A food chain approach to control of Shiga toxin-producing *Escherichia coli* in New Zealand

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

of Doctor of Philosophy in Veterinary Science

at Massey University, Palmerston North, New Zealand

Andrew Springer Browne

2018

ᵐEpiLab, Hopkirk Research Institute

School of Veterinary Science

Massey University

Palmerston North, New Zealand

Abstract

This thesis describes the prevalence and molecular epidemiology of Shiga toxin-producing *Escherichia coli* (STEC) in New Zealand using microbiological, genomic, molecular, and statistical methods. STEC are a zoonotic pathogen that can cause bloody diarrhoea and acute kidney failure. Cattle are a well-recognized STEC reservoir, and previous research has identified living near cattle and contact with their faeces as an increased risk for human infection. Seven STEC serogroups (O157, O26, O45, O103, O111, O121, O145), known as the 'Top 7' STEC, have been identified as an increased risk to human health, with the New Zealand meat industry undertaking testing to ensure that veal beef exports to some international markets are free of these 'Top 7' serogroups.

A random stratified cross-sectional study of 'Top 7' STEC prevalence of young dairy calves (n=1,508) on New Zealand dairy farms (n=102) found that approximately 20% of calves and 75% of farms were positive for one or more of the 'Top 7' STEC. 'Top 7' STEC prevalence was positively associated with increased number of calves in a calf pen, and prevalence significantly varied by region. This study utilized a new culture-independent diagnostic test, NeoSEEK (PCR/MALDI-TOF method), and used statistical and microbiological techniques to evaluate the sensitivity and specificity of the method for this and further studies.

A longitudinal study evaluating prevalence and transmission of 'Top 7' STEC in animals and the dairy farm environment found evidence of calf-to-calf, dam-to-calf, and environment-to-calf transmission. Whole genome sequencing analysis and

prevalence data revealed cross-contamination of young veal calf hides occurs during transport and lairage to processing plants.

Analysis of New Zealand serogroup O26 bacterial isolates (n=152), in comparison to publicly available genome sequence data (n=252) from other countries (n=14), suggested introduction of STEC and non-STEC O26 into New Zealand during few periods in the 20th and early 21st century. Populations of New Zealand serogroup O26 *E. coli* are monophyletic, possibly due to minimal live cattle importations into the country.

Further research in this area should focus on effective interventions at the farm and meat processing level to decrease the risk of veal beef contamination, while protecting public health.

List of Publications

1. Browne, AS, Midwinter, AC, Withers, H, Cookson, AL, Biggs, PJ, Marshall JC, Benschop, J, Hathaway, S, Haack, N, Akhter, R, and French, NP. Molecular epidemiology of Shiga toxin-producing *Escherichia coli* (STEC) on New Zealand dairy farms: application of a culture-independent assay and whole genome sequencing. Applied and Environmental Microbiology, accepted for publication.

2. Browne, AS, Midwinter, AC, Withers, H, Cookson, AL, Biggs, PJ, Marshall JC, Benschop, J, Hathaway, S, Hranac, R, Nisa, S, Rogers, L, Akhter, R, and French, NP. Evaluation of transmission dynamics and presence of Shiga toxin-producing *E. coli* (STEC) in animals and their environment on New Zealand dairy farms, and the potential impact on contamination of veal carcasses during slaughter and dressing. In preparation for Frontiers in Microbiology.

3. Browne, AS, Biggs, PJ, Cookson, AL, Wilkinson, D, Bloomfield, S, Midwinter, AC, Marshall, JC, Benschop, J, Rogers, L, Hranac, R, Withers, H, Hathaway, S, George, T, Jaros, P, Irshad, H, Fong, Y, Dufour, M, Kariki, N, Winkleman, T, and French, NP. A global genomic examination of Shiga-toxin producing *Escherichia coli* (STEC) serogroup O26 and non-toxigenic variants from multiple sources. Under review for Emerging Infectious Diseases.

List of Presentations

1. Browne, AS, Biggs, P, Cookson, A, Midwinter, A, Marshall, J, Benschop, J, Bloomfield, S, Wilkinson, D, Roger, L, Withers, H, Hathaway, S, George, T, Jaros, P, Irshad, H, and French, N. The local and global evolution and transmission of Shiga toxin-producing *E. coli* (STEC) serogroup O26. One Health Aotearoa Symposium, Wellington, New Zealand. December 14th, 2017. (Oral)

2. Browne, AS. Detection, prevalence, and transmission of STEC on dairy farms. Taranaki Veterinary Association. July 10th, 2017. (Oral)

3. Browne, AS, Biggs, P, Marshall, J, Cookson, A, Midwinter, A, Benschop, J, Withers, H, Hathaway, S, and French, N. Characterisation of *Escherichia coli* Serogroup O26 Isolates from New Zealand Cattle and Humans Compared to International O26 Isolates. Proceedings of the American Society of Microbiology Microbe Conference, New Orleans, USA, June 1-5, 2017. (Oral)

4. Browne, AS. Detection, prevalence, and transmission of STEC on dairy farms. Taranaki District Health Board. May 11th, 2017. (Oral)

5. Browne, AS. Detection, prevalence, and transmission of STEC on dairy farms. Starship Children's Hospital, Paediatric Nephrology Unit. March 28th, 2017. (Oral)

6. Browne, AS. Detection, prevalence, and transmission of STEC on dairy farms and how this translates to potential contamination of bobby veal. 2017 Meat

Industry Workshop at AgResearch, Ruakura, Wednesday 15 March 2017. Invited Speaker. (Oral)

7. Browne, AS, Biggs, P, Marshall, J, Cookson, A, Midwinter, A, Benschop, J, Withers, H, Hathaway, S, and French, N. Whole genome based comparison of *Escherichia coli* O26 serogroup isolates from New Zealand dairy calves. Proceedings from the 4th International One Health Conference, Melbourne, Australia, December 3-7, 2016. (Poster)

8. Browne, AS. A food chain approach to control STEC in NZ. Meat Industry Association Annual Technical Meeting, July 8th, 2016. Invited Speaker. (Oral)

9. Browne, AS. New molecular and genomic technologies to assess Shiga toxin-producing *E. coli* (STEC) on New Zealand dairy farms and meat product. Proceedings from the New Zealand Institute of Food Science and Technology Conference. July 4th-7th, 2016, Rotorua, New Zealand. Invited Speaker. (Oral)

10. Browne, AS, Midwinter, A, Withers, H, Cookson, A, Biggs, P, Marshall, J, Benschop, J, Hathaway, S, and French, N. Epidemiology of Shiga toxin-producing *E. coli* on New Zealand dairy farms using new molecular and genomic technologies. Proceedings of the New Zealand Veterinary Association Annual Conference, 21-24 June 2016, Hamilton, New Zealand. Invited Speaker. (Oral)

11. Browne, AS. The application of new molecular and genomic technologies to understand the epidemiology of the Top7 STEC in dairy cattle in New Zealand. STEC Workshop hosted by Dairy Companies Association of New Zealand, the

Meat Industry Association, Ministry of Primary Industries. May 11th, 2016. Invited Speaker. (Oral)

12. Browne, AS. The application of new molecular and genomic technologies to understand the epidemiology of STEC 7 in cattle in New Zealand. STEC Workshop hosted by AgResearch, Massey University, and Ministry of Primary Industries. December 15th, 2015. Invited Speaker. (Oral)

13. Browne, AS, Midwinter, A, Withers, H, Cookson, A, Biggs, P, Marshall, J, Benschop, J, Hathaway, S, and French, N. Prevalence, risk factors, and spatial distribution of Shiga-toxin producing *E. coli* (STEC) on dairy farms in New Zealand. Proceedings of the 9th Triennial International Symposium on Shiga Toxin (Verocytotoxin)- producing *Escherichia coli* (VTEC) meeting in Boston, September 13-16, 2015. (Poster)

14. Browne, AS. STEC on Dairy Farms. Meat Industry Association: STEC Workshop. Wellington, New Zealand. February 9th, 2015. Invited Speaker. (Oral)

Research Grants

1. Co-Investigator, Metabolic characteristics of *Escherichia coli* serogroup O145, Institute of Veterinary, Animal, and Biological Sciences, Massey University, (Amount Awarded: $0 to $5,000), Date Awarded: 04/2017

2. Co-Investigator, Evaluation of transmission of Shiga toxin-producing *E. coli* between dairy cattle and deer, Institute of Veterinary, Animal, and Biological Sciences, Massey University, (Amount Awarded: $0 to $5,000), Date Awarded: 04/2015

3. Co-Investigator, Genome evaluation of Shiga toxin-producing *E. coli* from deer and humans, Institute of Veterinary, Animal, and Biological Sciences, Massey University, (Amount Awarded: $0 to $5,000), Date Awarded: 11/2016

4. IVABS Postgraduate Conference Travel Fund, Institute of Veterinary, Animal, and Biological Sciences, Massey University, (Amount Awarded: $0 to $5,000), Date Awarded: 04/2017

5. IVABS Postgraduate Conference Travel Fund, Institute of Veterinary, Animal, and Biological Sciences, Massey University, (Amount Awarded: $0 to $5,000), Date Awarded: 04/2015

6. IVABS Postgraduate Conference Travel Fund, Institute of Veterinary, Animal, and Biological Sciences, Massey University, (Amount Awarded: $0 to $5,000), Date Awarded: 10/2016

7. ASM Student and Post Doctoral Award, American Society of Microbiology, (Amount Awarded: $0 to $5,000), Date Awarded: 05/2017

Acknowledgements

I can still remember getting off the bus outside Massey University, climbing up the hill to the Vet tower on a (rare) hot day in Palmerston North. After changing into some respectable yet crumpled clothes, I snuck into Kevin Stafford's office to discuss a possible PhD at Massey University. The former dean of my Veterinary School in Ireland, Boyd Jones, had put me in touch with Kevin to discuss potential PhD projects. I was wandering aimless, adventuring and doing some productive work, but needing a new direction. I interviewed with three people that day, one of whom was Nigel French. So first off, many thanks to Kevin and Boyd for helping me find this great PhD.

A huge thank you to my eight supervisors...the "cricket team" (as an American...I'm not sure how many people are actually on a cricket team). I owe a ton to Nigel French, who took a chance on me, and has been supportive over the past four years. Special thanks also to Steve Hathaway, for bringing in his vast expertise but also helping create my thesis project. A massive thank you to Anne Midwinter, who listened to my wingeing, always had time to hear my crazy ideas, and is an excellent supervisor who cares so much for her students. Many thanks to Adrian Cookson, who was always cheery, calm, and had another idea up his sleeve. Thanks to Patrick Biggs, who tolerated my blundering attempts at genomics and always keeps the objectives in view. Thank you to Jackie Benschop, who was always calm and thoughtful with her feedback. Thanks to Jonathan Marshall ("Mr. Wizard"), for helping me with his statistical mastery. And last but not least, many thanks to Helen Withers, who joined me in the field for bobby calf adventures and helped so much in the writing outputs for this project. Most of all, thank you to all my supervisors who have helped me find a love and appreciation for scientific research, which I'll keep for many years to come.

Table of Contents

List of Figures

List of Tables

Glossary and Abbreviations

Adulterant

A poisonous or deleterious substance on a carcass or meat product that can be injurious to human health

Allele

An alternate form of a gene that arises due to a fixed substitution in a nucleotide

Antibiotic

A medicine that inhibits the growth or destroys bacteria

Beef trim

Smaller pieces of beef muscle used in the production of ground beef products

Bobby calf

In New Zealand, a calf between the ages of four and ten days that is slaughtered for veal meat. The calf usually is born in a dairy herd, where the calf is surplus to requirements for replacement animals in the herd and is not viable for meat production.

CIDT

Culture independent diagnostic test; in comparison to methods where bacteria are isolated on nutrient agar

Clade

A group of descendants of a common evolutionary ancestor

Dam

The bovine mother of a calf

eae

intimin; a virulence gene that facilitates attachment of *E. coli* to the epithelial cells in the intestine

Enrichment broth

A nutrient broth that is mixed with bacteria and incubated at a specific temperature over a specific time in

	order to increase the number of bacteria present
HUS	Haemolytic uremic syndrome; a clinical presentation of haemolytic anemia (low red blood cell count due to destruction of red blood cells), acute kidney failure (anuria, lack of urine production), and thrombocytopenia (low platelet count); associated with severe clinical cases of STEC
MLST	Multilocus sequence typing; a method of differentiating organisms based on the variations (alleles) in seven housekeeping genes, in order to assign a sequence type (ST)
MPI	Ministry of Primary Industries; a public service department of New Zealand, in charge of overseeing, managing, and regulating the farming, food, and biosecurity sectors in New Zealand
PCR	Polymerase chain reaction; a molecular detection method where a pair of primers, sequences of DNA that are specific markers for a gene or number of genes, are amplified and detected in an agarose gel by the length of the sequence
PCR/MALDI-TOF	Polymerase chain reaction / Matrix Assisted Laser Desorption/Ionization – Time of Flight; a culture independent diagnostic test where a sample is ionized and then molecules

are detected using time of flight mass spectrometry, with specific molecular mass indicating specific targets for detection; this method is used by the NeoSEEK assay

PFGE   Pulse field gel electrophoresis; a DNA fragmentation technique to produce a "DNA fingerprint" of particular bacteria

Phylogenetic tree   A branching diagram to illustrate evolutionary relationships of organisms based on similarities or differences of genetic characteristics

Potential STEC   In this thesis, this refers to an enrichment sample that tests positive for a *stx* gene as well as the *eae* gene, but may or may not have an STEC bacterium (*stx* and *eae* present) present in the sample

Prebiotic   In animals, a non-digestible carbohydrate that promotes the growth of microorganisms in the intestines which may benefit health

Probiotic   A mixture of microorganisms that are ingested by animals that may promote intestinal health

R_o Basic reproduction number; in epidemiology, this refers to the number of cases of disease caused by one infective individual

RAMS   Recto-anal mucosal swab; a sterile cotton tipped swab is inserted into

	the rectum of a cow; this sample is then enriched in liquid media to increase detection of STEC
RT-PCR	Real time polymerase chain reaction; similar to PCR where a specific DNA sequence between primers is amplified, but a colour based probe reacts to binding in the region and is detected by a machine, leading to real time recognition of the amplification of the DNA sequence
SNP	Single nucleotide polymorphism; Single nucleotide differences between genes that are shared between organisms
Spring calving season	For dairy farms in New Zealand, this usually begins in late June to early July, and ends in September to October. Dairy farming in New Zealand typically follows an annual cycle, although some farms may allow for an Autumn calving season.
ST	Sequence type; a number assigned through the MLST method to differentiate groups of bacteria
STEC	Shiga toxin-producing *Escherichia coli*; *E. coli* bacteria that contain the *stx* gene and therefore may be able to produce Shiga toxin; also called verocytotoxigenic *E. coli* (VTEC), due to its pathogenicity to vero (kidney) cells
Strain	A genetic variant of an organism

stx	A virulence gene that leads to the production of Shiga toxin
'Top 7' STEC	The seven O serogroups (O157, O26, O45, O103, O111, O121, O145) of STEC declared adulterants of beef by the USDA-FSIS, and recognized as a significant risk to human health
UK	United Kingdom
USA	United States of America
USDA-FSIS	United States Department of Agriculture–Food Safety and Inspection Service; in charge of protecting public health by ensuring the safety of meat, poultry, and processed egg products in the USA
Zoonoses	Pathogens (bacterial, viral, fungal, prion) that are transmissible between animals and humans

Nel mezzo del cammin di nostra vita

mi ritrovai per una selva oscura,

ché la diritta via era smarrita.

Canto I, *The Inferno*, Dante Alighieri

There is something to be learned from a rainstorm.
When meeting with a sudden shower, you try not to get wet and
run quickly along the road. But doing such things as passing
under the eaves of houses, you still get wet. When you are
resolved from the beginning, you will not be perplexed, though
you will still get the same soaking.
This understanding extends to everything.
The Hagakure, Tsunetomo Yamamoto

TONIGHT AT THE MAGICAL THEATER
FOR MADMEN ONLY
PRICE OF ADMITTANCE YOUR MIND
Steppenwolf, Herman Hesse

1 Introduction to thesis

1.1 Background

Worldwide, Shiga toxin-producing *Escherichia coli* (STEC) are a growing public health concern. Large scale outbreaks in Europe (1) and the United States (2) continue to occur, while human STEC cases in Argentina have a high rate of serious clinical complications (3). Although STEC may have a lower incidence than other notifiable zoonotic diseases (4), the pathogen's propensity to affect very young children, leading to potential long term kidney and brain damage (5), is a concerning public health issue. STEC are primarily transmitted via the faecal-oral route, and ruminant animals, particularly cattle, have been identified as the most important reservoir (6).

1.2 STEC in New Zealand

4.3.5 Human disease

New Zealand is a geographically isolated island nation that offers a unique opportunity to interpret the effects of importation and biosecurity measures on control and transmission of zoonotic diseases (7). New Zealand has a relatively high incidence of notified STEC infection in humans, with 8.9 STEC cases per 100,000 population reported in 2016 (8), compared to 2.85 in the USA in 2016 (9), and several incidence rates in 2015: 12.92 in Ireland, 5.08 in the Netherlands, and 2.05 in the United Kingdom (10). Since it became a notifiable disease in New Zealand there has been a general increase in annual notifications, with both STEC O157:H7 and non-O157:H7 STEC causing human disease (11). A New Zealand case-control study identified contact with animal manure and the presence of cattle in the local area, along with contact with recreational waters, as significant risk factors for human STEC

infection (12). This study did not identify food as a statistically significant exposure pathway in New Zealand (12).

4.3.6 Implications for trade and the New Zealand beef industry

Having found STEC in raw ground beef, plus outbreaks associated with consumption of undercooked beef patties, the USA declared STEC O157 an adulterant of beef in 1994, followed by declaration of 6 additional serogroups (O26, O45, O103, O111, O121, O145) as adulterants in late 2011 (13). Mandated testing of the additional six STEC serogroups began on March 5th, 2012 (13). These seven serogroups are commonly known as the 'Top 7' STEC. In 2016-2017, 50% of New Zealand beef exports, valued at $1.16 billion NZD, were sent to the United States (14). Given the importance of agricultural exports for the New Zealand economy, STEC is both an economic and public health concern.

The Ministry of Primary Industries (MPI) and the Strategic Directions Group (SDG) of the Meat Industry Association (MIA) have a joint work programme for STEC. Key objectives of the programme include:

- Validation of practical and cost-efficient laboratory methods for detection and confirmation of New Zealand STEC isolates

- A high level of understanding of the epidemiology of STEC in animal populations as it impacts on the level of contamination on fresh meat carcasses

- Better knowledge on the risk factors and interventions that can be applied at farm level and pre-slaughter to significantly reduce the level of contamination on beef

1.3   Goals of thesis research

Overall, the goals of this thesis are to identify risk factors for STEC carriage at the farm and animal level, increase understanding of STEC transmission dynamics on New Zealand dairy farms, and improve detection methods for New Zealand STEC. These goals are achieved through the use of advanced microbiological, molecular, and statistical techniques. The information gained identifies possible interventions at the farm level and helps protect public health and inform the New Zealand beef industry.

1.4   Research Questions

The questions for each research chapter are outlined below:

4.3.7   Chapter 3: Molecular epidemiology of STEC on New Zealand dairy farms: application of a culture-independent assay and whole genome sequencing

- What proportions of dairy farms in New Zealand are carrying the 'Top 7' STEC?

- What proportions of calves are shedding these organisms?

- What is the spatial distribution of STEC positive farms in New Zealand? Is there evidence of localised farm to farm transmission of particular genotypes?

- What are the risk factors for carriage of STEC and can they be targeted for control?

- Is the PCR/MALDI-TOF method NeoSEEK suitable for the detection and characterisation of STEC in cattle faeces and beef trim?

- What is the sensitivity and specificity of the NeoSEEK assay for the detection of 'Top 7' STEC serogroups in New Zealand cattle?

4.3.8 Chapter 4: Evaluation of transmission dynamics and presence of Shiga toxin-producing *E. coli* (STEC) in animals and their environment on New Zealand dairy farms, and the potential impact on contamination of veal carcasses during slaughter and dressing

- What are the broad dynamics of STEC infection in dairy farms during the calving season?

- Are animal, group and herd-level infections single clones, or the result of dynamic transmission and evolution of multiple strains?

- What are the longitudinal trends in slaughter populations, especially very young calves, and eventual contamination rates of veal carcasses?

4.3.9 Chapter 5: A global genomic examination of STEC serogroup O26 and non-toxigenic variants from multiple sources

- What is the genetic diversity and population structure of STEC serogroup O26 in New Zealand cattle and humans, and how rapidly are they evolving?

- When was STEC O26 imported into New Zealand?

- How do the virulence genes and antibiotic resistance traits of STEC O26 in New Zealand compare to STEC O26 isolated in other countries?

1.5 Structure of Thesis

This PhD thesis is composed of a literature review in Chapter 2 of relevant facets of STEC that relate to this research, followed by three research chapters (

Figure 1-1). In **Chapter 3**, the national prevalence of STEC in young dairy calves is estimated, while identifying risk factors at the animal, environmental, and farm management level that could be used as targets for future interventions. This chapter also involves the assessment of a new molecular detection method for STEC (NeoSEEK) using advanced statistical methods. In **Chapter 4**, the transmission dynamics of STEC carriage in cattle and the environment on New Zealand dairy farms

is examined, with follow-through to the carriage and contamination of veal carcasses in processing plants. In **Chapter 5**, whole genome sequencing of *E. coli* serogroup O26 isolated from humans and cattle in New Zealand is performed and data compared to publicly available sequence data obtained from *E. coli* serogroup O26 isolated from around the world. Sequence data is used to estimate the most recent common ancestor of STEC O26 in New Zealand, as well as the virulence and resistance profiles of New Zealand bacterial isolates. **Chapter 6** includes a holistic discussion of the research chapters and recommendations for future directions for STEC research in New Zealand. Overall, this thesis adds comprehensive understanding of the prevalence and transmission dynamics of STEC in New Zealand dairy farms, while identifying future research projects to develop novel interventions that reduce risk for the meat industry as well as protect public health.

Figure 1-1: Structure of PhD thesis: A food chain approach to control of Shiga toxin-producing *E. coli* in New Zealand

As a writer, you should not judge.

You should understand.

Monologue to the Maestro: A High Seas Letter,

Ernest Hemingway

2 Literature Review

This literature review establishes the background for the following thesis research, by providing an overview of the prevalence, transmission, and molecular epidemiology of Shiga toxin-producing *Escherichia coli* (STEC) around the world. Due to the vast amount of research on STEC, with over 10,000 articles identified on the PubMed database following a search for "STEC" (www.pubmed.gov, accessed February 1st, 2018), this review is focused on specific facets of STEC research that are directly applicable to this thesis. After each main topic is evaluated, a section highlights how the literature is relevant to the following research chapters.

2.1 Introduction

The first Shiga toxin-producing *Escherichia coli* (STEC) associated with severe illness was linked to an outbreak of bloody diarrhoea in 47 people, associated with beef served by a fast-food retail chain in the United States in 1982 (15). The causative bacterium was identified as *E. coli* O157:H7, based on the lipopolysaccharide antigen (O157) and flagella antigen (H7). This serotype had been previously identified in a case of haemorrhagic colitis in 1975, but was not associated with being invasive or toxigenic by standard tests at that time (15). Ten years later, STEC O157:H7 was linked to an outbreak due to consumption of undercooked beef patties in the Jack-in-the-Box fast-food retail chain; this outbreak led to 732 people being infected, and subsequently 4 deaths and 178 people with permanent disability due to kidney or brain damage (16). The severity of this outbreak led to global recognition of the importance of STEC in food safety and public health, and STEC O157:H7 was classified as an adulterant of

meat in the United States in 1994. In New Zealand, STEC O157 infection became a notifiable disease in 1997 (17).

Following a study which identified the most common non-O157 STEC causing human disease in the United States (18), the United States Department of Agriculture Food Safety and Inspection Service (USDA-FSIS) classified six additional *E. coli* serogroups, O26, O103, O45, O111, O121, and O145, as adulterants of raw beef on September 13th, 2011. This declaration increased the scope of STEC detection from one to seven serogroups, known as the 'Top 7' STEC, leading to the need for greater and more sophisticated methods of detection and isolation of these pathogenic bacteria.

Understanding STEC and the complex interrelationship between the farm environment and its main reservoir, cattle, offers the chance to address this pathogen at its source to protect public health and the vital agricultural economic trade.

2.2   Shiga toxin-producing *E. coli*: the bacterial pathogen

4.3.10   Taxonomy and molecular typing

Escherichia coli are gram-negative facultative anaerobic bacteria, belonging to the gamma subdivision of the Proteobacteria phylum (19). A myriad of classifications of *E. coli* exist based on pathogenicity (e.g. commensal or pathogenic), clinical manifestations in humans (e.g. uropathogenic, extra-intestinal pathogenic), virulence mechanisms (e.g. enteroinvasive, enterohaemorrhagic), serology (e.g. O157:H7*, K12)*, and multilocus sequence typing (MLST) (e.g. ST-10, ST-29) (19). MLST in *E. coli* is based on the alleles of seven housekeeping genes: *adk, fumC, gyrB, icd, mdh, purA*, and *recA* (20). According to the modified Kauffman scheme of serotyping, O (somatic), H (flagellar), and K (capsular) surface antigens are used to differentiate between various *E. coli* (21). While these surface markers do not in themselves confer virulence, certain

combinations of O and H antigens are associated with particularly virulent groups of *E. coli* (21).

Even greater molecular and genetic resolution of *E. coli* can be attained through the use of techniques such as pulsed-field gel electrophoresis (PFGE) and whole genome sequencing (WGS). During PFGE, bacterial genomic DNA is first cleaved with restriction endonucleases with fractionation of high molecular weight DNA molecules as they pass through a gel medium under electric current, leading to unique patterns of DNA separation that can be attributed to specific strains of bacteria (22). Whole genome shotgun sequencing collects data on an entire DNA sequence of an organism by breaking up DNA in small random fragments, amplifying those fragments and detecting the DNA bases to create reads, and then assembling these reads to infer a complete DNA sequence. These typing methodologies were historically used to deduce small variations between bacterial strains, but the advent of universal databases and a global system of interconnected computers has allowed comparisons to be made from large datasets across the world.

4.3.11 Bacterial gene transfer

Horizontal gene transfer is the transfer of genetic material between bacteria of the same generation (23). This method of gene transfer is important in STEC research, as most of the genes that cause pathogenicity in STEC are transferred via this method.

Mobile genetic elements are genes that insert in various locations in the host genome. Numerous mobile genetic elements, comprising 9-20% of the genome, were found in STEC and non-STEC O26, indicating a complex history and variety of recombination (24). While some mobile genetic elements, such as those that express antibiotic

resistance, are acquired through cell to cell contact (conjugation), bacteriophage transfer of genetic material (transduction) is a more important method of genetic transfer for STEC (24).

A bacteriophage is a virus of bacteria. These bacteriophage can attach to a bacterium and insert a DNA sequence, known as a prophage, directly into the host bacterium genome or as a plasmid, a small circle of DNA that contains a small number of genes and an origin of replication (25). These prophage can include pathogenicity islands, which include large clusters of virulence genes (26). Acquiring new genes can increase fitness and survival of a bacterial pathogen (27).

4.3.12 Virulence factors

The *stx*-toxins expressed by STEC are also known as Shiga-like toxins, due to their similarity to the toxin produced by *Shigella dysenteriae*, and were discovered by Dr. Kioshi Shiga in 1898 (28). In fact, phylogenetic evidence using 16S rRNA gene sequence alignments notes that *Shigella* spp. could be classified within *E. coli*, and they are as related to *E.coli* as they are to each other (19). Two Shiga toxins are observed in STEC: Shiga toxin 1 and Shiga toxin 2, produced by the *stx1* gene and *stx2* gene respectively (19). The genes for Shiga toxin 1 and 2 only share 60% nucleotide identity, which indicates that these genes underwent separate evolutionary pathways, and were only recently acquired by *E. coli*. In fact, *stx2* is found in other members of the Enterobacteriaceae family, indicating that other bacterial species have acquired this gene through horizontal gene transfer (29). Phage-mediated *stx* transduction can occur in food and water, but fails to occur with low pH or low temperatures (30). It has been suggested that STEC are more resistant to consumption by protozoa due to the presence of Shiga toxin, which offers a survival advantage to the bacteria (31).

STEC are defined by the presence of the *stx* virulence gene, but commonly carry the *eae* gene (intimin) as well, in both international food safety regulations and in this thesis. The *eae* gene, which is an abbreviation for "*E. coli* attaching and effacing", facilitates attachment of the bacterium to the intestinal wall leading to enterocyte membrane and effacement of the microvilli. The gene lies within the Locus of Enterocyte Effacement pathogenicity island (LEE); this island is a distinct region of DNA that confers pathogenicity to STEC.

4.3.13 Antimicrobial resistance factors

Resistance genes can be transferred via plasmids or mobile genetic elements via integrons. Integrons are genetic mechanisms that allow for the acquisition and dissemination of antibiotic resistance genes in antibiotic resistant bacteria through genetic elements that allow absorption of these elements using genes like integrase (32). In most cases, resistance is conferred to the bacterium by a single gene (e.g. *bla*$_{TEM}$ confers resistance to the antibiotic ampicillin).

A comprehensive study of over two hundred *E. coli* strains containing *eae* isolated from various livestock species found that over 65% of the strains were resistant to tetracycline, streptomycin, erythromycin, and sulfamethoxazole, with resistance to ampicillin, chloramphenicol, trimethoprim, and trimethoprim-sulfamethoxazole being less common (approximately 30%) (33). The predominant resistance genes for each resistant strain of bacteria were ampicillin (*blaTEM*, 97.1%), tetracycline (*tetA*, 76.7%), gentamicin (*aac(3)II*, 80%), streptomycin (*strA/strB*, 76.7%, *aadA*, 76.7%), chloramphenicol (*catI*, 85.1%), trimethoprim (*dhfrI*, 76.3%), and sulfamethoxazole (*sul1*, 60%; *sul2* (63.3%).

4.3.14 Pathogenesis

Shiga toxins cause disease by direct (inhibition of protein synthesis) and indirect (proinflammatory cytokine expression) methods (34). In humans, Shiga toxin inhibits protein synthesis and immune response locally in the gastrointestinal tract, leading to haemorrhagic diarrhoea. Once absorbed through the intestinal epithelium, they preferentially act on endothelial cell subtypes found in the brain and kidney (34). Absorption of the toxin into the bloodstream leads to haemolytic uremic syndrome (HUS), due to the toxin damaging renal endothelium and widespread inflammation and destruction of erythrocytes, causing acute renal failure. This complication is more common in young children, who express higher levels of Shiga toxin receptors in their renal glomeruli (35). Shiga toxin 2 is associated with more severe clinical outcomes than Shiga toxin 1, and *in vitro* assays estimate Shiga toxin 2 is 40 to 400 times more potent than Shiga toxin 1 (36). The incubation period varies from 2-10 days, and 5-15% of people who ingest the bacterium will succumb to HUS, which can be fatal (37).

STEC are apathogenic to ruminant animals, particularly cattle, who are identified as an important reservoir of STEC (6). Cattle lack vascular receptors for Shiga toxin and it has no detrimental effects (enterotoxicity) in their digestive tract (38). The primary site of colonisation of STEC O157:H7 in cattle is reported to be at the distal gastrointestinal tract (e.g. colon) (39). Once *E. coli* with *stx* and *eae* colonise a calf's digestive tract, they are more likely to persist in the digestive tract compared to other non-STEC *E. coli* flora (40).

4.3.15 Origin of STEC

Phylogenetic evidence indicates STEC O157 may have originated from a non-*stx* producing *E. coli* serotype O55:H7, which acquired the *stx2* gene and the O157 plasmid

(29). Other studies have suggested, through the analysis of housekeeping genes, that multiple lineages of *E. coli* acquired the LEE pathogenicity island with *eae*, plasmid-borne enterohaemolysin (*ehxA*), and the prophage bearing *stx*; this finding is a classic example of parallel evolution (41). For instance, an evaluation of *ehxA* encoding virulence plasmids found that a particular subtype (subtype D) had an evolutionary lineage completely distinct from the other *ehxA* encoding plasmids, indicating a separate instance of evolution (42). It is reasonable to suspect that the acquisition of various virulence genes may lend a competitive advantage for these bacterial strains against other bacteria or protozoa.

4.3.16 Evolution of STEC

With the potential for horizontal transfer of virulence genes, STEC continues to evolve. Evolution based on recombination and horizontal gene transfer is explained by the Public Goods Hypothesis for the evolution of life on earth (43). In contrast to a Tree of Life evolutionary model (44), the Public Goods Hypothesis proposes that horizontal exchange of widely available DNA sequences, notably genes, are the primary driver for local bacterial evolution (43). Applying this hypothesis to the evolution of STEC would allow for continued adaptation of STEC in local geographical environments as they continue to acquire and share virulence genes, which could lead to changes in the pathogenicity of STEC. Molecular typing of STEC O157 from New Zealand, Australia, and the USA indicated a unique differentiation of genotype distribution for each country (45). Molecular analysis of STEC O157 in New Zealand revealed patterns of local spatial clustering of genotypes (45).

Genomic analyses can evaluate the number of shared genes between *E. coli* (core genome) and genes not common between bacteria (accessory genome). Analysis of 53

E. coli genomes from various serogroups estimated a core genome of 1,472 and an accessory genome of 13,296 (19). Remarkably, this would indicate that only 20% of genes were shared between the *E. coli* in this study, while the rest of the genes were not found in all genomes. *E. coli* are capable of remarkable genetic diversity.

One example of ongoing STEC evolution is the emergence of *E. coli* serogroup O26 ST-29 in Europe. Progenitor STEC O26 strains harboured the *stx1* gene and were classified as ST-21. Beginning in the late 1990's, a new STEC O26 variant, ST-29, with the more virulent *stx2* gene emerged in Germany (46), and more recently in France (47) and Switzerland (48). This strain has enhanced pathogenicity and is associated with increased incidence of HUS even in healthy adults. Controlled experiments revealed that *stx2* can be acquired and lost in serogroup O26 isolates, leading to dynamic changes in pathogenicity (49). The emergence of a highly virulent STEC O26 ST-29 clade has been identified in Japan since 2012, and single nucleotide polymorphism (SNP) phylogenetic analysis revealed it was a separate clade from the European variant, indicating another example of an acquisition of *stx2* in this serogroup (50).

Another example of recent STEC evolution was the emergence of a highly virulent O104:H4 in Germany in 2011 (51). The previously avirulent strain acquired the prophage coding for *stx2* gene, along with an extended beta-lactamase antimicrobial resistance factor (CTX-M-15), through horizontal gene transfer (51). The bacterial strain also caused severe disease without the presence of the LEE pathogenicity island, or *eae*. 810 cases and 39 deaths were attributed to the STEC O104:H4 outbreak in Germany (1). The majority of HUS (90%) cases occurred in adults, rather than children.

Objectives of this thesis include evaluation of the genetic diversity, virulence, antimicrobial resistance, and population structure of STEC. Whole genome sequencing is employed to understand differences on the farm and region level in a national cross-sectional study in **Chapter 3**, as well as to understand transmission dynamics from the farm environment to meat processing plants in **Chapter 4**. A more holistic approach is pursued in **Chapter 5**, where STEC O26 from New Zealand humans and cattle are thoroughly evaluated against bacterial isolates around the world, in order to understand the arrival of STEC into the country, as well as its evolution since its introduction. These components of each chapter inform the fundamental nature of STEC in New Zealand on a cellular level.

2.3   Methods and difficulties of detection and isolation of STEC

The detection and isolation of a single bacterium within a complicated matrix of micro-organisms and organic material is difficult. *E. coli* may be grown on MacConkey agar at 37°C, which takes advantage of their salt tolerance as well as lactose fermentation to differentiate them from other Enterobacteriaceae. However, both due to the severe virulence associated with some STEC serogroups, and their difficulty of isolation, many methods for the detection and isolation of STEC have been developed. Detection methods are primarily focused on food items and beef trim due to the regulatory nature of the international trade of beef, and public health concern regarding food-associated outbreaks.

4.3.18 Overview of USDA-FSIS detection and isolation of STEC

The current well recognized and legally required screening and isolation for the 'Top 7' STEC in meat products in the USA is dictated in the USDA Microbiology Laboratory

Guidebook (MLG) 5A.04 (52) for STEC O157 and MLG 5B.05 (53) for the other 'Top 7' STEC. The testing of ground veal meat is also described in the Market Access Requirements documentation (54). The process, following retrieval of the sample material, is labour intensive and requires: enrichment of the sample in a liquid nutrient broth, molecular screening for virulence genes (*stx* and *eae*) and the 'Top 7' serogroups (O157, O26, O45, O103, O111, O121, O145), an immunomagnetic separation (IMS) procedure with antibody coated beads, isolation of bacterial colonies on culture media, and confirmation of identity of bacterial isolates using latex agglutination and molecular assays. At a bare minimum the process requires four days. This method has also been verified in Canada as leading to detection of one or two viable STEC cells per 65 g of beef trim (55).

4.3.19 Sample retrieval

Recto-anal mucosal swabs (RAMS) are ideal for sampling of STEC in cattle due to the propensity for STEC bacteria to colonise the recto-anal mucosal junction. RAMS have been found to be a better indicator of colonisation of STEC O157:H7 than faecal culture (56). RAMS is also superior to faecal samples for detection of the 'Top 7' STEC using a PCR/MALDI-TOF method, NeoSEEK, which will be discussed in detail in a later section (57).

Sterile cellulose sponges can be used to collect environmental and calf hide samples (58). Other methods of collecting faecal material include overboot swab sampling. In Sweden, overboot swab sampling and dust sampling of the environment correctly classified over 80% of STEC O157:H7 positive herds that were confirmed to be positive using individual RAMS sampling (59). Overboot swab sampling was also successfully used in research on *Mycobacterium avian* subsp. *paratuberculosis*, where over 90% of

herds that had cows that were individually confirmed as positive were also found to be positive via environmental overboot swab sampling (60).

4.3.20 Common DNA-based culture-independent diagnostic tests (CIDTs)

In order to decrease the amount of time and resources used in searching for STEC bacterial isolates, DNA can be prepared from a microbial sample through boiling or other chemical methods, and specific DNA sequences, notably target genes (e.g. *stx* or *eae*), can be detected using various methods such as PCR. It is important to reach a low detection limit as most adult cattle excrete less than 100 CFU/g of STEC (61). The most common culture-independent diagnostic tests (CIDTs) used for STEC rely on the polymerase chain reaction (PCR), or its more modern variant, real time PCR (RT-PCR).

Early molecular screening involved the use of PCR. PCR allows for massive duplication of specific target DNA regions through the use of oligonucleotide primers (forward and reverse) that are sensitive and specific to allow amplification of the target gene. Multiple primer targets can be duplicated in the same sample at the same time, but the lengths of each primer pair product have to be different in order to visualize on agarose gels, after separation using electrophoresis. One of the most widely used PCR panels was developed in the late 1990's for the screening of *stx1, stx2, eae, hlxA* (*ehxA*), and the O157 surface antigen (62). Multiplex PCR panels, screening for all the 'Top 7' STEC serogroups, have also been developed (63).

Real time PCR (RT-PCR) uses traditional primers with an additional component of a fluorescently-labelled oligonucleotide probe; this probe is a DNA sequence that lies between the two primers of the target gene, and also has a colour-based fluorophore

molecule that is released when the probe binds to a DNA target. A device that heats and cools the samples also has a specific light detector to pick up the intensity of light at a specific wavelength, indicating the amount of DNA that has bound, and therefore light released, with the probe. Advanced RT-PCR devices allow for the detection of four different wavelength ranges, therefore allowing for detection of four different gene targets in the same sample. The amplitude of light emission can be compared to a standard curve of DNA concentrations, leading to inferences on DNA quantity in the original sample (quantitative RT-PCR, qPCR). The minimum information for publication of quantitative real-time PCR experiments (MIQE guidelines) was established to offer consensus and reproducibility among researchers using RT-PCR (64). Early RT-PCR panels for *E. coli* could identify over 22 virulence genes (65). A sensitive and 100% specific multiplex RT-PCR protocol was developed that allowed for screening of virulence genes (*stx1*, *stx2*, and *eae*), as well as all 'Top 7' serogroups in three runs (66).

The inherent difficulty with molecular screening is that a vast microbial matrix may contain a bacterium with an *stx* gene, a different bacterium with an *eae* gene, and a third bacterium that possess gene sequences specific for a 'Top 7' STEC O-antigen serogroup, but PCR-based methods cannot differentiate whether a 'Top 7' STEC is present in the sample. However, the sensitivity of this assay is useful as a screening procedure before more specific methods are used.

4.3.21 Enrichment

Enrichment involves the microbiological technique of putting a sample into a nutrient liquid and incubating it for a period of time and at a particular temperature, in order to allow multiplication of the desired target bacteria and increase the chances of its

detection. Several different enrichment broths have been used for STEC. Universal

pre-enrichment broth (UPB) and modified Tryptone Soya Broth (mTSB) were found to

be more effective than *E. coli* broth in enrichment of non-O157 STEC (67). Some

methods of enrichment of STEC include the addition of novobiocin, an antibiotic, in

order to decrease the proliferation of background flora in the sample. However,

novobiocin added to TSB has been found to have an inhibitory effect on non-O157

STEC, while STEC O157:H7 was significantly more resistant to its effects (68). Another

study evaluating the addition of novobiocin to *E. coli* Broth (EB) and mTSB found it

should not be used in the enrichment of non-O157 bacteria in samples (67) .

Therefore, the addition of novobiocin to TSB would likely lead to false-negative results

for detection of non-O157 STEC. Originally, the USDA-FSIS protocols (MLG 5B.01)

recommend the use of 8 mg/L novobiocin to modified TSB broth for enrichment of

the 'Top 7' STEC, but this addition has been removed in the new protocol (MLG

5B.05).

Other factors have been studied regarding time and temperature of incubation. An

evaluation of enrichment times found no significant difference between enrichment

times of 6 and 24 hours (69). Increasing the incubation temperature to 42°C was

found to significantly increase recovery of STEC O157:H7 in comparison with lower

temperatures traditionally used for bacteria (70). Furthermore, this temperature (42°C

+/- 1) is required by the USDA-FSIS protocol MLG 5B.01 for isolation of the 'Top 7'

STEC.

4.3.22 Culture

Culture-based approaches to detection of STEC have disadvantages, due to the lack of

specificity of molecular methods that screen bacterial enrichments for the presence of

the 'Top 7' STEC defining virulence genes (*stx1*, *stx2*, and *eae*) and the 'Top 7'

serogroups, and the low sensitivity of culture-based isolation methods (71). However,

bacterial isolation is essential for phenotypic assays and whole genome sequencing of

individual bacteria.

Several different agars are used to isolate STEC. Sorbitol MacConkey agar

supplemented with cefixime and potassium tellurite (CT-SMAC) is used for the

isolation of STEC O157. Most STEC O157 are unable to ferment D-sorbitol and grow as

grey colonies on CT-SMAC (72), but some different lineages of STEC O157 have been

identified that do ferment sorbitol and grow as purple colonies (73). CT-RMAC is a

variation of CT-SMAC that includes the substitution of rhamnose for lactose (74).

STEC O26 do not usually ferment rhamnose (74), but some rhamnose fermenting

STEC O26 have also been identified in New Zealand (T. George, unpublished data).

Possé agar was designed to use chromogenic compounds to indicate β-galactosidase

activity and particular fermentative carbon sources (75). However, when the media

was tested it was found to be effective for detection of STEC in food material, but

could only isolate 15.8% of STEC in artificially contaminated cow faeces (76).

CHROMagar STEC medium (CHROMagar Microbiology, Paris, France) uses

chromogenic compounds to identify STEC compared to non-virulent *E. coli* bacteria. A

comprehensive assessment of STEC isolates (n=362) on this agar found that 81.6% of

the isolates were capable of growing after direct plating onto the media, but two of the

'Top 7' STEC serogroups, O157 and O103, grew poorly (33-39%) (77). The incorporation

of tellurite in the media was directly linked to the finding that only isolates with the

terD gene, which confers tellurite resistance, were able to grow on the medium.

Another study found that CHROMagar STEC inhibited the growth of 27/96 (28.1%) of

'Top 7' STEC strains tested (78). These findings suggest that CHROMagar STEC is selective for STEC but may be overly inhibitive of growth of all 'Top 7' STEC. The USDA recommended formulation for Rainbow Agar was found to inhibit growth of 20/96 examined STEC strains (78). This study also noticed a large range of colony colours of the STEC strains, despite recommendations that certain colony colours indicate whether a colony is STEC or non-STEC (78). A comprehensive study of several different methods of three enrichment broths and four agars found that the standard MacConkey agar was the most sensitive (79). However, MacConkey agar does not allow the differentiation between the 'Top 7' STEC and other *E. coli*, which would necessitate additional testing of all colonies on a plate; this would be an expensive and time-consuming process.

Culture often involves the use of immuno-magnetic separation (IMS) using beads coated with a serogroup specific antibody which binds to the target O-antigen of the bacteria in order to increase sensitivity. A comparison of direct plating after enrichment and using IMS beads before plating found that only the O157 serogroup benefited from IMS, and IMS did not have any beneficial effect on isolation of the four other 'Top 7' STEC serogroups evaluated (O26, O103, O111, and O145) (69). Another study of the non-O157 'Top 7' STEC showed that experimentally spiked faecal samples had varying success rates of retrieval using IMS: 92.3% for O26 and 100% for O103, in comparison to 38.5% for O121 and 43.8% for O111 (80). For faecal samples that tested positive for these serogroups in a field trial, the retrieval was much lower: 21-25% for O26 and O45, and 6-10% for O145, O121, O111, and O103 (80). Another study found that bacterial isolation procedures identified 33.1% (53/160) and 35.5% (11/31) of PCR-positive samples for *E. coli* O157 and non-O157 serogroups, respectively (63). Further

research has found that despite a PCR prevalence of O157, O26, O45, and O26 exceeding 78%, only 4-10% of the samples had STEC successfully isolated (81). USDA-FSIS methods were used to detect 44.5% prevalence of a non-O157 'Top 7' serogroup, *eae*, and *stx* in samples from Australian cattle processing plants, while only 1.3% were confirmed by isolation (82). Though prevalence ranged from 5.7% to 33.5%, no O45, O103, O121, or O145 were isolated in the study (82). In a previous ^{m}EpiLab study, 10 of 55 samples where serogroup O157 was detected had STEC O157 isolated (83). It is possible that non-viable bacteria are being detected with molecular methods, therefore culture isolation will not be successful. However, the disparity between traditional molecular detection and bacterial isolation of STEC can make estimates of prevalence difficult.

4.3.23 Other culture-independent diagnostic tests (CIDT)

Other methods such as enzyme-linked immunosorbent assay (ELISA) have been used for STEC serogroup detection in food and meat trim, however these require a high pathogen load (10^5 CFU/ml) and do not allow for identification of the pathogenicity of that bacteria within the serogroup (84).

While whole genome sequencing is typically used to provide a full molecular evaluation of a single bacterium post-isolation, metagenomic sequencing has been utilised as a screening method. Metagenomic sequencing involves purification and sequencing of DNA or RNA from a mixed sample without selective amplification, leading to the identification of a variety of organisms, which can then be classified according to various marker genes. This method was employed to evaluate STEC in artificially spiked spinach, and found it was able to detect 10 CFU/100 g of produce

(85). While this method is experimental and may not be cost-effective at this time, it may become widely used in the future.

One molecular method of particular value is NeoSEEK (NeoSEEK STEC Confirmation, NeoGen Corporation, Lansing, MI, USA). The NeoSEEK method uses PCR/MALDI-TOF; this method involves PCR amplification, primer extension to generate allele-specific DNA products of different masses, and chip-based mass spectrometry to analyse the extension products in order to detect pathogenic STEC. This method has been employed for use in identifying STEC in outbreaks (86), as well as identifying specific sequence types (87). NeoSEEK has a Letter of No Objection (USDA-FSIS), indicating that the USDA found this technology to be reliable and appropriate for use as a STEC confirmation method. NeoSEEK uses 89 gene targets, as well as virulence markers and *eae* subtypes, to detect the presence of STEC from an enrichment broth, without the need for individual plating of bacteria. The technology is based on initial findings that single nucleotide polymorphisms (SNPs) in the O-antigen gene cluster could identify both which serogroup the bacterium belongs to, and whether that bacterium contains *stx* (88). This finding was ground breaking, in that it bypassed the predominant problems with molecular screening for virulence genes and O serogroups: a failure to effectively isolate bacteria from molecular screen-positive enrichments to confirm the presence of STEC bacteria. Possible disadvantages of NeoSEEK include the limitation of its database of relevant SNPs, which it uses to compare true STEC to non-toxigenic strains. However, this database can be updated with the continued evaluation of new whole genome sequenced bacterial isolates from STEC.

4.3.24 Relevance

The use of culture-independent diagnostic testing in the place of traditional microbiology is controversial (89). In general, CIDTs offer better sensitivity than culture methods, thus reducing the underestimation of disease burden (90). These tests also have the benefit of being faster and less expensive, in particular with regards to labour costs (90). However, an argument was made in the United States that clinical and local laboratories need to continue to perform microbiological culture of isolates rather than only use CIDT, which transfers the entire burden of culture methods to the state and federal level (89). The utilisation of CIDTs to detect bacterial zoonoses is rapidly increasing, and STEC is diagnosed via these methods more than any other pathogen (91). Furthermore, the utilisation of CIDTs had led to increased detection via culture-confirmed STEC in the United States, perhaps due to the ease of testing or the increased sensitivity of the methods (9). The economic and employment effects of the adoption of CIDTs should not be overlooked.

One extensive study that evaluated four enrichment broths and three agars found that no enrichment protocol would work for all artificially inoculated ground beef samples, and significant variation occurred with naturally contaminated samples (79). In addition, neither one nor a combination of agars could identify all 'Top 7' STEC serogroups in artificially or naturally contaminated beef (79). The authors concluded that no single method or combination of methods was capable of identifying all STEC serogroups (79).

Phylogenetic analyses indicate that O-antigens do not necessarily relate exclusively to an STEC pathotype of a particular serogroup, and these serogroups can be represented by a number of different multilocus sequence types (MLST) (88). In particular, non-

STEC and STEC strains of the same serogroup are commonly of different lineages (88). Therefore, the assumption that the 'Top 7' STEC will have unique traits that can be exploited using specific agar media and immunomagnetic beads may be inappropriate. For example, sorbitol fermenting STEC O157 and rhamnose-fermenting STEC O26 have both been linked to human disease, but these bacteria would not be detected using traditional methods on CT-SMAC and CT-RMAC agar (92, 93).

Further research is required to develop microbiological methods for STEC detection and for continued evaluation of the bacteria at a phenotypic and cellular level. More specific molecular methods to provide fast and economical results to determine the prevalence of STEC, will also be beneficial for both scientific research, industry, and public health.

Traditional as well as advanced molecular methods for the detection and isolation of STEC are used in this thesis. Following an assessment of peer-reviewed literature, a modified USDA-FSIS method was used for traditional screening and isolation of STEC, which mainly differed from the USDA-FSIS described method by the addition of a direct plating step onto agar media in addition to IMS. Also a more efficient and sensitive RT-PCR screening protocol (66) was used for evaluation of virulence genes and serogroups using a more robust "mastermix" and a 100-well sample holder to increase throughput of samples. This thesis research required testing of over 10,000 individual wells via single-plex and multiplex RT-PCR for screening of enrichments and individual isolates.

The evaluation and use of the NeoSEEK assay is a major component of this thesis, therefore it is combined with traditional methods in **Chapter 3** and **Chapter 4** for

evaluation or screening, but it is used as a method in itself to detect the 'Top 7' STEC. While microbiological methods are essential for whole genome sequencing and in-depth analysis of an individual bacterium, the use of culture-independent diagnostic tests may be better suited for epidemiological studies and food safety.

2.4   Public health and epidemiology of STEC

4.3.25   STEC prevalence, risk factors, and transmission dynamics: humans

In 1993, an 11-month old boy from Whakatane, New Zealand presented with bloody diarrhoea; he was hospitalized for two days and then discharged after no dehydration or anorexia was noted (94). He presented later that day with seizures, no urine production (anuria), and vomiting, at which time a diagnosis of haemolytic uremic syndrome (HUS) was made, and STEC O157 was cultured from his faeces (94). This was the first reported case of STEC O157:H7 in New Zealand (94).

STEC infection is globally estimated to cause 2.8 million cases of illness, 3,890 cases of HUS, and 230 deaths annually (4). New Zealand has a relatively high incidence of notified STEC infection in humans, with 8.9 STEC cases per 100,000 population reported in 2016 (8), compared to 2.85 in the USA in 2016 (9), and 12.92 in Ireland, 5.08 in the Netherlands, and 2.05 in the United Kingdom in 2015 (8). Since it became a notifiable disease in New Zealand the incidence has shown a general annual increase (Figure 2-1), with both STEC O157:H7 and STEC non-O157:H7 causing human disease (8). The introduction of PCR screening of all diarrhoea samples in the Auckland District Health Board since July 2015 has likely influenced the rapid rise in notifications in the past few years.
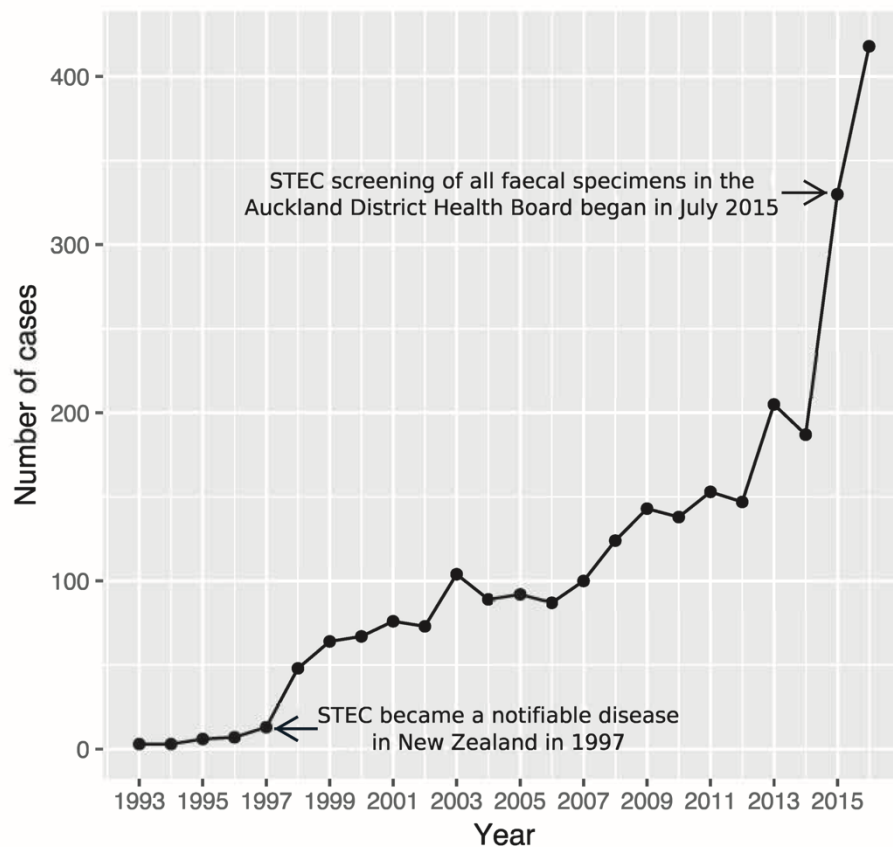
Figure 2-1: Number of STEC cases per year in New Zealand from 1993 to 2016

HUS has been linked to over 24 STEC serotypes, but STEC O157 is the most commonly implicated (95). An evaluation of non-O157 STEC reported in the USA between 2000 and 2010 identified the 'Top 7' STEC (serogroups O157, O26, O45, O103, O111, O121, O145) as the most common cause of human diseases (96). A particular clade (Clade 8) of STEC O157 were found to have increased virulence, causing HUS and more hospitalization (97). Unique combinations of virulence genes may affect the toxicity and duration of STEC disease in humans (98). A high leukocyte count and antibiotic treatment during the first week of illness have been associated with increased risk of HUS in children (99). Furthermore, clinical cerebral involvement with STEC cases is associated with increased risk of death (100). In general, non-O157 STEC infections

were associated with shorter dialysis treatments and less bloody diarrhoea, compared to STEC O157 (100).

The risk of human consumption of cattle products such as milk and beef contaminated with faecal material was the first focus of STEC public health research and response. An early case-control study in the United States from 1990 to 1992 identified eating undercooked hamburger meat as the primary risk factor for STEC O157 infection (101). Several other food sources have also been identified as a risk. Raw milk cheeses in Switzerland were contaminated with *stx1* and *stx2*-positive strains of STEC (102), and similar contamination of cheese products occurred in Egypt (103). An outbreak of STEC O26 in Italy and Romania was linked to a milk processing facility (104). A mixed serotype outbreak of STEC O145:H28 and STEC O26:H11 due to contamination of ice cream occurred in Belgium (105). STEC O26 has recently been linked to two outbreaks in the United States due to consumption of flour (106) and raw produce at a fast-food burrito chain (107). Other outbreaks linked to produce have included watercress (108) and sprouts (109). Outbreaks due to the consumption of raw produce may be associated with the use of cattle faeces as fertiliser material.

Cattle are the primary reservoir of STEC (72). The importance of animal contact for transmission of zoonotic enteric disease has been well recognized (110). A case-control study of human STEC cases in New Zealand identified living near cattle and contact with animal manure as significant risk factors for infection (12). This study also found that outbreaks were rare, and food was not a significant exposure pathway for infection in New Zealand (12). The connection between contact with cattle and their faeces has been recognized in other human populations around the world. In England,

a case-control study found exposure to the farming environment was a significant risk factor in STEC O157 infection in humans (111). A series of enteric outbreaks, including STEC, were associated with young children bottle feeding calves at a day camp in Minnesota, and hand-washing was protective for infection (112). In France, a five year study of HUS found that dairy cattle density and the ratio of calves to children under 15 years of age was significantly associated with HUS incidence (113). A case-control study of an STEC O157 outbreak in Pennsylvania found that contact with calves increased risk of infection, while handwashing was protective (114). Similar discoveries were made in Argentina, where a case-control study of STEC infection in children found that eating undercooked beef and living in or visiting a place with farm animals were risk factors for infection, while handwashing practices after handling raw beef were protective (3). A survey of dairy families in Canada found that 67.7% of the eighty families had at least one family member seropositive for *stx1*, and 30% seropositive for the O157 antigen (115). The prevalence of *stx1* antibodies was highest in children under 5 years of age (78%) (115).

Contaminated water has been recognized as a transmission route for STEC. Geostatistical investigation of STEC O26 and STEC O157 in Ireland found that private water well usage, cattle density, and home domestic wastewater treatment systems were associated with human infection (116). A New Zealand case control study also recognized contact with recreational waters as a significant risk factors for infection, in addition to cattle associated factors (12).

A summary of the main transmission routes of STEC to humans is illustrated in Figure 2-2.

Transmission cycle of Shiga toxin-producing *E. coli* to humans

Maintenance of STEC in cows, calves, and the farm environment

Ingestion of faecally contaminated produce and cow products

Direct contact with cattle and faecal material on farm environment

Drinking and recreational use of contaminated water sources

STEC infection highest in children under 5 years of age
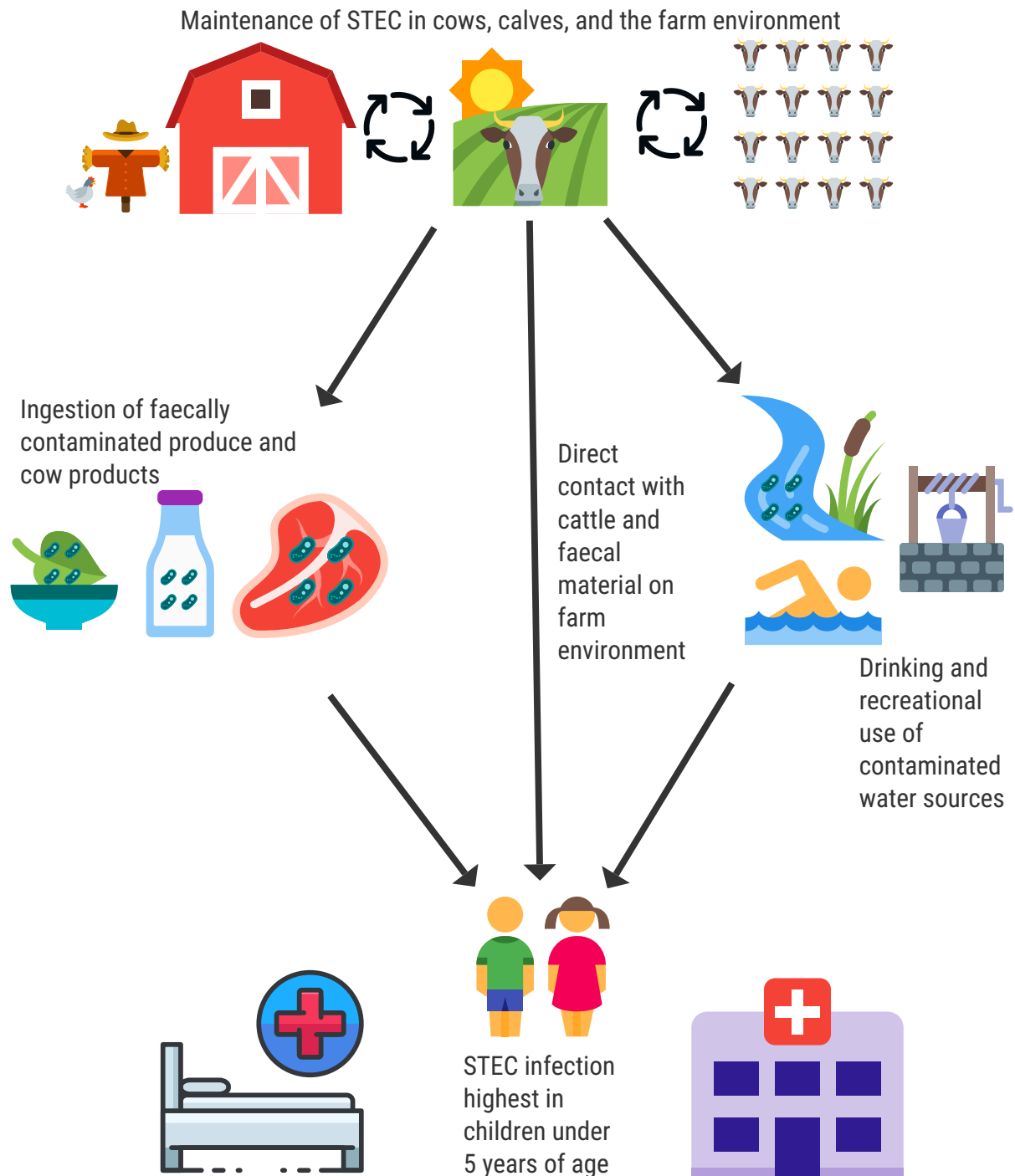
Figure 2-2: Illustration of STEC transmission cycle

4.3.26 Medical treatment of STEC infection in humans

Medical treatment of STEC is supportive, and mainly involves fluid therapy and dialysis to protect the kidneys, as well as blood products as needed. Early physician advice emphasized not consuming undercooked ground beef, raw milk, raw fruit juices, and the importance of hand-washing (117). The use of antibiotics is usually contraindicated (99). Even if children recover from HUS, 30% can have long term effects including persistent hypertension, neurological symptoms, proteinuria, and decreased kidney filtering function (5).

Recognizing the symptoms of STEC is important for public health providers. In an assessment of physician knowledge of STEC, it was found that while 83% of physicians submitted bloody faecal samples for culture, less than half (49%) thought their lab tested for STEC O157, and less than a third (30%) thought they tested for non-O157 STEC (118). Only a third (34%) of respondents correctly correlated a positive Shiga toxin test as most likely associated with STEC, rather than *Shigella* sp., and these were respondents who had a superior knowledge of STEC in other questions (118).

While not helpful to those already infected with the pathogen, global databases such as FoodNet (https://www.cdc.gov/foodnet) can provide source attribution information to help prevent further STEC infection from common food sources (119). The identification of bacterial strains across geographical boundaries can help assist with recalls of the contaminated food items.

4.3.27 Antimicrobial resistance of STEC: human and cattle prevalence

Since it is currently contraindicated to treat STEC infections with antibiotics, antimicrobial resistance is not a leading topic in STEC research. However, evaluation of antimicrobial resistance offers a view of evolution of the bacterium and STEC may

serve as a sentinel for other bacteria that obtain resistance factors. Antimicrobial resistance evaluation of over a hundred *E. coli* from calves in Canada found no association between the presence of virulence genes (e.g. *eae*, *stx*) and antimicrobial resistance, which was common (89.6% of selected isolates had at least one resistance gene) (120). In a study of Minnesota dairy farms (n=40), STEC isolates (n=83) had 23% resistance to tetracycline and 48% resistance to sulfadimethoxine; at the animal level, 77% of isolates from calves and 39% of isolates from cows had antimicrobial resistance (121). In Belgium, antimicrobial resistance was more common in non-O157 strains from humans than O157 isolates, and an extended-spectrum beta-lactamase gene $bla_{\text{TEM-52}}$ was identified in a serogroup O26 isolate (122).

4.3.28 STEC prevalence, risk factors, and transmission dynamics: cattle

The Food and Agriculture Organization of the United Nations estimates that 1.4 billion cattle are farmed worldwide (123). Several studies have estimated the global prevalence of STEC in cattle. Worldwide dairy cattle prevalence of STEC O157 ranges from 0.2% to 48.8%, and from 0.4% to 74% for non-O157 STEC (95). A global meta-analysis of STEC O157 found that 5.68% of cattle (95% CI 5.16-6.2) carry the pathogen, and feedlot cattle had a higher prevalence than other groups (124). Fifteen countries in the continent of Africa have reported STEC O157 from either humans, animals, or food (125).

STEC were isolated in faecal samples from New Zealand cattle and sheep as early as 2002 (126). As confirmed by culture isolation and inferred through sampling at four processing plants, the New Zealand farm level prevalence of STEC O157 from adult cattle and calves was 2.8%, and 6% for STEC O26 (127). This study of STEC O26 and STEC O157 prevalence found that calves had a significantly higher prevalence (6.0%)

than adult cattle (1.8%) (127). Furthermore, dairy farms had a higher prevalence than beef farms (127).

Calves have been recognized to have a higher prevalence of STEC. While initial research focused on STEC as a pathogen for calves, a systematic review and meta-analysis of STEC in calves found that STEC carriage was not associated with diarrhoea, and the average prevalence of STEC was 19.8% for calves (128). A study of diarrheic calves in Iran found a 5% prevalence of STEC (129). One to 14 week old weanling calves were determined more likely to be shedding STEC O157 and STEC O26 on Australian dairy farms than cows (130). In Denmark, there was a prevalence of 7.3% STEC O157 in bull calves and a prevalence of 25.4% of STEC that were both *eae* and *stx*-positive (131). In Minnesota, USA, STEC O157 was isolated from 4.5% of weaned calves on 4/17 (36.8%) farms examined (132). In Brazil, calves were more likely to harbour STEC than cows, while season had no effect on prevalence (133). A Scottish study evaluating STEC O26 on a single farm found that over 94% of calves had the serogroup O26 detected via PCR during the study period, and 17% of these were confirmed via isolation (61). In the Netherlands, 10.6% of cattle and only 0.5% of veal calves were positive for STEC O157, and the PFGE profiles were closely related to human isolates (134). *stx* genes have been identified in 25% of calves less than one day old on dairy farms in Argentina (135). Unweaned calves on Minnesota dairy farms had a higher risk of shedding STEC than adult cattle (OR=2.6) (136). Taken together, these studies indicate that very young calves are more rapidly colonised and shed STEC in higher proportions than older cows.

Risk factors related to farm management for STEC carriage in cattle has been evaluated in many studies (Table 2-1).

Table 2-1: STEC risk factors related to farm management by country

Risk factor	Not a risk factor	Country	Reference
	Farm size; introduction of new animals	Belgium	(137)
Milk pails, rather than nipple feeders; calves 2 weeks to 3 months		Canada	(138)
Calves 2 to 6 months	Herd size, housing conditions	Denmark	(139)
Smaller farms (<100 cattle)		Minnesota, USA	(140)
Indoor housing of cattle (OR=4.9)		Netherlands	(141)
Spreading slurry on pasture; housing cattle rather than putting on pasture		Scotland	(142)
	Organic versus conventional farming	Switzerland	(143)
	Chlorinated vs. non-chlorinated water sources in water troughs	Washington, USA	(144)

Some studies have suggested stress may cause specific cattle groups to have higher shedding of STEC than others. Observations regarding increased STEC prevalence post-weaning due to stress may be hampered by observation bias, as sampling was not performed earlier in the calf's life. One study implicated the weaning period and specific diets as a risk for increased STEC O157 shedding, but the sample design involved relocating 9 month calves from pasture into possibly contaminated sheds for the duration of the study (145). Stressors, such as high ambient temperatures, first lactation, or early lactation were implicated in increased STEC shedding of any serogroup in dairy cattle (146). However this study was only performed between May and August (summer) in Michigan, leading to biased conclusions on shedding given that most cattle would be in early lactation and high temperatures prevail during

these months, with no adequate comparison to winter milking months (146). In another study in the USA, isolation of STEC O157 pre- and post- weaning was not found to cause a significant change in prevalence (147). Weaned heifers had a higher prevalence (1.8%) compared to unweaned calves (0.9%) in one longitudinal study in the northwestern USA (148). STEC O157 was not found to vary in a cohort of calves in Italy after transport, weaning, or two months after weaning (149). A metagenomic study also found no influence of weaning on STEC shedding prevalence (150). The effect of weaning on STEC shedding is not well established.

STEC and *stx* has been detected in various sources in the farm and feedlot environment (Table 2-2).

Table 2-2: Prevalence of environmental contamination of STEC

Prevalence of sample source	Risk factor for cattle	Outcome variable	Location	Reference
37% of ground, 18% of feeder, and 15% of water samples, with 43% of samples in the rearing calf environment contaminated		*stx*	Argentina	(151)
35% of faecal samples and 20% of environmental samples		*stx*	France	(152)
3 water troughs		STEC	Ohio, USA, 20 farms	(153)
17.1% of animals, water, wildlife	Contamination of water	STEC	19 farms in USA	(154)
Feed, water		*stx*	USA	(155)
21.7% of water trough samples		STEC O157	USA	(144)

The transmission of STEC in calf pen and farm environments has been evaluated in several studies. A longitudinal study over five months using PFGE profiles of STEC O157 isolated in twenty feedlot pens indicated that environmental contamination, rather than incoming cattle, was the most important factor for STEC transmission on

farms (144). Similar PFGE profiles were seen through the study in the same feedlot, and there was also little to no evidence of transmission of organisms between pens in each feedlot, indicating highly local dynamics of transmission (144). In experimental conditions following direct inoculation of calves, STEC O157 were detected on hides, pen floors, water troughs, and pen walls as well as in RAMS of calves (58). A longitudinal study of Wisconsin dairy farms identified marked variation in STEC O157 prevalence, but one farm had a two-year period where calves, drinking water, flies, pigeon faeces, and adult cattle all tested positive to the same strain of STEC O157 (via PFGE) (156). A controlled study indicated that calf-to-calf transmission was only associated with 20% of serogroup O26 infection, while environmental exposure was a more predominant factor related to transmission (157). Modelling of the transmission dynamics revealed two-thirds of STEC serogroups were lost by calves in a day, and 40% of infections were attributable to calf-to-calf transmission, with 60% from environmental sources (158). Increased environmental contamination with STEC was associated with increased transmission rates in a pen infection-transmission experiment (159). Increased faecal contamination of calf pens was associated with increased STEC carriage and hide contamination of calves (159). Increased levels of faecal contamination of pen floors with STEC O157 were associated with increased faecal contamination of hides (160). Overall, many environmental and vector related factors may be associated with STEC transmission in cattle, and there is no predominant transmission pathway identified at this time.

Cattle may shed STEC for variable periods of time. STEC shedding in calves is more likely in the first six months of life, and decreases as the calf matures (150). In Japan, faecal shedding of STEC O157 was a few days to ten weeks in calves, while shedding of

STEC O26 persisted from one to three weeks (161). In Canada, 25% of calves under one week of age were shedding STEC O157, and continued shedding for seven weeks while in calving pens, but the prevalence dropped dramatically after they were moved out to pasture (162). Metagenomic analyses suggested that STEC shedding was more common when the calf had a lower diversity of gut microflora (150). Artificial inoculation of cows and calves with STEC O157 found that calves shed the inoculated strains for longer periods of time, but that shedding patterns varied widely (163).

STEC have been known to persist in the farm environment for long periods of time. Using PFGE profiling, a study on French dairy farms found that some strains persisted in the farm environment for up to a year, and strains were unique between farms (152). Other studies have shown that STEC O157 can survive for at least two months in sterile water (164). In cheese, STEC O26:H11 was found to tolerate a more acidic and higher salt concentration environment than STEC O157 (103).

The importance of direct transmission from cows to calves has been recognized in several studies. Udders were found to be a source of STEC contamination, with STEC detected on udders (54/175) and teats (49/161) from 13 randomly selected dairy farms in France (152). In Italy STEC O26 and STEC O157 have been detected in milk filters from dairy farms that sell cheese and raw milk (165). In contrast, an Australian study did not find any correlation between O26:H11 isolates from dams and calves on farms (166).

Other animal and insect vectors have been suggested to be involved with STEC transmission in farm environments. Under experimental conditions, inoculation of flies with STEC O157 led to transmission to all calves in a pen, as well as drinking

trough water (167). European starlings were capable of shedding STEC O157 and transmitting it to calves within three days in experimental conditions (168). Bird droppings sampled from two farms 32.5 km apart found the same restriction endonuclease digestion (REDP) subtype of O157:H7, indicating birds are a possible vector of STEC (153).

Studies of calf to calf transmission have shown variation in their results. A comparison of STEC strains between calves in sheds showed low variability, indicating within shed transmission (169). Sampling of calves on a Swedish dairy farm implicated in a human STEC outbreak found that once calves were moved to pasture, they did not continue to shed STEC, while calves kept in sheds continued to shed STEC (170). Under experimental conditions, the average R_0 for STEC O157 among beef calves was 4.3, indicating that each infected calf infected over four other calves in the same pen environment (171). Horizontal transmission of unique STEC O157 strains was not observed in weaned calves preconditioned in a feedlot setting, indicating that colonisation occurs early and the animals may not be prone to colonisation at a later time (172). In a controlled experiment, all calves which shared the same shed as an STEC O157 inoculated calf began shedding the same strain in less than 3 weeks (173). This study also found that very low (<300 CFU/ml) doses were sufficient to cause infection (173). Even in a pasture environment, all of the susceptible calves were shedding STEC O157 within six days after introduction of inoculated calves, with a R_0 of 7.3, indicating infected calves spread STEC to seven other calves within the 28 days of the assumed infectious period (174). This study also noted that susceptible calves grazing a previously infected STEC O157 pasture that had been fallow over the winter did not become colonised for over 14 weeks of the study (174). A controlled experiment

found that group penning led to increased transmission of a traceable STEC strain among all calves in four days, with pen floors and calf hides also becoming contaminated (175).

Much of the molecular evidence of STEC transmission indicates localized, on-farm, or within-shed transmission. Sequence-based PCR fingerprinting of STEC isolates indicated that isolates taken from pastoral farms in Virginia, USA, were similar, suggesting within-farm transmission, but no links were found between calves and dams (169). California dairy farms with high prevalence of STEC O157 had unique strains on each farm (176). SNP (single nucleotide polymorphism) analysis, indicating single base changes on common genes between several bacteria, indicated that STEC O157 populations were dominated by a single sequence, but differences occurred between farms, and some sequences remained the same after resampling 11 months later (176). In New Zealand, PFGE analysis of *E. coli* serogroup O26, the most common 'Top 7' serogroup in the study, found that isolates from the same farm were more similar than isolates from other farms (177). This study did not find any significant differentiation based on geographical location, but it was focused on a 15 km² area (177). All STEC O157 isolates from a dairy herd in England that was implicated in a human outbreak had the same phage type, toxin genotype, and plasmid profile (178). PFGE of O26:H11 isolates on Australian farms also found unique strains at each farm (166).

STEC is commonly found in many environmental sources, and several potential transmission routes have been identified in many countries. Overall, much of the

evidence suggests that transmission and persistence of STEC is linked with local environmental transmission.

4.3.29 Super-shedders

STEC super-shedder cattle have been extensively evaluated in research (179). STEC super-shedders are defined as animals that excrete >10^4 CFU per gram of faeces (180). The idea that specific cattle were mainly responsible for the transmission and infection of other cattle led to significant interest that detecting and removing these super-shedders from herds could be a valuable intervention to decrease STEC prevalence (179). However, there is growing evidence that all cattle may shed STEC at super-shedder levels at one point in time, and it is not a feature that would be unique or identifiable to individual animals.

Numerous studies have supported the heterogeneity of shedding by individual animals. In one study, 24 of a 52 heifer cohort sampled weekly over 18 weeks were detected as super-shedders; the majority (19/24) of heifers detected as super-shedders were only detected on a single sampling occasion, indicating that the trait was not frequently associated with the same animal (181). Furthermore in the same study, daily shedding prevalence in heifers was highly variable (16.7%-100%), and no significant association existed between shedding of STEC O157 and the shedding status two days prior (181). In another study, populations of STEC O157 shed by animals varied between undetectable to >10^3 CFU per gram within a few hours; two cohorts of calves (n=14, n=16) sampled over 5 and 15 days all shed O157 at some point (182). A study to evaluate super-shedding of STEC O157 and STEC O26 in Irish dairy herds identified only 4 super-shedders from 529 animals sampled, and none of these were ever identified more than once (183). A dairy herd that was implicated in a STEC O157

outbreak had all cattle sampled twice a month for 15 months, and found that shedding stopped for several months (November to May), and 74% of the animals that tested positive for STEC O157 only did so on a single sampling occasion (178). A New Zealand study that pre-evaluated "positive" and "negative" farms for its study design found that some farms switched to positive and negative throughout the duration of the study (July to September, 2010) (184).

An in-depth molecular analysis of STEC O157 isolates retrieved from super-shedding cattle identified 52 unique PFGE patterns, and 19 phage types, indicating that no exclusive STEC genotype explained their supershedder classification (185). Another study investigated pathogen excretion through modelling of STEC O157:H7 super-shedding from feedlot cattle (186). This study concluded that the importance of super-shedders in transmission is overstated, with a third of the herd having been found to have been super-shedding throughout the experiment (186). Another study concluded that within-animal variation of STEC shedding is greater than between animal shedding over time, and resources would be better spent sampling fewer animals more times to detect super-shedding (187).

Logistically, enumeration of colonies is limited by the effectiveness of selective media; the sensitivity of methods to isolate STEC on agar was previously discussed. Evaluation of super-shedders would also require daily sampling of the same animals for at least 30 days (181). Finally, no realistic and economic farm level intervention exists to identify super-shedders in herds and remove them from farms on a national level.

4.3.30 Interventions to decrease or eliminate STEC in cattle

Several therapeutic interventions have been evaluated to decrease shedding of STEC by cattle. The most commonly used method is increasing immunity in cattle to STEC, either through colostrum or vaccination. Increased antibody titres to serogroup O26, O111, and O157 have been demonstrated in calves after administration of colostrum (188).

Several vaccines have been evaluated for decreasing STEC shedding in cattle. Vaccination of pregnant cows did allow for transfer of Shiga toxin 2 neutralizing antibodies via colostrum to calves (189). One vaccine study used bacterial ghosts of STEC O157, a cell with no cytoplasm, and found a decrease but not elimination of STEC O157 shedding after two vaccinations (190). A separate vaccine study evaluated the use of vaccines both on the pre-partum dam and on their calves, and found no significant differences in faecal shedding of STEC O157 between the treatment and control groups (191). A multivalent vaccine based on various virulence genes showed STEC O157 shedding was reduced in calves, but had no effect on the proportion of calves colonised (192). However, the proportion of calves colonised decreased when serotype specific targets (H7 flagellin) were used (192). Vaccine studies have also assessed the potential ability to decrease Shiga toxin 2 shedding *in vitro* (193).

There is some doubt that an immune response would help prevent the shedding of STEC in the digestive tract of a reservoir animal. Serological testing of fifteen beef herds across the Midwestern USA found, using serological anti-O157 antibody testing, that most calves (83%) and all herds had been exposed to STEC O157 (194). This would indicate that most calves already have the ability to produce a protective immune response to STEC without the use of a vaccine intervention. Dexamethasone-induced

immunosuppression of calves challenged with STEC O157 showed no difference in duration of shedding from those who were not immunosuppressed, and little difference in prevalence of bacterial shedding (195). Also, serology of dairy cows with STEC O157 found no correlation between shedding and antibody titres (196). A New Zealand study found carriage of STEC O157 was not associated with the concentration of IgG in the serum of calves at processing plants (83).

With a previously mentioned R_o of 4.3 of STEC transmission among calves, 65-86% of the herd would need to be effectively vaccinated to eliminate STEC from the herd (171). The key question is whether the cost and effort required in individually dosing cattle is realistic; the practicality of vaccination becomes even more tenuous when considering interventions for young bobby calves, which in New Zealand are predominantly sent to slaughter at four days of age.

Other therapeutic interventions include the feeding of antibiotics, prebiotics, or probiotics. Supplementation of the antibiotic oxytetracycline in milk replacer for calves did not result in any difference in STEC O157 shedding from the control group (197). An evaluation of a prebiotic, Celmanax®, suggested it decreased carriage of STEC in calves, but this relationship was only seen once and not consistently during other seasons when the research was performed (198). An assessment of probiotic treatment found no significant difference in faecal shedding of STEC O157, STEC O26, and STEC O111 inoculated calves during the first week of life, although the probiotic group had less long term shedding for the O26 and O111 serogroups (199).

4.3.31 Prevalence, risk factors, and transmission dynamics: cattle processing plants

The prevalence of the 'Top 7' STEC has been evaluated in cattle processing plants in many countries (Table 2-3). An Australian study identified veal as the highest prevalence group when compared with older cattle, both from PCR testing (46.0%) and bacterial isolation (12.7%) of the 'Top 7' STEC.

Table 2-3: Prevalence of STEC at cattle processing plants

Prevalence of sample source	Animal class	Diagnostic target	Location	Reference
10.5% faecal, 2.6% carcass	Calves	STEC O157 bacteria	Argentina	(200)
44.5% faecal samples	Calves and cows	non-O157 'Top 7' STEC serogroup, *eae*, *stx*	Australia	(82)
25.9% faeces, 64.9% hides, 7% carcass	Dairy cattle	'Top 7' STEC, NeoSEEK	USA	(201)
6.5% faeces, 15.6% hides, 1% carcass	Dairy cattle	'Top 7' STEC bacteria	USA	(201)
90% hides, 68% carcass	Calves	Non-O157 STEC, NeoSEEK	USA	(202)
6.3% RAMS and carcass	Cattle	STEC O157 bacteria	Turkey	(203)
16% faecal, 4% carcass	Calves	STEC O157 bacteria	Italy	(204)
17.6% hides, 2.3% faeces, 0.7% carcass	Beef cattle	STEC O157 bacteria	Ireland	(205)
20% hides, 7% carcass	Calves	STEC O157 bacteria	USA	(202)

Various risk factors can be linked to carcass contamination. For instance, the connection between hide contamination and carcass contamination is well established. Greater concentrations of STEC O157 on hides were correlated to greater concentrations of STEC O157 on pre-evisceration carcasses (206). Another study in Australia also found that detection of a non-O157 STEC serogroup in a faecal sample was associated with prevalence in a hide sample (201). STEC biofilms in processing plants may be a significant barrier to proper disinfection, and allow the persistence of STEC in the processing plant environment (207). Carriage of STEC O157 was not

associated with carcass weight or gender of calves in a study of veal processing plants in the North Island of New Zealand (83).

Transmission during transportation can likely influence STEC prevalence at processing plants. A study of STEC prevalence in cattle trailers in Canada found a high prevalence of *E. coli* serogroups via PCR (>90% for O45 and O103, 75-85% for O26 and O157, 50-70% for O121, and <10% for O111 and O145) (81). STEC O157 hide prevalence increased from 50.3% to 94.4% between the time cattle were loaded for transport and stunned at the processing plant (208). Only 29% of STEC O157 isolates collected at the processing plant matched isolates collected on farm using PFGE, indicating cross-contamination of carcasses from other sources (208).

4.3.32 Interventions for STEC in cattle at processing plants

The evaluation of chemical and physical interventions on processing lines at cattle processing plants is a well described and researched topic that is not assessed in this thesis. However, the value of educational outreach and food safety training at cattle processing plants cannot be overlooked. An assessment of current processing plant practices in the Netherlands revealed half the plants had structural and procedural inadequacies that led to contamination of carcasses (209). This study led to significant improvements in hygiene practices. A previously mentioned USDA study in veal processing plants found that after three processors improved their harvest process as recommended by the USDA-FSIS, there were significant drops of almost 50% in the contamination levels of hides compared to carcasses (202). Last year, the Meat Industry Association in New Zealand worked together with the Ministry of Primary Industries to introduce nine initiatives aimed at reducing the risk of STEC contamination of veal, including hosting STEC workshops targeted at senior

operators, supervisors, technical staff and on-site verification staff. This was repeated prior to the 2017 spring calving season with excellent attendances at four locations throughout the country (14).

4.3.33 Relevance

STEC is a zoonotic pathogen, therefore an understanding of the interaction between the reservoir host, the environment, and humans is critical for the development of new transmission mitigation opportunities and assessing their effectiveness. While several previous New Zealand studies have effectively recognized the importance of cattle in the farm environment as a risk to humans, as well as performed prevalence studies with cattle processing plants to estimate prevalence, there is no direct knowledge of the STEC prevalence nationally, nor how STEC is transmitted in the farm environment to high risk cattle populations (e.g. young dairy calves processed for veal). Furthermore, while previous studies in New Zealand have evaluated STEC prevalence at cattle processing plants, no research has assessed the link between the farm environment and eventual contamination of carcass at processing plants. These knowledge gaps are addressed in this thesis.

Advanced molecular methods of detection and evaluation are utilized in this thesis. The application of the PCR/MALDI-TOF assay allows for a superior assessment of true prevalence of STEC both in animals and the environment, which is a great benefit to this epidemiological research. While previous studies have utilized PFGE typing to understand source attribution or transmission dynamics, this method does not have the unparalleled sensitivity of whole genome sequencing. This thesis utilizes whole genome sequencing to evaluate STEC at the cellular level, as well as applying the

technology to assess the broader evolution and population dynamics to understand the broad epidemiology of this pathogen in New Zealand and around the world.

Current direct interventions to decrease STEC in cattle, such as the use of vaccines or probiotics, have mixed results. These methods also may be expensive and time-consuming; making them unrealistic for application by farmers at a national level. Evaluating the prevalence and transmission dynamics of STEC on dairy farms may lead to effective and realistic interventions that could benefit public health, as well as decrease the risk of STEC contamination of beef.

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Andrew Springer Browne

Name/Title of Principal Supervisor: Distinguished Professor Nigel French

Name of Published Research Output and full reference:
Browne, AS, Midwinter, AC. Withers. H, Cookson, AL, Biggs. PJ, Marshall JC, Benschop, J, Hathaway, S, Haack, N, Akhter, R, and French, NP. Molecular epidemiology of Shiga toxin-producing Escherichia coli (STEC) on New Zealand dairy farms: application of a culture-independent assay and whole genome sequencing. Applied and Environmental Microbiology, under second review after revision.

In which Chapter is the Published Work: 3

Please indicate either:

• The percentage of the Published Work that was contributed by the candidate:

and / or

• Describe the contribution that the candidate has made to the Published Work:
AS Browne designed the study with input from all supervisors.
AS Browne performed 100% of the field work and 80% of the laboratory work (N Haack and R Akhter performed initial laboratory processing while AS Browne was in the field). AS Browne drafted the manuscript and created all figures, with input from supervisors.

_____
Candidate's Signature

13/2/18
Date

_____
Principal Supervisor's signature

14/2/18
Date

49

I'm taking this bloody car to Invercargill, boy!

John, Good Bye Pork Pie

3 Molecular epidemiology of Shiga toxin-producing *Escherichia coli* (STEC) on New Zealand dairy farms: application of a culture-independent diagnostic test and whole genome sequencing

3.1 Abstract

New Zealand has a relatively high incidence of human cases of Shiga toxin-producing *Escherichia coli* (STEC), with 8.9 STEC cases per 100,000 people reported in 2016. Previous research showed living near cattle and contact with cattle faeces as significant risk factors for STEC infections in humans in New Zealand, but infection was not linked to food-associated factors. During the 2014 spring calving season, a random stratified cross-sectional study of dairy farms (n=102) in six regions across New Zealand assessed the prevalence of the 'Top 7' STEC (serogroups O157, O26, O45, O103, O111, O121, O145) in young calves (n=1,508) using a culture-independent diagnostic test (PCR/MALDI-TOF). Twenty percent (306/1,508) of calves on 75% (76/102) of dairy farms were positive for at least one of the 'Top 7' STEC. STEC carriage by calves was associated with environmental factors, increased calf age, region, and increased number of calves in a shared calf pen. Strong clustering of 'Top 7' STEC positive calves was observed for the O157, O26, and O45 serogroups within the same pens and farms, indicating that if one calf was positive, others in the same immediate environment were more likely to be positive as well. This finding was further

evaluated with whole genome sequencing that indicated a single *E. coli* O26 clonal strain could be found in calves in the same pen or farm, but diverse strains existed on different farms. This study provides evidence that would be useful for designing on-farm interventions to reduce direct and indirect human exposure to STEC.

3.2   Introduction

Worldwide, Shiga toxin-producing *Escherichia coli* (STEC) are a growing public health concern. Large scale outbreaks in Europe (1) and the United States (2) have continued to occur. Furthermore, human STEC cases in Argentina have a high rate of serious clinical complications (3). Although STEC may have a lower prevalence than other notifiable zoonotic diseases (4), the pathogen's propensity to affect very young children, leading to potential long term kidney and brain damage (5), is a concerning public health issue. STEC are primarily transmitted via the faecal-oral route. Ruminant animals, particularly cattle, have been identified as the most important reservoir (6).

New Zealand has a relatively high incidence of notified STEC infection in humans, with 8.9 STEC cases per 100,000 population reported in 2016 (210), compared to 2.85 in the USA in 2016 (9), and 12.92 in Ireland, 5.08 in the Netherlands, and 2.05 in the United Kingdom in 2015 (10). Since it became a notifiable disease in New Zealand, there has been a general increase of STEC cases annually, with both STEC O157:H7 and non-O157 STEC causing human disease (11). A New Zealand case-control study identified contact with animal manure and the presence of cattle in the local area, along with contact with recreational waters, as significant risk factors for human STEC infection (12). Interestingly, the same study did not identify food as a statistically significant exposure pathway in New Zealand (12). Previous research findings overseas

have highlighted beef food products and raw produce as the main sources of human infection (6, 72, 211), but findings in the United Kingdom also identified an important contribution from environmental sources (111). Determining the carriage of STEC in ruminant hosts, through targeted national studies, will help our understanding of the epidemiology of this important pathogen.

Since the 1993 outbreak of STEC O157:H7 in the United States (16), monitoring and regulatory requirements regarding this pathogen have increased. Having found STEC O157:H7 in raw ground beef, plus outbreaks associated with consumption of undercooked beef patties, the US declared STEC O157:H7 an adulterant of beef in 1994, followed by declaration of six additional serogroups (O26, O45, O103, O111, O121, O145) as adulterants in 2011 (13). These six additional serogroups and STEC O157 are known as the 'Top 7' STEC. In 2015-2016, 50% of New Zealand beef exports were sent to the United States (212). Given the importance of agricultural exports for the New Zealand economy, STEC is an economic as well as a public health concern.

Previous research in New Zealand identified a higher STEC O157 and STEC O26 prevalence in young calves compared to adult cattle (127), and this finding has been supported by 'Top 7' STEC research in other countries (133, 150, 213, 214). New Zealand dairy farms follow a seasonal calving strategy, where surplus dairy calves, known as bobby calves, may be slaughtered at a very young age (four to ten days old). The higher prevalence of STEC among very young calves means that preventing inadvertent contamination of veal during dressing of carcasses at primary processing is an important risk management goal. Similarly, reducing children's contact with calves is likely to lessen the risk to human health.

Culture-based approaches for confirmation of STEC are the gold standard method. However, they have a low sensitivity, particularly with regards to isolation of non-O157 STEC, and this makes it difficult to estimate the true prevalence in epidemiological studies (90). The USDA recommended formulation for Rainbow Agar was found to inhibit growth of 20 of 96 examined STEC strains in a laboratory study (78). There were also a large range of STEC colony colours on Rainbow Agar, despite recommendations that certain colony colours indicate whether a colony is STEC or non-STEC (78). Culture isolation methods often involve the use of immuno-magnetic separation (IMS) using beads which bind to the target O antigen of the bacteria to increase sensitivity. A comparison of direct plating after enrichment and using immuno-magnetic separation (IMS) beads before plating found that only the O157 serogroup benefited from IMS, and this method did not have any beneficial effect on isolation of the four of the other 'Top 7 STEC' serogroups evaluated (O26, O103, O111, O145) (69). Another study of the non-O157 'Top 7' STEC showed that experimentally spiked faecal samples had varying success rates of retrieval using IMS: 92.3% for O26 and 100% for O103, in comparison to 38.5% for O121 and 43.8% for O111 (80). For naturally occurring faecal samples that tested positive for these serogroups using molecular screening methods, the retrieval was much lower: 21-25% for O26 and O45, and 6-10% for O145, O121, O111, and O103 (80). A study in the USA using faecal samples from feedlot cattle (n=108) and dairy cattle (n=108) isolated 33.1% (53/160) and 35.5% (11/31) of PCR-positive samples for *E. coli* O157 and non-O157 serogroups using bacterial isolation methods, respectively (63). Despite a PCR prevalence of serogroups O157, O26, O45, and O26 exceeding 78% in one research study, only 4-10% of the samples had STEC isolates successfully isolated (81). A Scottish study evaluating STEC

O26 on a single farm found that over 94% of calves (n=664) had the serogroup O26 detected via PCR during the study period, and 17% of these were confirmed via isolation (61). One extensive study that evaluated several methods for bacterial isolation of 'Top 7' STEC found that no enrichment protocol would work for all artificially inoculated samples, and significant variation occurred with naturally contaminated samples (79). In addition, neither one nor a combination of agars could identify all STEC serogroups in artificially or naturally contaminated beef. The authors concluded that no single method or combination of methods was capable of identifying all STEC serogroups (79). Overall, there is a marked disparity between the detection of samples containing 'Top 7' STEC serogroups and virulence genes (*stx*, *eae*), and the bacterial isolation of STEC from those samples.

The utilization of culture independent diagnostic tests (CIDTs) to detect bacterial zoonoses is rapidly increasing, and such methods for detecting STEC are widely utilized (91). Furthermore, the utilization of CIDTs had led to an increased detection of STEC via culture dependent methods in the United States (9). Thus, the development of culture-independent assays for STEC detection which are sensitive and specific will improve epidemiological studies of this pathogen. In this study, we used an established molecular confirmation method: NeoSEEK (NeoSEEK STEC Confirmation, Neogen Corporation, Lansing, MI, USA). NeoSEEK uses PCR amplification to generate allele-specific DNA products of different masses, and chip-based mass spectrometry to analyse the extension products. The assay is based on the presence of single nucleotide polymorphisms (SNPs) in the O-antigen gene cluster that can differentiate between STEC and non-STEC bacterial strains of the same serogroup (88). NeoSEEK uses over 89 gene targets via PCR/MALDI-TOF to detect the

presence of the 'Top 7' STEC without the need for agar-based culture isolation. This assay has a "Letter of No Objection" from the USDA-FSIS, and is used commercially as a confirmation method for detection of STEC in ground beef and beef trim. The evaluation and application of this technology in this study to detect faecal carriage of STEC in calves is unique. Using culture-based and real-time PCR methods, we assessed the use of NeoSEEK for detection of New Zealand STEC.

Whole genome sequencing (WGS) offers an unparalleled resolution to assess the genetic similarity between closely related bacterial strains. We retrieved serogroup O26 bacterial isolates, and a subset of these underwent WGS and subsequent comparative analysis. The primary aim of this evaluation was to compare the similarity of the isolates retrieved from calves in the same calf pen, same farm, and same region, in comparison to other calves.

This study examined the prevalence of young calves shedding 'Top 7' STEC (O157, O26, O45, O103, O111, O121 and O145) on dairy farms in New Zealand. We estimated the spatial distribution of STEC-positive farms, clonal relationships of STEC bacteria in calves by pen and farm, and determined risk factors for STEC carriage by calves which could potentially be targeted for control. By understanding and reducing STEC from its source, we hope to decrease the risk of both veal meat contamination and human exposure to STEC on farms.

3.3 Materials and Methods

The Animal Ethics Committee of Massey University, Palmerston North, New Zealand approved this study on April 17th, 2014, under protocol number 14/29.

4.3.34 Sample size calculations

We performed sample size calculations using a cluster-sample calculation with a
design effect of 3.6, based on a previous repeated cross-sectional study of STEC O26
and STEC O157 at cattle processing plants in New Zealand (127). Table 3-1 contains
sample size calculations for the number of farms and calves required to be 95% certain
that the prevalence estimate is within +/-20% of the true prevalence. Given previous
estimates of STEC O26 and STEC O157 prevalence in calves in New Zealand, we used a
conservative estimate of 20% farm prevalence of the 'Top 7' STEC, and aimed to
recruit a minimum of 93 farms and sample a maximum of 15 calves per farm. The
critical probability for all statistical analyses was $p < 0.05$.

Table 3-1: Sample size calculations for farms and calves with cluster-sample design effect of 3.6

Estimated 'Top 7' STEC farm prevalence*	Farms needed	Calves per farm	Total samples
25%	53	20	1060
25%	70	15	1050
25%	105	10	1050
20%	70	20	1400
20%	93	15	1395
20%	139	10	1390
15%	99	20	1980
15%	132	15	1980
15%	197	10	1970

*Based on one animal being positive for any of the 'Top 7' STEC

4.3.35 Random stratified farm selection

We selected farms using a stratified random sampling scheme based on regionally
proportioned sampling of the number of farms in each region. We targeted the six

largest dairy regions, which account for approximately 75% of the dairy farms in New Zealand: Northland, Waikato, Taranaki, Manawatu-Wellington, Canterbury, and Southland (215). Given a 60-day calving period, only farms with a documented herd size of more than 150 milking cows were eligible, to ensure enough calves would be present on the day of sampling. Potential farms were selected randomly from a national farm database (Agribase[TM], AsureQuality Limited, Auckland, New Zealand) and contacted by telephone, leading to the recruitment of 102 dairy farms (Table 3-2).

Table 3-2: New Zealand dairy farms per region sampled and milking herd size, based on farm manager records

Region (number of farms)	Median herd size	Range*
Northland (10)	210	35 to 920
Waikato (35)	360	120 to 960
Taranaki (19)	320	140 to 560
Manawatu-Wellington (12)	440	320 to 1100
Canterbury (14)	800	200 to 1960
Southland (12)	580	440 to 980

*Farms in three regions contained fewer than the target 150 milking herd size on the day of sampling; adequate numbers of calves were sampled

4.3.36 Random animal selection and sampling within calf pens on farms

We categorized calves into two groups: young calves from two to nine days of age, and older calves from 10 to 21 days of age. Given the focus on 'Top 7' STEC prevalence in very young calves, where possible ten calves were sampled in the young age group, and five in the older age group.

Sampling of calves occurred during a single farm visit from July 28th to September 24th during the 2014 Spring calving season. For this study, 'pen' was defined as an enclosed area where calves had direct contact with each other and shared water and feeding

resources. After identifying calf ages, up to three pens were selected that allowed for the maximum number of animals in the two age groups to be sampled, with equal numbers per pen where possible. A random number generator was used to select pens if more than three suitable pens were available for sampling. If more than five animals were present in a pen, a spin-pointer mobile phone application was used to randomly select the first calf to be sampled, after which animals were chosen in an alternating manner in the clockwise direction, in proportion to the total calves in the pen. Calves were marked for selection and then again following sampling to maintain the random selection and prevent resampling.

Animals were excluded from sampling if they appeared injured or sick, based on visual clinical assessment by the sampler (A.S. Browne: a registered veterinarian). In total, 1,508 young calves from 267 pens were sampled by collecting recto-anal mucosal swabs (RAMS) from each calf using Amies transport swabs (Copan Diagnostics Inc., Brescia, Italy). All RAMS were kept on ice in an insulated container immediately after sampling, until they were shipped for processing the same day as they were collected.

4.3.37   Initial laboratory processing

All RAMS were shipped on ice overnight to ^{m}EpiLab, Massey University, Palmerston North, and enriched in modified Tryptone Soya broth (mTSB, Oxoid Limited, Hampshire, United Kingdom) at 42°C for 15-21 hours. Genomic DNA was extracted from 1mL of enrichment broth using a double-wash boil preparation method, according to the GeneSeek laboratory's instructions, and frozen at -80°C. The DNA samples were shipped to GeneSeek Operations (Lincoln, Nebraska, USA) on dry ice. All samples were analysed using the PCR/MALDI-TOF assay, NeoSEEK, for presence of the 'Top 7' STEC.

4.3.38 Evaluation of NeoSEEK for New Zealand 'Top 7' STEC detection

A technical report, including summary data from the study conducted for NeoSEEK to receive a Letter of No Objection, is available online (216). Prior to field collection of samples for this study, 100 characterised New Zealand STEC and non-STEC isolates from the six serogroups (O26, O45, O103, O121, O145, and O157; n=88) as well as non-'Top 7' serogroups (n=12) from the Ministry of Primary Industries (n=64) and the Hopkirk Institute (n=36) culture collections were obtained and used by the Institute of Environmental Science and Research to evaluate the detection efficacy of the NeoSEEK assay. One Australian STEC O111 isolate was also tested as no STEC O111 has been isolated in New Zealand. All 101 isolates had undergone serological analysis and previously been characterized by PCR for the presence of *stx1*, *stx2*, and *eae* virulence markers; there was 100% concordance with the NeoSEEK assay.

All DNA samples derived from the calf faecal samples, in addition to being submitted for NeoSEEK analysis, were tested for the 'Top 7' serogroups at ^{m}EpiLab using a real time PCR (RT-PCR) assay (66). All DNA samples were run with positive, negative, and blank template controls using PerfeCTa® Multiplex qPCR ToughMix® (Quanta Biosciences, Beverly, Massachusetts, USA) on a Rotor-Gene Q 5plex HRM Platform (Qiagen, Hilden, Germany). In-house validation of the RT-PCR method revealed a limit of detection (LOD) of 10^2 colony forming units (CFU) per mL for all serogroups evaluated, except for O157 and O103 where the LOD was 10^1. The LOD of the NeoSEEK assay was approximately 10^3 CFU/mL (E. Hosking, pers. comm.).

Latent class modelling (217) was used to estimate the sensitivity and specificity of serogroup detection of the 'Top 7' STEC serogroups. This modelling technique is used to compare two diagnostic tests, when neither is considered a "gold standard". Latent

class analyses were performed (https://github.com/jmarshallnz/lcar) to calculate a 95% CI for the sensitivity and specificity of the NeoSEEK and RT-PCR methods for detection of all seven serogroups for the 1,508 DNA samples (Appendix Table 1, Appendix Table 2, Figure 2-1). Latent class analyses also produced prevalence estimates of all seven serogroups by three factors: region (n=6), age of calf (young and old), and location in the North or South Island.
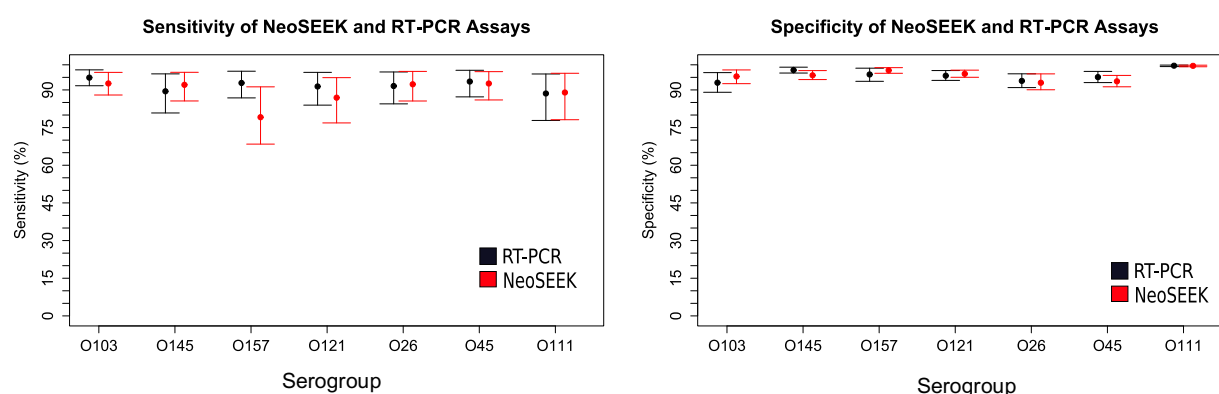
Figure 3-1: Sensitivity and specificity of NeoSEEK and RT-PCR assays for detection of the 'Top 7' serogroups in calf faecal enrichment samples (n=1,508)

All 1,508 calf RAMS samples collected were enriched and stored in a glycerol (4:1 ratio) suspension in a -80°C freezer. The isolation of individual STEC from frozen enrichment broth was important for confirming the STEC detection using molecular methods (NeoSEEK, RT-PCR), as well as analysis of the bacteria using whole genome sequencing. Recovery of bacteria from frozen enrichment broth samples was attempted based on the NeoSEEK assay results for 'Top 7' STEC positive samples.

Due to the large number of STEC bacteria detected by analysis of all 1,508 samples using the NeoSEEK assay (n=408), and the costs and labour required for testing and isolation of bacteria from enrichment broth samples, isolation was prioritized based on serogroup. Due to their public health importance, recovery was attempted on all STEC O157 (n=29) and STEC O26-positive (n=109) samples using a modification of

USDA-FSIS methods (52, 53). STEC recovery was attempted on samples using sorbitol

MacConkey agar supplemented with cefixime and tellurite (CT-SMAC) and rhamnose

MacConkey agar supplemented with cefixime and tellurite (CT-RMAC) (Fort Richard

Laboratories, Auckland, New Zealand) for STEC O157 and STEC O26 respectively,

using immuno-magnetic separation beads (IMS) (Abraxis, Warminister, Pennsylvania,

USA). The methods used were adapted from the USDA-FSIS methods (52, 53) to

include an initial "direct" culture screen, where frozen glycerol enrichment culture

was plated directly onto selective agar (CT-SMAC for O157 and CT-RMAC for O26). If

target STEC serogroups were not identified, frozen glycerol enrichment broth was re-

enriched in mTSB broth and immunomagnetic separation (IMS) was attempted

according to manufacturer's instructions. Up to ten colonies were tested for the

specific serogroups on a plate using latex agglutination, and up to four positive

individual isolates were subcultured and stored frozen with glycerol. Subcultured

isolates were confirmed for serogroup and tested for virulence-associated genes using

an in-house RT-PCR (66).

4.3.39 Whole genome sequencing, assembly, and analysis of *E. coli* serogroup O26 isolates retrieved from calf faecal samples

We used random stratified selection by region, farm, and calf pen to select 66

serogroup O26 bacterial isolates (45/66 STEC O26, 21/66 non-toxigenic O26) for

whole genome sequencing. Multiple isolates were selected from four calves to evaluate

within-animal diversity. We performed DNA extraction from a single colony picked

from Columbia Horse Blood Agar (Fort Richard Laboratories, Auckland, New Zealand)

using the QIAamp® DNA MiniKit (Qiagen, Hilden, Germany), and prepared the

libraries using the Nextera XT DNA Library Preparation Kit (Illumina, San Diego,

United States). Prepared libraries were submitted to New Zealand Genomics Limited (University of Otago, Dunedin, New Zealand), who performed sequencing using Illumina MiSeq 2x250 PE and Illumina HiSeq 2x125bp PE v4. Processed reads are publicly available on the NCBI Sequence Read Archive (SRA) under BioProject ID PRJNA396667.

Raw sequences were evaluated, assembled, annotated, and analysed using the Nullarbor pipeline in the "accurate" mode (218). RaxML maximum-likelihood trees were generated from Roary data for core genes via single nucleotide polymorphism (SNP) analysis of core genes, and accessory genes via a presence/absence matrix (219). Assembled genomes were batch uploaded to the Center for Genomic Epidemiology server for identification of virulence factors, multilocus sequence type (ST), antimicrobial resistance genes and somatic (O) and flagellar (H) type (220). A distance matrix was created from the SNP distances between isolates, and a dissimilarity matrix was created from the presence/absence matrix of the accessory genome from Roary, as well as the 26 virulence genes predicted by the Center for Genomic Epidemiology output, and all three were evaluated in PERMANOVA and CLUSTER (PRIMER-E, Quest Research Limited, Auckland, New Zealand) with region and farm as independent factors.

Figures depicting the phylogenetic relationships and associated variables were created using iTOL (Interactive Tree of Life) software (221), and further amended using Inkscape open source software version 0.92.2 (https://inkscape.org).

4.3.40 Data Retrieval and Statistical Analysis

At the time of the visit, written consent to participate in the study was obtained from a manager on every farm. Animal and farm level data, including management and environmental factors, were collected from each farm through observation, electronic devices, and interviewing a manager on every farm (Appendix Table 3).

All statistical analyses were performed using R Version 3.2.1 (222). Eight outcome variables were considered: the presence or absence of each of the 'Top 7' STEC, and an additional variable specifying the presence or absence of any of the 'Top 7' STEC. All factors were first assessed using machine learning techniques from the "randomForest" package (223). The most important 10% of factors identified in the "randomForest" analysis were considered as explanatory fixed effects in a linear mixed-effects model, with "pen" within "farm" included as random effects variables (28). A preliminary model was generated by stepwise backward elimination of the least significant variables, and eliminated variables were assessed for confounding. Confounding variables, determined by a change of >30% in the main variable coefficient, were kept in the model even if they were non-significant. Intraclass correlation (ρ) was calculated using the "aod" package with a Monte Carlo 1-way generalized linear mixed model (224). A description of the strength of correlation is as follows: 0.00-0.19: "very weak"; 0.20-0.39: "weak"; 0.40-0.59: "moderate"; 0.60-0.79: "strong"; and 0.80-1.0: "very strong".

3.4   Results

4.3.41 Prevalence of 'Top 7' STEC serogroups by latent class analysis of RT-PCR and NeoSEEK

Our in-house RT-PCR assay was only able to detect the presence of the O-serogroup in a sample, whereas the NeoSEEK assay was able to discriminate between *stx-*

positive and *stx*-negative *E. coli* of a 'Top 7' STEC serogroup (e.g. STEC O26 versus non-toxigenic O26). By using latent class modelling techniques, the prevalence of these serogroups was determined using both assays to give a more robust estimation of serogroup prevalence (Figure 3-2, Figure 3-3). This modelling technique required the calf population to be divided into groups for comparison; we therefore stratified by region, North and South Island, and age: young (2 to 9 days) and old (10 to 21 days).
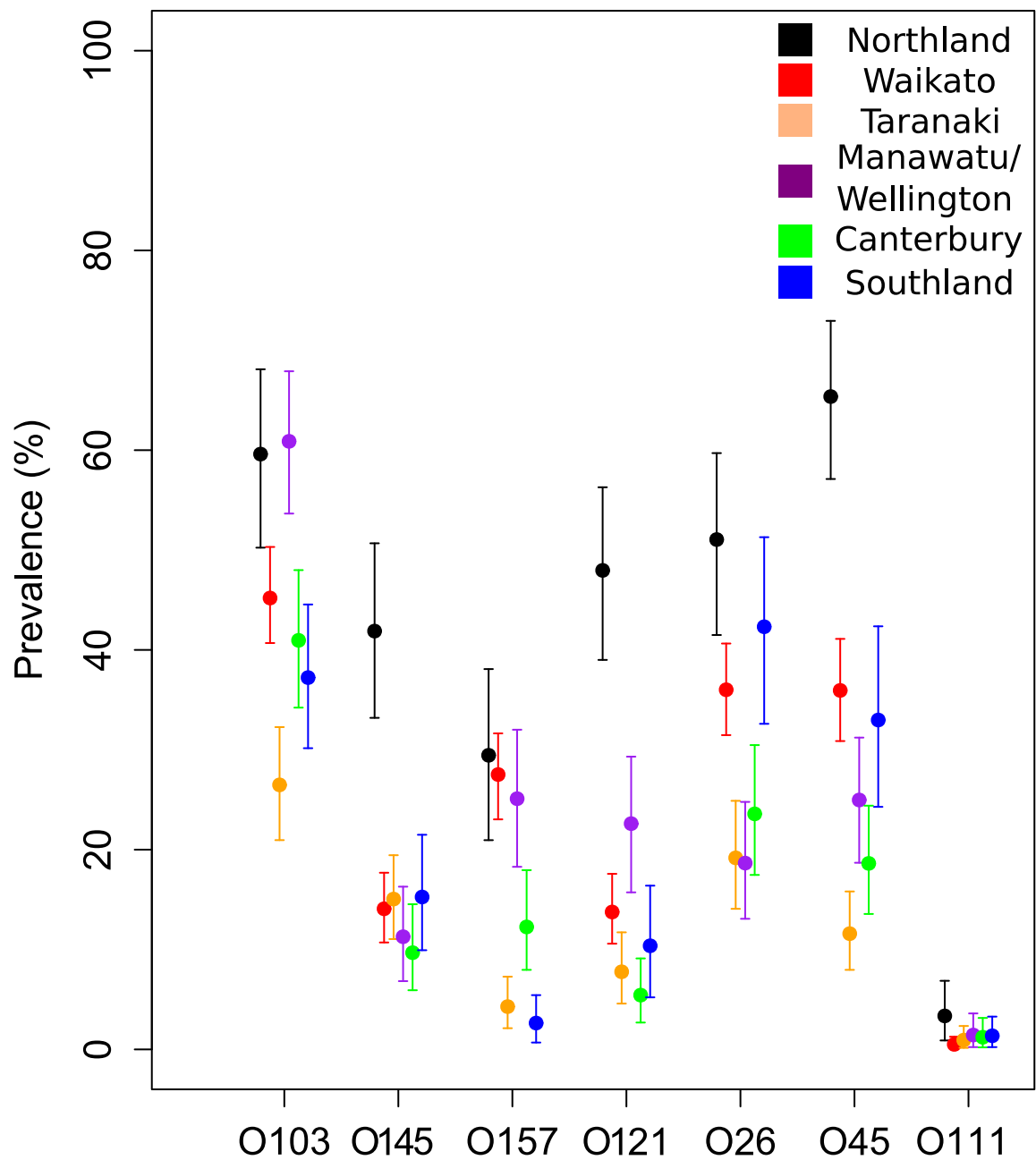
Figure 3-2: 'Top 7' serogroup prevalence (with 95% CI), including both STEC and non-STEC, detected in calves (n=1508) by region, using latent class analysis of NeoSEEK and RT-PCR results
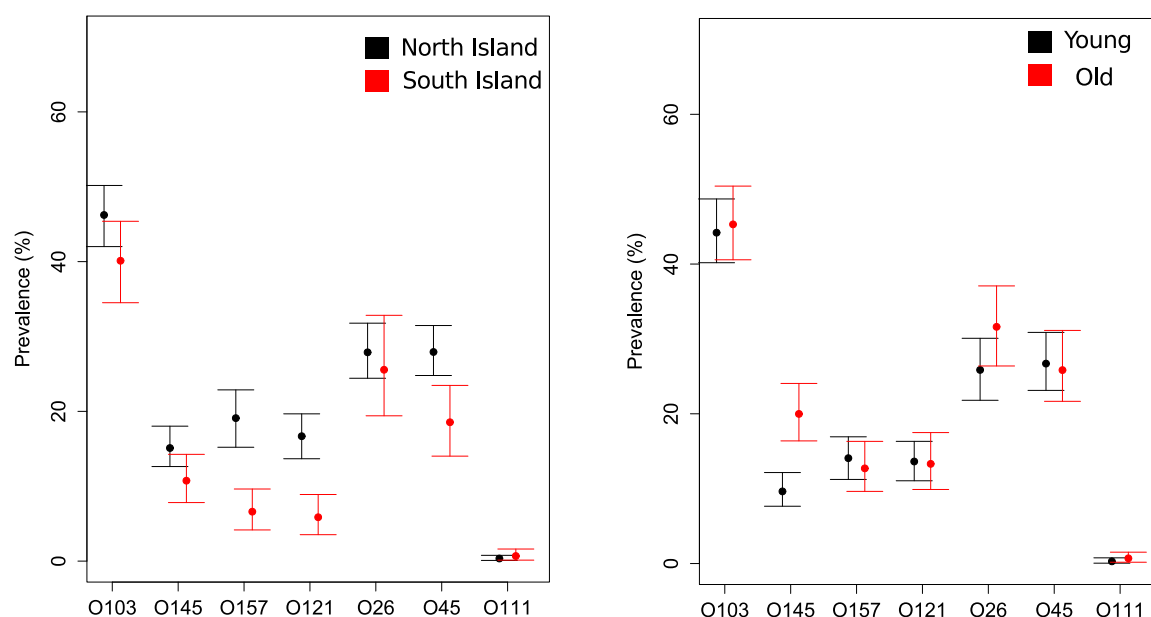
Figure 3-3: 'Top 7' serogroup prevalence (with 95% CI), including both STEC and non-STEC, detected in calves (n=1508) by island and age (young, 2 to 9 days; old, 10 to 21 days), using latent class analysis of NeoSEEK and RT-PCR results

There were notable differences in estimated serogroup prevalence between groups. Northland, Manawatu-Wellington, and Waikato had high prevalence of several serogroups compared to other regions, particularly serogroup O26 and O45 *E. coli* (Figure 3-2). Prevalence between older and younger calves was similar, but older calves had a higher prevalence of O145 and O26 serogroups (Figure 3-3). Finally, the North Island had a higher prevalence of most serogroups, with the exception of O26, which was similar between the two islands (Figure 3-3).

4.3.42 'Top 7' STEC detection via culture-independent methods

NeoSEEK detected 20.3% (95% CI 16.1-24.5) of the calves on 75% (76/102) of the dairy farms as positive for at least one of the 'Top 7' STEC (Table 3-3). NeoSEEK identifies both the presence of a 'Top 7' serogroup, as well as the presence of *eae* and *stx* genes, within the same serogroup. All 'Top 7' STEC, except for STEC O121, were detected in

samples taken from the New Zealand dairy farms tested. The highest estimated STEC

prevalence at the farm and calf level was STEC O145 and STEC O26, while STEC O111

was only detected in recto-anal mucosal swabs (RAMS) from three calves located on

one farm in the Northland region. Prevalence maps illustrate the regional variability of

prevalence of 'Top 7' STEC in New Zealand (Figure 3-4). 'Top 7' STEC prevalence

varied between serogroups, with STEC O26 more commonly detected in the South

Island (Canterbury and Southland), and a much higher prevalence of STEC O45

detected in Northland compared to other regions.

Table 3-3: Farm (n=102) and calf (n=1,508) level prevalence of the 'Top 7' STEC on New Zealand dairy farms

	STEC O103	STEC O121	STEC O111	STEC O145	STEC O157	STEC O26	STEC O45	Any 'Top 7' STEC[#]
Calves (number +ve)	75	0	3	148	29	109	44	306
% +ve	5.0%	0%	0.2%	9.8%	1.9%	7.2%	2.9%	20.3%
[95% CI]	[2.7-7.2]		[0.0-0.6]	[6.7-12.9]	[0.5-3.3]	[4.5-9.9]	[1.2-4.7]	[16.1-24.5]
Farms (number +ve)	36	0	1	44	15	23	18	76
% +ve	35%	0%	1%	43%	15%	23%	18%	75%

[#] The detection of at least one of the 'Top 7' STEC in an individual calf. 408 instances of 'Top 7' STEC were detected, but some calves shed multiple STEC serogroups: 1 serogroup (n=217), 2 serogroups (n=76), 3 serogroups (n=13) (Appendix Table 4)
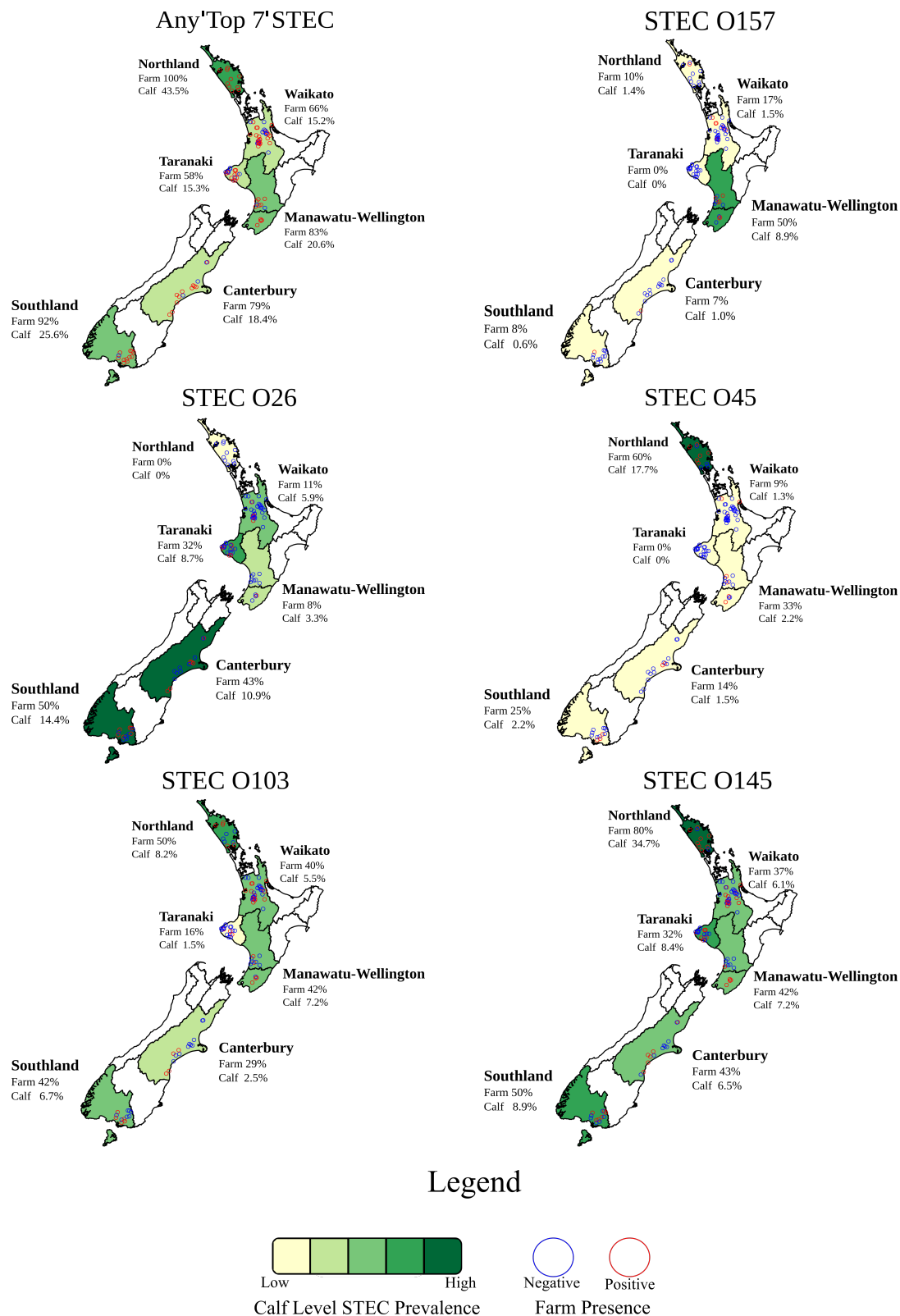
Figure 3-4: Calf (n=1,508) and farm (n=102) level prevalence of the 'Top 7' STEC on New Zealand dairy farms by region (n=6)

The virulence genes used to define 'Top 7' STEC, *stx* and *eae,* were common in calf samples; NeoSEEK detected *stx* in 70.5% of the calf samples, and *eae* in 57.6% of calf samples. Both *eae* and *stx* genes were detected in 45.4% of calf samples, however it is important to note that this did not necessarily indicate a 'Top 7' STEC was present. *stx* was detected in at least one calf sample from all farms in the study, while *eae* was detected in at least one calf sample from 101 of the 102 farms. These detections were based on only a single target gene, therefore they could be present in different *E. coli* bacteria; these data are provided for comparison with PCR or RT-PCR detection of virulence genes in enrichment broths in other research studies. Statistical analysis revealed strong clustering of 'Top 7' STEC positive calves within pens, and some strong clustering of calves on farms, most notably with the STEC O26, STEC O157, and STEC O45 serogroups (Table 3-4).

Table 3-4: Intraclass correlation (ρ) values of STEC using farm (n=102) and calf pen (n=267) as a random factor

	STEC O103	STEC O145	STEC O157	STEC O26	STEC O45	Any 'Top 7' STEC
Farm	0.13	0.29	0.61[*]	0.68[*]	0.62[*]	0.24
Calf pen	0.57	0.60[*]	0.71[*]	0.79[*]	0.77[*]	0.34

[*]Strong clustering observed

Calf and farm level risk factors were evaluated for the three most prevalent STEC serogroups (STEC O26, STEC O103, and STEC O145) and the presence of any 'Top 7' STEC. Due to the low calf-level prevalence of STEC O157 (n=29 calves), STEC O45 (n=44 calves), and STEC O111 (n=3 calves), it was not possible to create a final model using the same statistical technique for these serogroups, therefore significant risk factors were not identified. Region, higher humidity measured inside the calf pen

compared to outside the calf housing area, older calf age, and increased number of calves in a pen were all identified as significant risk factors for the presence of any 'Top 7' STEC (Table 3-5). Individual STEC serogroup analysis revealed increased number of calves in a pen (STEC O26; Table 3-6), increased pen humidity and a high ammonia presence (determined subjectively) in a pen (STEC O103; Table 3-7), and region, increased age, and increased pen humidity (STEC O145; Table 3-8) as significant risk factors.

Table 3-5: Logistic mixed effects regression model of factors associated with prevalence for any 'Top 7' STEC

Factor	OR	95% CI	p-value
Humidity: Difference between inside pen vs. outside the calf housing area (increase in 1% relative humidity)	1.09	1.02, 1.16	0.006*
Region (compared to Northland)			0.001†
Waikato	0.09	0.03, 0.29	<0.001*
Taranaki	0.11	0.03, 0.39	<0.001*
Manawatu-Wellington	0.23	0.06, 0.87	0.030*
Canterbury	0.19	0.05, 0.72	0.014*
Southland	0.30	0.08, 1.13	0.076
Number of calves in calf pen: Increase of one calf	1.04	1.01, 1.07	0.003*
Temperature: Difference between inside pen vs. outside the calf housing area (increase of 1°C)	1.20	0.96, 1.49	0.114#
Age: Young calves (2 to 9 days of age) vs. older calves (10 to 21 days of age)	0.43	0.27, 0.68	<0.001*

*Significant variable (p<0.05)
#Confounding factor for calf pen humidity, left in model
†Likelihood-ratio test of factor
Random effects variance: Calf pen within Farm (Variance = 1.09), Farm (Variance = 1.34)

Table 3-6: Logistic mixed effects regression model of factors associated with prevalence for any STEC O26

Factor	OR	95% CI	p-value
Number of calves in calf pen (increase of 1 calf)	1.07	1.02, 1.13	0.012*

*Significant variable (p<0.05)
Random Effects Variance: Pen within Farm (Variance = 3.26), Farm (Variance = 33.36)

Table 3-7: Logistic mixed effects regression model of factors associated with prevalence for any STEC O103

Factor	OR	95% CI	p-value
Humidity: Inside calf pen vs. outside the calf housing area (increase of 1% relative humidity)	1.02	1.01, 1.03	0.002*
High ammonia in calf pen (subjective measurement)	1.17	1.00, 1.36	0.047*

*Significant variable (p<0.05)
Random Effects Variance: Pen within Farm (Variance = 4.49), Farm (Variance = 0.90)

Table 3-8: Logistic mixed effects regression model of factors associated with prevalence for any STEC O145

Factor	OR	95% CI	p-value
Region (compared to Canterbury)			0.002#
Northland	1.38	1.15, 1.66	<0.001*
Waikato	0.98	0.84, 1.14	0.805
Taranaki	1.01	0.85, 1.20	0.886
Manawatu-Wellington	1.03	0.86, 1.24	0.742
Southland	1.02	0.85, 1.23	0.811
Number of calves in calf pen (increase of 1 calf)	1.04	0.999, 1.08	0.059†
Age: Young calves (2 to 9 days of age) vs. older calves (10 to 21 days of age)	0.89	0.84, 0.95	<0.001*

*Significant variable (p<0.05)

#Likelihood-ratio test of variable as a whole

†Factor left in model as it approaches significant value

Random Effects Variance: Pen within farm (Variance = 1.56), Farm (Variance = 2.69)

4.3.43   Bacterial isolation of *E. coli* serogroup O26 and O157

A total of 31 STEC O157 isolates, 123 STEC O26 isolates, and 69 non-toxigenic O26 isolates were retrieved from 109 calf faecal enrichment broths. The results of bacterial isolation of *E. coli* O157 and O26 *E. coli* serotypes from calf faecal enrichment broths are shown in Table 3-9, where results are based on the successful recovery or failure of recovery of at least one isolate from a calf faecal enrichment broth. In some cases, more than one isolate were recovered from a calf faecal enrichment broth, but this is not reflected in Table 3-9.

Table 3-9: Bacterial isolation of STEC and non-STEC isolates of serogroup O157 and O26 from faecal calf enrichment broths

Serogroup	Number of samples detected as STEC by NeoSEEK	Isolate Recovered from Calf Sample*	STEC Isolate Recovered*	Overall STEC Recovery
O157	29	14/29 (48%)	14/14 (100%)	14/29 (48%)
O26	109	70/109 (64%)	49/70 (70%)	49/109 (45%)

*At least one isolate was recovered from the enrichment broth

4.3.44 Whole genome sequencing (WGS) of serogroup O26 bacterial isolates

WGS data of serogroup O26 isolates (n=66, 45/66 STEC O26) from 24 sheds on 18 farms in five regions of New Zealand were processed using the Nullarbor pipeline and the Center for Genomic Epidemiology output, to evaluate the core genome, accessory genome, virulence genes, and antibiotic resistance genes (218, 220). The core genome (Figure 3-5) and accessory genome (Figure 3-6) were annotated with region, antimicrobial resistance gene class (n=1), and virulence gene (n=26) presence or absence. Clear clustering of STEC O26 isolates (n=45) apart from non-toxigenic isolates (n=21) was visible in both Figure 3-5 and Figure 3-6, but no obvious clustering by region was visible. The heatmap of virulence genes detected (n=26) indicated that STEC O26 and non-toxigenic O26 had similar virulence gene profiles (Figure 3-5, Figure 3-6). Antimicrobial resistance gene detection was rare, with only aminoglycoside resistance class genes detected (*strA, strB, aph(3')-IIa-like*) in eight isolates from the Manawatu-Wellington and Canterbury regions. All genomes sequenced from O26 bacterial isolates retrieved from calves in this study were identified as multilocus sequence type 21 (ST-21), and serotype O26:H11.
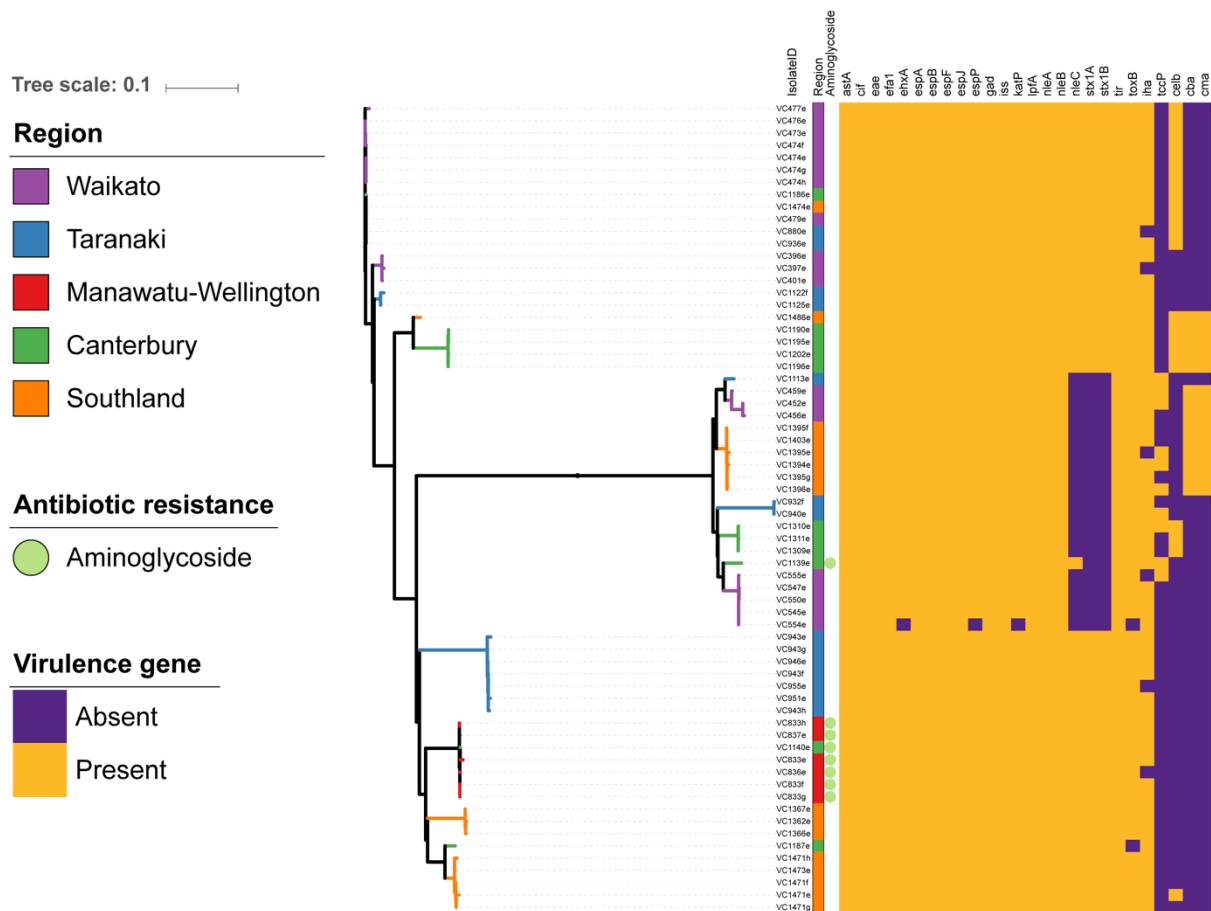
Figure 3-5: Maximum-likelihood core genome tree of serogroup O26 calf isolates (n=66), annotated with region (n=6), antibiotic resistance gene class (n=1), and virulence genes (n=26)
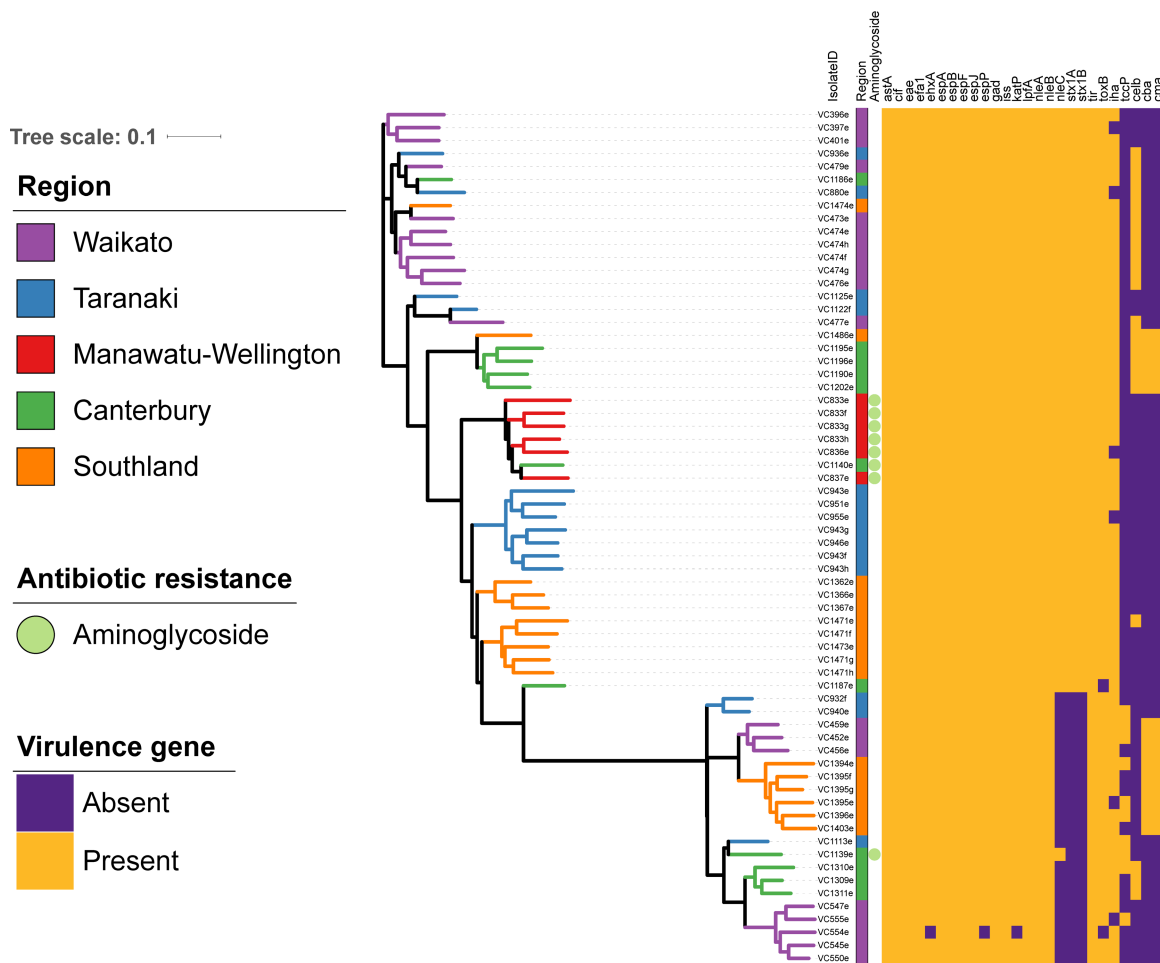
Figure 3-6: Maximum-likelihood accessory genome tree of serogroup O26 calf isolates (n=66) annotated with region (n=6), antibiotic resistance gene class (n=1), and virulence genes (n=26)

SNP analysis between serogroup O26 isolates indicated that the same clonal strain existed in calves in the same pen and the same farm, while strains between farms were different (Appendix Figure 1). For analysis of isolates from calves (n=42) on the same farm (n=14), as well as in the same pens (n=20), 0 to 29 SNPs separated all isolates. A subset of calves (n=5) had multiple isolates (n=19) sequenced from the same animal; only 0 to 17 SNPs separated isolates retrieved from the same animal sample. Two exceptions were noted in

the analysis, where two calves had markedly different (214 SNPs, 223 SNPs) O26 strains compared to other calves in the same farm and pen, indicating multiple serogroup O26 strains were present in the farm environment at the same time.

PERMANOVA analysis was used to compare region and farm with the variability of the core genome (SNP distance), accessory genome (presence or absence of accessory genes), and virulence genes (presence or absence of virulence genes) (Table 3-10). Farm was a significant predictor of variability (69.7-88.5%), indicating that the majority of the genetic variability at the core, accessory, and virulence gene level could be associated with each calf's presence in a specific farm environment. The importance of farm was further evaluated in hierarchical cluster plots (Figure 3-7), where a clear differentiation based on farm is visible, with the exception of farms which contain both *stx* positive and *stx* negative isolates. The hierarchical cluster analysis of core, accessory and virulence gene profiles also separated *stx* positive and *stx* negative isolates into different clonal groups, despite the same multilocus sequence type (ST-21).

Table 3-10: PERMANOVA analysis of core genome (SNP distance matrix), accessory genome (presence or absence of accessory genes), and virulence genes by region (n=5) and farm (n=18)

Factor evaluated	Genomic dataset	df	Pseudo-F	p-value	Component of variation
	Core	4	1.9	0.0975	NS
Region	Accessory	4	2.69	0.0016	11.6%
	Virulence	4	1.36	0.245	NS
	Core	17	28.6	0.0001	88.5%
Farm	Accessory	17	9.3	0.0001	69.7%
	Virulence	17	24.7	0.0001	86.8%

Residual variation: core (11.5%), accessory (30.2%), virulence (13.2%)
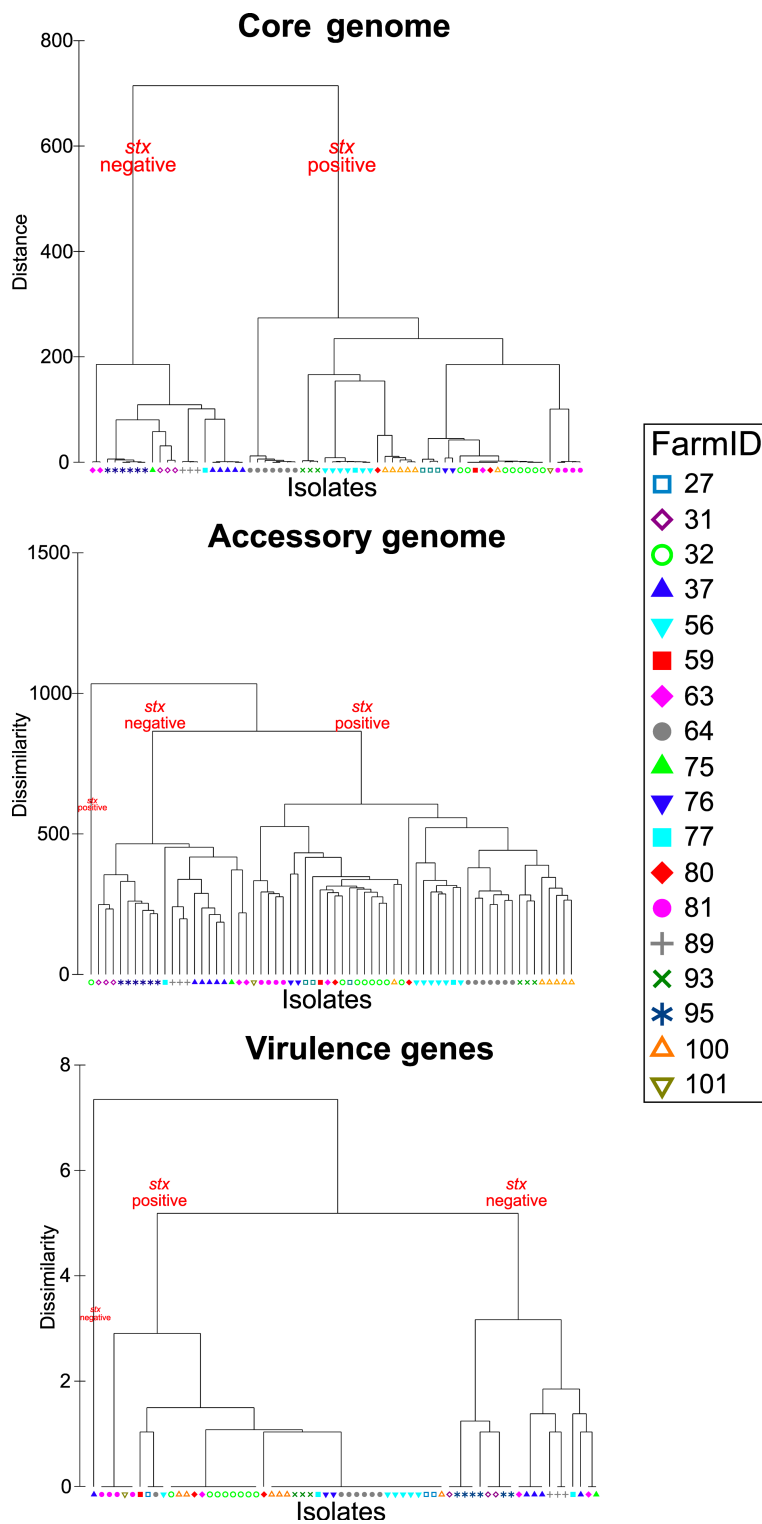
Figure 3-7: Hierarchical cluster trees of core, accessory, and virulence genes by farm (n=18)

3.5   Discussion

This study utilized an established molecular method that distinguishes STEC and non-STEC variants, along with random stratified sampling to estimate the national prevalence of the 'Top 7' STEC in young calves on dairy farms throughout New Zealand. Statistical analyses evaluated risk factors for positive prevalence in calves, while WGS and further statistical analysis determined the similarity of 'Top 7' STEC isolates between calves in a shared environment.

Systematic review and meta-analysis estimated an 8.7% prevalence of STEC (both *eae* and *stx* present in a single bacterium) in calves from 19 countries (128). A comprehensive national prevalence study of cattle and calves at 31 Australian processing plants showed a 6.3% prevalence of STEC O157, with a 1.7% prevalence for the other 'Top 7' STEC using culture methods (213). This Australian study also found that veal calves had the highest potential STEC prevalence (12.7%) using PCR methods, compared to young beef, young dairy, and adult cattle, with 51% of all samples testing positive for both *eae* and *stx* virulence markers (213). Our results indicated a higher 'Top 7' STEC prevalence of 20.3% in young calves; our use of a culture independent diagnostic test may have increased the sensitivity of detection of STEC.

Several results in our analysis suggested that STEC transmission occurs between calves or within the immediate calf pen environment: strong clustering of 'Top 7' STEC positive calves observed in pens for STEC O26, STEC O157, and STEC O45; increased risk of 'Top 7' STEC prevalence with increasing numbers of calves in a single pen; and clonal strains of serogroup O26 *E. coli* observed in specific farms and pens. In a controlled transmission study, a calf infected with a low dose of STEC O157 began

shedding the bacteria within six days, and that calf subsequently colonised all other calves in the same pen within four to 11 days after the initial calf began shedding (173). A separate comparison study of calves housed in individual pens versus an open group pen showed that a single calf inoculated with a control STEC strain in a group pen infected all other calves in that group over 10 days (130). Modelling studies have deduced a R_0 of 4.3 to 7.3 for STEC O157 in young calves from both natural and induced infection, suggesting calves in shared environments infect numerous other individuals when shedding (171, 174).

Our WGS analysis indicated that *stx* positive and *stx* negative *E. coli* O26 form distinct clones with divergent core, accessory and virulence gene profiles. Further epidemiological analysis also demonstrated that unique *E. coli* O26 clones disseminate among calves in a farm environment. PERMANOVA results indicated that farm, but not region, was a significant predictor of genetic variability (Table 3-10). The lack of similarity among strains in the same region, as well as the difference between isolates on farms, suggests relatively low transmission between farms in the same region. Only a minority of farms sampled in this study brought animals from outside the farm onto their property in the past two calving seasons: 7/102 farms brought in calves, while 15/102 farms had brought in cows. It is more likely that once established, specific strains proliferate in farms, leading to transmission between animals on the same farm. This finding has been reflected in other studies, where STEC strains isolated from calves from the same pen showed low variability, indicating high within-pen transmission (169). Unique STEC O157 lineages also proliferated among cattle on US dairy farms with a high STEC O157 prevalence (176). SNP analysis indicated that STEC O157 populations were dominated by a single clonal type on farm, but differences

occurred between farms, and some clonal types were still present during resampling 11 months later (176). Pulsed field gel electrophoresis (PFGE) analysis of O26:H11 isolates (n=11) on three Australian farms also found unique strains at each farm (166).

Increased relative humidity inside the pen environment compared to outside the calf housing area was associated with increased 'Top 7' STEC prevalence. Higher humidity has been associated with increased risk of shedding STEC O157 (225), but it is unclear whether this is due to environmental factors that would benefit bacterial growth, or high humidity causing stress of the animal. The increase in STEC prevalence with calf age may be associated with the duration of STEC exposure within the pen. The longer the calf is present with other infected animals and in a STEC contaminated environment, the increased likelihood of STEC ingestion and colonisation.

The calf pen environment is an important potential intervention point. Decreasing the number of young calves in pens is a practical intervention that may decrease STEC carriage. On average, thirteen calves were present in the pens in this study, with a range from two to eighty calves. This may also have animal welfare benefits. Recent legislation in New Zealand has focused on young calf welfare, and mandatory management changes could lead to opportunities for interventions at the farm level (226). Individual outdoor calf hutches, though used in other countries, are not widely used in New Zealand and may not be a realistic intervention for dairy farmers from either a time management or an economic viewpoint.

Limitations of this study included the cross-sectional study design that estimated STEC prevalence based on a single sampling event. It is well documented that calves may shed STEC intermittently, showing daily or even hourly variations (181, 182). By

sampling many calves from multiple pens on each farm, we estimated the farm level prevalence, as well as the proportion of calves shedding any 'Top 7' STEC on a dairy farm at a single point in time. Our use of a culture-independent diagnostic test (NeoSEEK) for this epidemiological study may have led to false positives due to a lower than 100% specificity compared to culture; PCR/MALDI-TOF techniques may have detected DNA rather than viable STEC bacteria. However, our study was interested in the prevalence of carriage of STEC in very young calves on dairy farms, and the presence of non-viable STEC bacteria may indicate colonisation of these calves. We evaluated the assay on New Zealand 'Top 7' STEC, and several other USDA studies in the United States have shown successful bacterial isolation of 84% (61/73) 'Top 7' STEC (214), and 55.7% (305/548) of non-STEC (202) following 'Top 7' STEC detection using the NeoSEEK assay. The New Zealand Ministry of Primary Industries has approved and utilized NeoSEEK as part of the regulatory testing and holding programme for veal beef exports to the United States, and the use of the NeoSEEK assay in a research context was beneficial for this epidemiological research.

3.6   Conclusion

This cross-sectional study of young calves on New Zealand dairy farms identified the widespread presence of 'Top 7' STEC. Future work using similar molecular confirmation methods, along with WGS, will permit the evaluation of the transmission dynamics of the 'Top 7' STEC on New Zealand dairy farms by sampling calves, cows, and their immediate environment throughout the calving season. Data from this research will provide further information as to the importance of specific environmental sources of infection for calves, as well as the persistence and spread of STEC throughout the dairy farm environment.

Practical and economic factors are often key drivers influencing the uptake and adoption of on-farm interventions by dairy farmers. While the use of vaccines or dietary supplements may decrease STEC O157 shedding in cattle (227), there is currently limited economic incentive for New Zealand dairy farmers to allocate time and money to prevent a bacterium colonising what are considered 'surplus' relatively low-value animals with no clinical signs. STEC and other *E. coli* are considered part of the normal bovine microbiota, therefore elimination of STEC from a herd and farm environment may be an unrealistic goal. Previously validated on-farm intervention strategies that are easily adopted, cost-effective, and that target mutual critical control points for several pathogens (e.g. STEC, *Campylobacter*, *Salmonella*, and *Cryptosporidium*), could form the basis of multiple agent control methods to reduce the overall level of zoonoses. This could impact overall animal prevalence levels and reduce the likelihood of human infection. Given STEC are found in cattle throughout the world, focusing on methods to decrease human exposure by minimising the presence of STEC in food and minimising environmental exposure is likely to be more beneficial than attempting to eliminate the presence of STEC in ruminant reservoirs.

The findings of this study provide important base-line data regarding the national prevalence of a zoonotic pathogen on New Zealand dairy farms. Future goals for STEC research should be multi-modal, addressing issues that could benefit the meat industry and protect public health using social science, epidemiology, and molecular biology.

3.7   Acknowledgments

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS

[To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis]

We, the candidate and the candidate's Principal Supervisor, certify that all co-authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Andrew Springer Browne

Name/Title of Principal Supervisor: Distinguished Professor Nigel French

Name of Published Research Output and full reference:
Browne, AS, Midwinter, AC, Withers, H, Cookson, AL, Biggs, PJ, Marshall JC, Benschop, J, Hathaway, S, Haranac, R, Nisa, S, Rogers, L, and French, NP. Evaluation of transmission dynamics and presence of Shiga toxin-producing E. coli in animals and their environment on New Zealand dairy farms, and potential contamination of veal beef. In preparation for Applied and Environmental Microbiology.

In which Chapter is the Published Work: 4
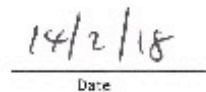
Please indicate either:

* The percentage of the Published Work that was contributed by the candidate:

and / or

* Describe the contribution that the candidate has made to the Published Work:
AS Browne designed the study with input from all supervisors.
AS Browne performed 80% of the field work, with assistance from H Withers and Tessa Handcock, and 80% of the laboratory work (S Nisa and L Rogers performed initial laboratory processing while AS Browne was in the field). AS Browne drafted the manuscript and created all figures, with input from supervisors.

Candidate's signature

13/2/18
Date

Principal Supervisor's signature

14/2/18
Date

Little fly,
Thy summer's play
My thoughtless hand
Has brushed away.

Am not I
A fly like thee?
Or art not thou
A man like me?

For I dance
And drink and sing,
Till some blind hand
Shall brush my wing.

If thought is life
And strength and breath,
And the want
Of thought is death,

Then am I
A happy fly,
If I live,
Or if I die.

The Fly, *Songs of Innocence*, William Blake

4 Evaluation of transmission dynamics and presence of Shiga toxin-producing *E. coli* (STEC) in animals and their environment on New Zealand dairy farms, and the potential impact on contamination of veal carcasses during slaughter and dressing

4.1 Abstract

This longitudinal study investigated factors associated with intestinal carriage and hide contamination of calves by Shiga toxin-producing *Escherichia coli* (STEC) on dairy farms and in processing plants in New Zealand. Management, environmental, and transport factors were evaluated for their potential impact on intestinal colonization and hide contamination of calves, as well as initial contamination of veal meat carcasses during slaughter and dressing. Animal and environmental samples (n=2580) were collected from six farms and three meat processing plants in the Waikato region over five sampling periods during each 2015 and 2016 Spring calving season.

Following screening for virulence genes (*eae, stx1, stx2* by RT-PCR), a selection of potential STEC positive samples (*eae* and *stx1* or *stx2*) were submitted for NeoSEEK analysis (n=1018). Of the submitted samples, 15.7% were positive for one or more of the 'Top 7' STEC. 'Top 7' STEC were identified in all 17 sample source types, including both environmental and animal-derived samples, with a prevalence range between 3% and

45%. A marked increase in 'Top 7' STEC prevalence was observed between calf hides on-farm (10.3% prevalence), and calf hides at processing plants (37.5% prevalence).

Factors associated with 'Top 7' STEC contamination of calf hides <u>on farm</u> included the number of calves in a shared pen environment (Odds Ratio (OR)=1.15, per increase of one calf in pen), the middle of the calving season compared to the early calving season (OR=5.6), and the proportion of colonised calves (recto-anal mucosal swabs, RAMS) in a shared pen environment on that day of sampling (OR=4.85, per 10% increase in prevalence). Calf colonisation with 'Top 7' STEC (RAMS) was associated with the proportion of 'Top 7' STEC hide contaminated calves in the same shared pen environment on the day of sampling (OR=2.01, per 10% increase in prevalence).

Factors associated with 'Top 7' STEC contamination of calf hides <u>at the processing plant</u> were visually apparent contamination of hides (OR=3.02) and a 'Top 7' STEC positive prevalence from a colonised (RAMS sample) cow on the farm on the same day (OR=9.4). Increasing the number of farms visited by the bobby calf transport truck was associated with an increased risk of 'Top 7' STEC contamination of pre-intervention carcasses (OR=1.09, per increase in one farm visited by a transport truck).

Whole genome sequencing was performed on a selection of 'Top 7' STEC bacterial isolates (n=40) recovered during this study. Core genome, accessory genome, and virulence gene analysis all showed that genetic diversity was low on individual farms, consistent with the presence of a resident strain disseminated widely across the farm environment. Bacteria isolated from cows, calves, feed and water sources, and the environment were essentially the same clonal strain (i.e. very similar genetically) on

each farm. Also, identical strains were shown to persist throughout a single calving season.

Prevalence data, molecular analysis, and statistical analysis all confirmed that transport and lairage of young veal calves were associated with increased contamination of hides at the processing plant, which is a likely transmission pathway for initial contamination of the veal carcass during slaughter and dressing.

Due to the large number of potential transmission routes identified in this study, preventing exposure of very young calves to STEC on dairy farms is likely to be difficult to achieve. Decreasing pen-occupancy rates of calves on farm may be a practical approach to reduce transmission between calves. Environmental sanitizers and topical disinfection of hides could be used to decrease the environmental contamination of STEC on farms and hide during transport and lairage. Using preliminary data generated from this study, it should be feasible to further quantify the impact of transport and lairage on levels of faecal carriage and hide contamination, and thus predict the value of interventions aimed at reducing initial contamination of veal carcasses during slaughter and dressing from these sources.

4.2 Introduction

Shiga toxin-producing *E. coli* (STEC) are globally estimated to cause 2.8 million cases of illness, 3,890 cases of haemolytic uremic syndrome (HUS), and 230 deaths annually (4). Illness and acute kidney failure are more common in young children, who express higher levels receptors for the Shiga toxin in their renal glomeruli (35). New Zealand has a relatively high incidence of notified STEC infection in humans, with 8.9 STEC cases per 100,000 population reported in 2016 (210). There has been a general increase in the incidence of STEC cases in New Zealand since it became a reportable disease in 1997 (210).

Cattle are the primary reservoir of STEC (72). In New Zealand, a previous cross-sectional study of randomly selected farms (n=102) from the six largest dairy regions revealed a relatively high prevalence of 'Top 7' STEC carriage of calves using recto-anal mucosal swabs (RAMS) based on NeoSEEK (PCR/MALDI-TOF) analysis (**Chapter 3**). A total of 20.3% (306/1508) of dairy calves under the age of 21 days on 75.0% (76/102) of the farms sampled were positive for at least one of the 'Top 7' STEC (O26, O45, O103, O111, O121, O145, O157) (**Chapter 3**). Previous research in New Zealand, as well as in other countries around the world, have identified calves as having a higher prevalence of STEC than adult cattle (127, 128, 136).

Having found STEC in raw ground beef, and following outbreaks associated with consumption of undercooked beef patties, the USA declared STEC O157 an adulterant of beef in 1994, followed by declaration of six additional serogroups (O26, O45, O103, O111, O121, O145) as adulterants in late 2011 (13). Mandatory testing of the additional six STEC serogroups began on March 5th, 2012 (13). These seven serogroups are collectively

known as the 'Top 7' STEC. In 2016-2017, 50% of New Zealand beef exports were sent to the United States, valued at $1.16 billion NZD (14). In New Zealand, very young calves born on dairy farms that are surplus to replacement needs of the herd are slaughtered between the ages of four to ten days; calves in this slaughter class are referred to as bobby calves. 'Top 7' STEC contamination of veal carcasses is considered reasonably likely to occur, therefore understanding colonisation and hide contamination of these calves may benefit food safety and market requirements for beef trade.

Given the relatively high prevalence of calves that are colonised by STEC in New Zealand, this study evaluated the risk factors that are likely to contribute to a higher initial level of contamination of veal carcasses. An understanding of these factors would add to the pool of knowledge that is drawn on in the design of slaughter and dressing interventions to minimise and hopefully eliminate the presence of STEC contamination in veal before it reaches the consumer. Determining the level of environmental contamination with 'Top 7' STEC is also important to determine transmission routes on farms that could be targeted for potential interventions. Whole genome sequencing (WGS) was used in this study to determine the transmission of particular STEC strains on farms to animals, as well as contamination of the environment, and to determine the between- and within-farm genetic variation of STEC in the farm environment over time. Understanding STEC transmission from cattle on farms offers the chance to identify potential mitigation strategies to reduce the on-farm prevalence of this pathogen at its source to protect public health and vital agricultural economic trade.

4.3 Materials and Methods

4.3.1 Farm and meat plant selection

Six farms in the Waikato region participated in the study during the 2015 and 2016

Spring calving seasons (July to September). Selection was determined using previous

'Top 7' STEC prevalence results from a cross-sectional study (**Chapter 3**) carried out in

2014, willingness to participate, and having very young calves processed by the meat

company Silver Fern Farms.

We enlisted the participation of Silver Fern Farms, who provided access to the Paeroa

and Waitoa veal processing plants for sampling, as well as logistical information

regarding calf transport to facilitate planning. A separate local meat processor also

participated in the study for two sampling periods during 2015, as some of the calves

from the selected study farms were sent to this plant for processing.

4.3.2 Sample collection

The majority of sampling was performed three times (early, middle, late) over the

calving season; these periods were specific for each farm in order to sample during the

period when the first 0-24% of calves were born (early), when 25-75% of calves were

born (middle), and when the final 76-100% of calves were born (late) for each calving

year. Sampling targeted very young calves (zero to three days of age), bobby calves and

replacement calves (four to ten days of age), and colostrum cows (post-partum cows

that had given birth within the past four days). Bobby calves that were sampled on-

farm were also sampled at processing plants. This entailed hide sampling pre-hide

removal, and carcass sampling immediately post-hide removal (pre-intervention

carcass). Selected farm environmental samples were collected before and after the

calving season including effluent, bird faeces, paddock overboot swabs, calf pen overboot swabs, and drinking water and feeding troughs.

Samples were collected using Amies swabs (Copan Diagnostics Inc., Brescia, Italy), sterile cellulose sponge swabs (EZ-Reach Sponge Sampler in 25 mL Buffered Peptone Water (BPW), World Bioproducts, Washington, USA), and overboot swabs (Envirobootie™, pre-moistened with double strength skim milk broth, Hardy Diagnostics, California, USA). For each farm visit, one milk filter was collected after the morning milking had concluded and placed in a dry sterile plastic bag.

Calves and cows were sampled using a rectal-anal mucosal swab (RAMS) technique with an Amies swab. After determining the age distribution of all calves, up to three calf pens were selected that allowed for the maximum number of calves (n=20) in the two age groups to be sampled, with equal numbers sampled per pen where possible. Sampling was prioritised to ensure the sample contained a maximum of ten bobby calves that were being shipped to the processing plant that day. A random number generator was used to select pens if more than three suitable pens were available for sampling. If more animals were present in a pen than needed, a spin-pointer mobile phone application was used to randomly select the first calf to be sampled, after which animals were chosen in an alternating manner in the clockwise direction, in proportion to the total calves in the pen. Calves were marked for selection and then again following sampling to maintain the random selection and prevent resampling. A maximum of ten cows that had calved within the past four days (colostrum herd) were selected for sampling. A random number generator was used to select the cows sampled while they were in the milking parlour during morning milking. Any animals

that appeared injured or sick based on visual clinical assessment, or under antibiotic treatment (e.g. cows for mastitis), were excluded from sampling.

Sponge swabs were used for sampling of the calf hide, cow udder, pre-intervention carcass, and environmental samples (colostrum feeder, water trough, concentrate feed trough, bird faeces, effluent). For calf hide and pre-intervention carcass sampling, calves were sampled on one side of the body, from the medial aspect of the carpus to the axilla, the entire ventral thorax and abdomen, and the medial aspect of the groin to the hock, using three back-forth passes in each area. These areas were selected due to their importance as likely sites of cross-contamination during slaughter and dressing of the carcass. The side of the body chosen for sampling was alternated between hide on farm, hide at processing plant, and pre-intervention carcass to avoid re-sampling of the same side.

Cow udders were sampled on the ventral aspect of the udder, lateral to each udder and between the teats, using three back-forth passes for each region. Colostrum feeders, water troughs, and concentrate feed trough samples were obtained by wiping the entire interior of each container with a sponge swab. Three to five bird faecal droppings were collected from each calf pen sampled on each visit using sterile forceps, and placed into the sterile cellulose sponge swab sampling bag. Effluent samples were obtained by inserting the swab into the effluent, at a designated location on each farm, chosen for its proximity to daily faecal outflow from the milking shed and safety of obtaining a sample. After sampling, each sponge swab was secured in its sterile bag and manually massaged to incorporate the 25mL of BPW into the sample.

Overboot swabs were used to sample calf pens, calving paddocks, and feed pads. Sterile plastic boot covers were placed over the boots before placing the overboot swabs to prevent cross-contamination. The sampler walked the entire perimeter of the sampling area, and then zig-zagged in equal transects across the area (six transects for the pens, three transects for the paddock and feed pad). During the transects in the calving paddock and feedpad, the sampler also walked the perimeter of high traffic areas, such as water troughs or feed areas.

All samples were placed in an ice filled insulated container in the field and shipped with fresh ice in insulated boxes to ^{m}EpiLab, Massey University, Palmerston North, New Zealand for processing.

4.3.3 Sample processing

Samples were shipped overnight on ice and processed at ^{m}EpiLab the day after collection. Processing methods were established for each sample type: Amies swabs, sponge swabs, overboot swabs, and milk filters. Amies swabs were placed in 20mL of modified Tryptone Soya Broth (mTSB, Oxoid Limited, Hampshire, United Kingdom). 25mL of double strength mTSB broth was added to the 25mL of BPW in each sponge swab Whirlflok bag, and processed in a stomacher machine for 120 seconds; 20mL of the liquid sample was then transferred into a sterile universal bottle (20mL volume). 100mL of normal strength mTSB was added to the overboot swab and milk filter samples, and processed in a stomacher machine for 120 seconds; 20 mL of the sample was transferred to a sterile universal bottle.

Each sample was enriched at 42°C for 15 to 18 hours. Enrichment broth was then stored with glycerol in a 4:1 ratio at -80°C, as well as processed for DNA extraction.

DNA processing in 2015 utilized the Kingfisher™ Flex Purification System per manufacturer's instructions. A subset of samples collected during 2015 was also processed using a double-wash boil preparation method, according to GeneSeek laboratory's instructions, to evaluate NeoSEEK detection using the two methods. Increased sensitivity was found using NeoSEEK with the double-wash boil preparation method. In 2016, all DNA processing was changed to a double-wash boil preparation method.

Each DNA sample was screened for virulence genes (*eae, stx1, stx2*) by RT-PCR using previously described methods (66). The limit of detection (LOD) of the virulence assay (RT-PCR screening for all three virulence factors) was estimated to be 9.9 x 10² colony forming units per mL (CFU/ml).

Samples selected for analysis using the NeoSEEK assay were based on the RT-PCR virulence gene results. All environmental and feed samples that tested positive for *eae* and at least one *stx* gene (*stx1* or *stx2*) were classified as "potential STEC" and submitted to Neogen for analysis. For animal samples (RAMS, hide, udder), a maximum of three potential STEC samples from one source type (e.g. RAMS) in a single pen or cow herd were submitted to Neogen. If more than three potential STEC samples of the same source type existed in a single pen or cow herd, three were chosen randomly using a random number generator. As part of the analysis of the cohort of bobby calves sampled at the processing plant, if one on-plant sample was detected as potential STEC, all samples from that calf (RAMS, hide on farm, hide at processing plant, and pre-intervention carcass) were sent for NeoSEEK analysis regardless of screening results. Selected DNA samples were shipped to GeneSeek Operations

(Lincoln, Nebraska, USA) on dry ice. Six percent (60/1018) of the samples submitted to NeoSEEK were judged by Neogen to require extra DNA processing, by repeating DNA purification using the Kingfisher™ Flex Purification System, including 29% (15/51) of the effluent sponge samples, and 49% (20/41) of the overboot paddock samples, likely due to the large amount of organic material present in these samples that may have inhibited the assay. The sample results after extra DNA processing were included in this study.

4.3.4   Data collection, database entry, and statistical analysis

A thermo-hygrometer was used to calculate the humidity and temperature within calf pens and outside of the calf building. Animal density was calculated by dividing the number of animals in each location (pen, calving paddock, feedpad) by the area (metres^2) of that location. Pen dimensions were measured manually, and calving paddock and feedpad areas were calculated using the area function on an eTrex GPS device (Garmin, Eastern Creek, Australia). All calves and cows sampled were given a hide cleanliness score (1 to 5) using the guidelines of the United Kingdom Food Standards Agency, where 1 = clean and dry, and 5 = filthy and wet (228). Calf transport information, including number of farms visited and number of calves in a truck, were obtained from processing plant records.

Weather data included the following variables: maximum wind gust, temperature, humidity, solar exposure, rainfall, runoff (excess rainfall not absorbed by saturated earth), and minimum grass temperature. Data were retrieved from the nearest weather station to each farm for the seven days prior to sampling, and accessed from the New Zealand National Climate Database (CliFlo, https://cliflo.niwa.co.nz/).

In order to evaluate potential transmission routes of STEC, positive 'Top 7' STEC (NeoSEEK) or potential STEC (RT-PCR) results from isolation sources (e.g. calving paddock, pen floor) were used as factors to evaluate associations with each outcome variable (Table 4-1). These were compiled into each individual dataset for analysis by attaching a positive or negative value on the specific farm visit on the day the animal was sampled.

Table 4-1: Outcome variables examined by statistical methods for both 'Top 7' STEC and potential STEC prevalence

Location	Outcome variable	Description
On farm	Calf colonisation (RAMS)	Recto-anal mucosal swab of calf on farm, positive or negative
	Calf hide on farm	Sponge swab of calf hide on farm, positive or negative
At processing plant	Calf hide at processing plant	Sponge swab of calf hide immediately post-stun on the processing line, positive or negative
	Pre-intervention carcass	Sponge swab of carcass after removal of the hide and before any decontamination intervention, positive or negative

All data were entered into a MySQL database. Statistical analyses were performed using a combination of univariable and multivariable regression models. The significance threshold was $p < 0.05$. Prevalence estimates with 95% confidence intervals were calculated using the 'effects' package (http://socserv.socsci.mcmaster.ca/jfox/) in R (222); in this method a generalised linear model for each output variable was created using farm, period, and/or sample source as a random effect in order to account for our clustered study design.

A random forest model (223) was used to identify important factors associated with positive 'Top 7' STEC or potential STEC prevalence outcome variables (Table 4-1). A logistic mixed effects model was used to evaluate the top ten factors identified by the random forest output for each outcome variable. This statistical technique allows for the use of random effects, which adjust for clustering (or lack of independence) of sample sources between farms. For our study, we used "farm" as a random variable

(i.e. random effect). Factors were sequentially removed and checked for confounding with other significant variables. Once a final model was determined, several biologically plausible risk factors were tested in the model to reassess their importance.

4.3.5 Bacterial isolation

Out of 1,018 samples screened with NeoSEEK, two hundred and four DNA samples were identified as positive for 'Top 7' STEC. Bacterial isolation was attempted for all positive samples. USDA-FSIS methods were used where immuno-magnetic separation beads (IMS; Abraxis, Warminister, USA) were utilized and plated onto three selective agars: CT-SMAC (O157; Fort Richard Laboratories, Auckland, New Zealand), CT-RMAC (O26; Fort Richard Laboratories, Auckland, New Zealand), and Rainbow® Agar O157 (O45, O111, O103, O145; Biolog, Hayward, USA). CHROMagar™ STEC was also used for several serogroups (O45, O111, O103, O145; CHROMagar Microbiology, Paris, France). In order to increase the success of bacterial isolate retrieval, frozen enrichment broth was also plated directly onto the same four agars, without the enrichment and IMS step.

4.3.6 Whole genome sequencing

One isolate was selected for whole genome sequencing from all enrichments which yielded a 'Top 7' serogroup isolate (n=40). DNA library preparation was completed in ^{m}EpiLab using the Nextera-XT protocol, and submitted to New Zealand Genomics Limited (University of Otago, Dunedin, New Zealand) for Illumina HiSeq sequencing. Raw sequence data was quality checked, assembled, and annotated using the Nullarbor bioinformatics pipeline (218). Virulence genes, antibiotic resistance genes, sequence type, and serotype were detected using the Center for Genomic

Epidemiology pipeline (220). A core genome alignment was created, indicating variability (single nucleotide polymorphisms, SNPs) among genes shared by all genomes. An accessory genome alignment, based on presence and absence of accessory genes that were not present in all genomes, was created with Roary (229). Distance (core genome SNP distance) and dissimilarity (accessory genome, virulence genes) matrices were created and evaluated in PERMANOVA and CLUSTER (PRIMER-E, Quest Research Limited, Auckland, New Zealand) with farm and isolation source as factors.

Raw sequence data and bacterial isolate metadata are available on the NCBI Sequence Read Archive (SRA) under BioProject PRJNA415994. Phylogenetic trees were visualised and annotated using iTOL software (221), and all figures were edited with Inkscape version 0.91 (https://inkscape.org/).

4.4 Results

Sampling was performed on all six farms at five time points during each year of the study (Table 4-2*)*.

Table 4-2: Sample numbers collected by farm and calving period (n=2580)

Year	Period	Farm 1	Farm 2	Farm 3	Farm 4	Farm 5	Farm 6	Total
	Pre Calving	5	5	5	5	5	5	**30**
	Early	65	92	80	69	82	73	**461**
2015	Middle	37	92	66	38	82	73	**388**
	Late	43	59	47	36	67	83	**335**
	Post Calving	5	5	5	5	5	5	**30**
	Pre Calving	5	5	5	5	5	5	**30**
	Early	69	70	79	55	92	83	**448**
2016	Middle	67	84	72	57	80	81	**441**
	Late	67	66	69	42	68	75	**387**
	Post Calving	5	5	5	5	5	5	**30**
Total		**368**	**483**	**433**	**317**	**491**	**488**	**2580**

A cohort of bobby calves (n=186) were sampled on-farm (RAMS and calf hide on farm), and then at the processing plant post-slaughter (calf hide at processing plant and pre-intervention carcass). These animals provided an important and separate cohort study to evaluate factors that impact STEC contamination from farm to processing plant.

In the following sections, prevalence is reported for both the 'Top 7' STEC (NeoSEEK) for relevant variables, and then for potential STEC (*eae* and *stx1* or *stx2* positive using the in-house RT-PCR assay).

4.4.1 'Top 7' STEC prevalence using NeoSEEK

The overall prevalence for any of the 'Top 7' STEC was 15.7% (5.9-25.5 95% CI), based on the subset of samples that was submitted for NeoSEEK detection following the initial screening by RT-PCR (n=1018). Prevalence varied by STEC serogroup, with STEC O103, STEC O26, and STEC O145 having the highest prevalence, while serogroups O103, O26, and O45 were the most common non-STEC samples detected (Table 4-3).

The relative proportions of STEC serogroups detected on the six farms in this study are similar to those detected during the 2014 cross-sectional study of STEC prevalence of RAMS samples from young calves on 102 New Zealand dairy farms (**Chapter 3**). STEC O121 was absent from the farm and processing plant environment, and STEC O111 was very rare.

Table 4-3: Prevalence of 'Top 7' STEC and non-STEC for all animal and environmental samples (n=1018)

Serogroup	O103	O111	O121	O145	O157	O26	O45
STEC prevalence (95% CI)	6.9% (0.1-13.6)	0.5% (0-2.6)	0%	4.4% (0-10.0)	1.3% (0-4.3)	5.5% (0-11.6%)	1.5% (0-4.7%)
Non-STEC prevalence (95% CI)	31.5% (19.0-44.0)	0.5% (0-2.4)	21.5% (10.5-32.6)	4.7% (0-10.4)	16.2% (6.3-26.1)	30.0% (17.6-42.3)	29.6% (17.3-41.8)

Prevalence estimates were stratified by farm, year, calving period, and isolation source. We used a modelling technique that incorporated random-effects terms (e.g. farm, calving period and source) in order to correct for bias due to the clustered sampling of animals and environments within farms (i.e. they were not independent observations). Although using this technique may provide more robust prevalence estimates and 95% confidence intervals, the results should be interpreted with caution due to the concurrent analysis of multiple sources, and variation in sampling of these sources between farms and time periods.

The 'Top 7' STEC prevalence estimates for each farm in the study are provided in Table 4-4. The farm level prevalence of 'Top 7' STEC varied significantly between farms (χ^2 (5) = 478, p < 0.0001).

Table 4-4: Prevalence estimates (95% CI) derived from a generalised linear model*
of 'Top 7' STEC in all animal and environmental sources by farm

Farm ID	Prevalence (95% CI)
F1 (n=156)	23.1% (15.9-32.4)
F2 (n=170)	9.1% (5.4-15.2)
F3 (n=182)	16.8% (11.3-24.2)
F4 (n=137)	4.7% (2.2-9.9)
F5 (n=183)	15.7% (10.4-23.0)
F6 (n=190)	17.4% (11.9-24.9)

*Calving period and sample source were included as random effects.

Prevalence data from both years and all five periods are shown in Table 4-5. The overall 'Top 7' STEC prevalence significantly increased between 2015 and 2016 (Table 4-5; χ^2 (1) = 5.4, p=0.02). This may have been influenced by the change in DNA processing from the Kingfisher™ Flex Purification System in 2015 to a double boil preparation in 2016, which was found to increase the sensitivity of the NeoSEEK assay in a subset of samples. The STEC prevalence was not significantly related to all five periods of the study (Table 4-5; Fisher's exact test, p = 0.87), or the three active calving periods (early, middle, late) of the calving season (χ^2 (2) =1.12, p=0.57).

Table 4-5: Prevalence estimates (95% CI) derived from a generalised linear model*
of 'Top 7' STEC in all animal and environmental sources on all study farms over
time (year and period)

	Variable	Prevalence (95% CI)
Year	2015 (n=395)	9.8% (6.2-15.3)
	2016 (n=623)	15.2% (10.3-21.9)
Period	Pre Calving (n=21)	11.2 (3.2-32.1)
	Early (n=346)	12.0 (7.6-18.4)
	Middle (n=276)	15.2 (9.8-22.9)
	Late (n=338)	12.7 (8.1-19.3)
	Post Calving (n=37)	11.2 (4.2-26.4)

*Farm and source of sample were included as random effects.

'Top 7' STEC were identified by the NeoSEEK method in all sample sources (n=17)
tested in this study (Table 4-6). Notably, RAMS samples from both calves and cows
represented the lowest prevalence, while environmental sources in the calf pen and
calf hides were higher. A significant difference was found between the prevalence of
sample source types (χ^2 (17) = 88.5, p<0.0001). Younger calves (age 1 to 3 days) had a
lower prevalence of colonisation (RAMS) (2.3%, n=45) and calf hide on farm (10.5%,
n=42) than older calves (age 4 to 10 days; RAMS prevalence 7.1%, n=156; calf hide on
farm prevalence 13.9%, n=151).

Table 4-6: Prevalence estimates (95% CI) derived from a generalised linear model*
of 'Top 7' STEC from sample sources (n=17)

Sample source	Prevalence (95% CI)
Bird faeces (n=20)	17.0% (5.9-40.3)
Paddock overboot swab (n=41)	12.8% (5.4-27.4)
RAMS: calf on farm (n=201)	4.7% (2.4-8.9)
RAMS: cow on farm (n=62)	2.7% (0.6-10.5)
Effluent (51)	16.1% (7.9-29.9)
Hide: calf on farm (n=193)	10.6% (6.3-17.3)
Feedpad overboot swab (n=5)	15.9% (1.9-64.4)
Pre-intervention calf carcass (n=122)	11.4% (6.3-19.7)
Milk filter (n=18)	19.6% (6.8-44.7)
Hide: calf at processing plant (n=128)	34.7% (24-47.2)
Calf pen overboot swab (early, middle, late calving period) (n=59)	26.0% (15-41.2)
Calf pen overboot swab (Pre or post calving period) (n=8)	11.9% (1.5-54.6)
Pen colostrum sponge sample (n=22)	11.2% (3.3-32.0)
Pen concentrates sponge sample (n=6)	45.2% (13.2-81.8)
Pen water trough sponge sample (n=11)	24.1% (7.0-56.9)
Dam udder sponge swab (n=61)	5.6% (2.0-14.9)
Water and concentrate sponge sample (pre or post calving period) (n=9)	10.3% (1.3-50.0)

*Farm and calving period were included as random effects.

4.4.2 Cohort study of bobby calves sampled on farm and at processing plants

A subset of calves (n = 118) from 186 bobby calves were selected, based on having at least one on-plant sample positive for virulence factor screens, to have all four samples (i.e. RAMS, hide on farm, hide at processing plant, pre-intervention carcass) from that calf submitted for NeoSEEK analysis (Table 4-7). These data indicate a significant difference in 'Top 7' STEC prevalence between sample source types (χ^2 (3) =51.32, p<0.00001); the increase of 10.3% to 37.5% of hide prevalence between the farm and processing plant is notable.

Table 4-7: Prevalence estimates (95% CI) derived from a generalised linear model*
of 'Top 7' STEC from calves (n=118) sampled from farm to processing plant

Sample source	Prevalence (95% CI)
RAMS: calf on farm (n=118)	5.0% (2.2-10.9)
Hide: calf on farm (n=118)	10.3% (5.6-18.1)
Hide: calf at processing plant (n=118)	37.5% (26.9-49.6)
Pre-intervention calf carcass (n=118)	12.7% (7.2-21.2)

*Farm and calving period were included as random effects

4.4.3 Potential STEC prevalence by in-house RT-PCR detection

Real time PCR (RT-PCR) testing of all sample enrichments (n=2580) was performed to
detect virulence factors (*eae*, *stx1*, *stx2*). While this method was mainly employed to
screen samples for submission to Neogen for the NeoSEEK assay, the data provided
allows for assessment of potential STEC (*eae* and *stx1* or *stx2*) on all samples collected
during the study.

 Over one-third of samples (37.3%) were screened as potential STEC (Table 4-8). *eae*
was detected in almost half the samples, and *stx2* was much more common than *stx1*.

Table 4-8: Prevalence of virulence factors (*eae, stx1, stx2*) and potential STEC (RT-PCR *eae* and *stx1 or stx2* positive) from all animal and environmental samples (n=2580)

	eae	*stx1*	*stx2*	Potential STEC
Samples	1205	269	1716	963
Positive (%)	(46.7%)	(10.4%)	(66.5%)	(37.3 %)

Potential STEC varied from 37.5% to 50.9% on the six farms in this study (Table 4-9). A significant difference between the prevalence values existed (χ^2 (5) =23.1, p<0.001).

Table 4-9: Prevalence estimates (95% CI) derived from a generalised linear model* of potential STEC (RT-PCR *eae* and *stx* positive) in all animal and environmental sources in each farm

Farm ID	Prevalence (95% CI)
F1 (n=368)	39.5% (30.8-48.9)
F2 (n=483)	37.5% (29.3-46.4)
F3 (n=433)	50.0% (41.0-59.1)
F4 (n=317)	48.0% (38.6-57.6)
F5 (n=491)	37.9% (29.7-46.8)
F6 (n=488)	50.9% (41.9-59.8)

* Calving period and sample source were included as random effects.

The potential STEC prevalence increased between 2015 and 2016 (Table 4-10), similar to 'Top 7' STEC prevalence values between the two years. This difference was also significant (χ^2 (1) =11.1, p<0.001). Overall the prevalence of potential STEC increased through the calving season (Table 4-10), but it should be noted that "Pre Calving" and "Post Calving" periods only contained environmental samples. The difference in prevalence was significant for all periods (χ^2 (4) =26.864, p<0.00001) and for the three calving periods (χ^2 (2) =11.06, p<0.004).

Table 4-10: Generalised linear model* prevalence (95% CI) of potential STEC (RT-PCR *eae* and *stx* positive) in all animal and environmental sources on all study farms over time (year and period)

	Variable	Prevalence (95% CI)
Year	2015 (n=1244)	36.7% (30.7-43.0)
	2016 (n=1336)	44.8% (38.4-51.4)
Period	Pre Calving (n=60)	21.9 (12.3-35.8)
	Early (n=909)	38.2 (31.5-45.3)
	Middle (n=829)	39.8 (33.0-47.1)
	Late (n=722)	46.4 (39.1-53.8)
	Post Calving (n=60)	55.4 (38.8-70.9)

*Farm and source of sample were included as random effects.

Potential STEC prevalence by sample source varied from 20.2% to 84.1% (Table 4-11). The highest prevalence values were noted in calf hides at the processing plant (75.9%) and environmental samples (effluent, paddock overboot swab, feedpad overboot swab; 74-84%). Younger calves (age 1 to 3 days; n=234) had a lower prevalence for RAMS (19.2%), and a similar prevalence for hide on farm (41.5%) compared to older calves (age 4 to 10 days; RAMS prevalence 28.8%; calf hide on farm prevalence 43.9%; n=319).

Table 4-11: Prevalence estimates (95% CI) derived from a generalised linear model*
of potential STEC (RT-PCR *eae* and *stx* positive) of sample sources (n=17)

Sample source	Prevalence (95% CI)
Bird faeces (n=60)	32.7% (21.4-46.4)
Paddock overboot swab (n=59)	70.2% (56.5-81.0)
RAMS: calf on farm (n=553)	23.7% (18.8-29.4)
RAMS: cow on farm (n=290)	22.2% (16.7-28.9)
Effluent (n=60)	84.1% (72.2-91.5)
Hide: calf on farm (n=553)	42.2% (35.6-49.1)
Feedpad overboot swab (n=7)	74.4% (34.6-94.1)
Pre-intervention calf carcass (n=186)	37.9% (29.6-47.0)
Milk filter (n=36)	49.2% (32.5-66.1)
Hide: calf at processing plant (n=186)	75.9% (67.8-82.4)
Calf pen overboot swab (early, middle, late calving period) (n=92)	66.1% (54.4-76.1)
Calf pen overboot swab (Pre or post calving period) (n=24)	33.7% (17.1-55.6)
Pen colostrum sponge sample (n=91)	20.7% (13.1-31.2)
Pen concentrates sponge sample (n=10)	52.3% (22.8-80.2)
Pen water trough sponge sample (n=23)	42.3% (23.5-63.5)
Dam udder sponge swab (n=290)	20.2% (15.0-26.6)
Water and concentrate sponge sample (Pre or post calving period) (n=9)	38.1% (20.3-59.8)

*Farm and calving period were included as random effects.

4.4.4 Cohort study of bobby calves sampled on farm and at processing plants

The cohort of young bobby calves that were sampled on farm as well as processing plants had similar patterns in potential STEC as the 'Top 7' STEC detection using NeoSEEK (Table 4-12). A notable increase between the calf hide at the farm (41.5%) and calf hide at the processing plant (76.2%) was observed. A significant difference between the prevalence values for the four sample sources was found (χ^2 (3) =102.78, p<0.00001).

Table 4-12: Prevalence estimates (95% CI) derived from a generalised linear model*
of potential STEC (RT-PCR *eae* and *stx* positive) of calves (n=186) sampled from
farm to processing plant

Sample source	Prevalence (95% CI)
RAMS: calf on farm (n=186)	22.4% (15.8-30.7)
Hide: calf on farm (n=186)	41.5% (32.6-51.2)
Hide: calf at processing plant (n=186)	76.2% (67.5-83.1)
Pre-intervention calf carcass (n=186)	39.0% (30.0-48.8)

*Farm and calving period were included as random effects

4.5 Factors associated with STEC hide contamination, pre-intervention carcass contamination, and carriage (RAMS) of calves

Risk factors were evaluated for two testing methods: 'Top 7' STEC positive (NeoSEEK assay) and potential STEC positive (RT-PCR *eae* and *stx* positive). Calf related outcome variables (RAMS on farm, hide on farm, hide at processing plant, and pre-intervention carcass) were evaluated against relevant factors.

4.5.1 Independent evaluation of outcome variables for calf colonization, hide contamination, and pre-intervention carcass contamination

The variables (Table 4-1) were independently examined, considering each as an explanatory variable, at the calf, calf pen, and farm level, using a generalised linear mixed model with farm as a random effect. The purpose of this evaluation was to test each variable on several levels, and then incorporate these biologically plausible variables in the analysis for all factors evaluated on farm and at processing plants.

For detection of 'Top 7' STEC, colonisation at the animal level was associated with an increased risk of hide contamination at the calf pen and farm level, but not at the individual level (a calf whose hide was contaminated was not necessarily colonised) (Table 4-13). Hide contamination on farm with 'Top 7' STEC at the animal level was

also strongly associated with an increased risk of a calf being colonised at all levels, particularly if any calf in the calf pen or farm was positive on the day of sampling. The presence of a hide contaminated with 'Top 7' STEC on the day of sampling was associated with increased risk of hide contamination at the processing plant, but not on the animal or calf pen level (Table 4-13). Hide contamination of the calf at the processing plant after the calf had been stunned was not a significant predictor of contamination of pre-intervention carcasses.

Table 4-13: Generalised linear mixed model of the interrelationships between the variables outlined in Table 4-1. Outcome variables are measured at the animal level, whereas explanatory variables are measured at the animal, calf pen, and farm level

Outcome variables location / sample type	Hypothesis (i.e. explanatory variable effect on outcome variable)	Level at which explanatory variable is measured	Outcome variable	Odds Ratio (95% CI)	p value
Farm / hide contamination	Effect of colonisation (RAMS) on calf hide on farm	Animal: Positive prevalence in same animal	'Top 7' STEC	NC#	0.721
			Potential STEC	2.35 (1.6, 3.5)	<0.00001
	Effect of colonisation (RAMS) on calf hide on farm	Calf pen: Presence or absence of a single positive in pen	'Top 7' STEC	11.4 (4.2, 31.0)	<0.00001
			Potential STEC	2.30 (1.5, 3.5)	<0.0001
	Effect of colonisation (RAMS) on calf hide on farm	Calf pen: Proportion of positives in pen	'Top 7' STEC	3.08 (1.8, 5.4)*	<0.0001
			Potential STEC	1.25 (1.2, 1.3)*	<0.00001
	Effect of colonisation (RAMS) on calf hide on farm	Farm: Presence or absence of a single positive on that farm visit	'Top 7' STEC	13.7 (4.81, 38.9)	<0.00001
			Potential STEC	3.98 (1.3, 12.0)	0.01

Farm / RAMS colonisation	Contaminated calf hide at farm on calf colonisation (RAMS)	Animal: Positive prevalence in same animal	'Top 7' STEC	1.40 (1.38, 1.41)	<0.000001
			Potential STEC	2.28 (1.5, 3.4)	<0.00001
	Contaminated calf hide at farm on calf colonisation (RAMS)	Calf pen: Presence or absence of a single positive in pen	'Top 7' STEC	37.3 (4.7, 298)	<0.00001
			Potential STEC	1.99 (1.1, 3.6)	0.02
	Contaminated calf hide at farm on calf colonisation (RAMS)	Calf pen: Proportion of positives in pen	'Top 7' STEC	2.01 (1.4, 3.0)*	<0.00001
			Potential STEC	NC#	0.08
	Contaminated calf hide at farm on calf colonisation (RAMS)	Farm: Presence or absence of a single positive on that farm visit	'Top 7' STEC	21.7 (2.6, 178)	0.0004
			Potential STEC	1.99 (1.1, 3.6)	0.04
Processing plant / hide contamination	Contaminated calf hide on farm on contaminated calf hide at processing plant	Animal: Positive prevalence in same animal	'Top 7' STEC	NC#	0.47
			Potential STEC	NC#	0.86
		Calf pen: Proportion of positives in pen	'Top 7' STEC	NC#	0.087

118

	Contaminated calf hide on farm on contaminated calf hide at processing plant		Potential STEC	1.15 (1.03, 1.28)*	0.011
	Contaminated calf hide on farm on contaminated calf hide at processing plant	Farm: Presence or absence of a single positive on that farm visit	'Top 7' STEC	2.9 (1.25, 6.73)	0.013
			Potential STEC	2.5 (1.12, 5.56)	0.025
Processing plant / carcass contamination	Contaminated calf hide at processing plant on pre-intervention carcass contamination	Animal: Positive prevalence in same animal	'Top 7' STEC	NC#	0.74
			Potential STEC	3.75 (1.6, 8.7)	0.02
	Contaminated calf hide at processing plant on pre-intervention carcass contamination	Processing plant: Proportion of contaminated calf hides	'Top 7' STEC	NC#	0.97
			Potential STEC	1.2 (1.1, 1.4)*	0.005
	Contaminated calf hide at processing plant on pre-intervention carcass contamination	Processing plant: Presence or absence of a single positive on that processing plant visit	'Top 7' STEC	NC#	0.43
			Potential STEC	NC#	0.73

* per 10% increase in calf pen or processing plant prevalence
#NC: not calculated due to non-significant finding

4.5.2 Factors associated with STEC contamination and carriage of calves on dairy farms

The extended analysis of the larger set of potential risk factors associated with 'Top 7' STEC and potential STEC carriage (RAMS) by calves on dairy farms is shown in Table 4-14. Significant risk factors from Table 4-13 were individually included in the generalised linear mixed models for the extended analysis, but only one level of the explanatory variables was included if significant, based on AIC (Aikaike information criterion) model evaluation. This was done to prevent the inclusion of the same sample result (i.e. one calf with a 'Top 7' STEC positive hide on farm) on multiple levels in the same model.

As indicated in Table 4-13, calves were more likely to be colonised when the proportion of calves within the same calf pen with contaminated hides increased (OR=2.01, per 10% increase in prevalence); this was the only significant association identified using the NeoSEEK dataset. Using the larger RT-PCR-based dataset, colonisation (RAMS) of a cow on the farm, contaminated milk filters, and contaminated colostrum samples were associated with a higher prevalence for potential STEC carriage in calf faeces (Table 4-14).

Table 4-14: Generalised linear models of 'Top 7' STEC and potential STEC prevalence of calf colonisation (RAMS) on farm

Outcome variable	Factor	OR	95% CI	p-value
'Top 7' STEC	Proportion of 'Top 7' STEC positive calf hide in same calf pen (per 10% increase in prevalence)	2.01	1.4, 3.0	0.0005*
Potential STEC	Potential STEC positive (RAMS) cow	2.23	1.1, 4.3	0.02*
	Potential STEC positive milk filter	1.88	1.2, 3.0	0.01*
	Potential STEC positive colostrum	1.56	0.9, 2.6	0.09#

*Significant variable (p<0.05)
#Confounding factor left in model
'Top 7' STEC random effects variance: Farm (0.08)
Potential STEC random effects variance: Farm (0.1533)

Generalised linear models including the extended set of explanatory variables were created for the contamination of the hide of calves on farms (Table 4-15). The proportion of 'Top 7' STEC positive colonised calves (RAMS sample) in the calf pen was associated with positive hide prevalence (OR=4.85, per 10% increase in calf pen prevalence). In addition, increased prevalence was noted in the middle of the calving period, compared to the early calving period (OR=5.56, middle compared to early calving period), and decreased prevalence was noted in the late period compared to the early period (OR=0.42, late compared to early calving period). Increasing numbers of calves in a shared pen environment led to increased risk of 'Top 7' STEC contamination of calf hides on farms (OR=1.15, per increase in one calf in the calf pen). Similar findings were noted for potential STEC, with the addition of a colonised cow (RAMS) present on the farm the same sampling day leading to an increased risk of hide contamination on farm (OR=3.35).

Table 4-15: Generalised linear model of 'Top 7' STEC and potential STEC prevalence for the hide of a calf on farm

Outcome variable	Factor	OR	95% CI	p-value
'Top 7' STEC	Proportion of 'Top 7' STEC colonised (RAMS) calves in same calf pen (per 10% increase in prevalence)	4.85	2.3, 10.0	<0.0001*
	Calving period:			0.002†
	Middle calving period compared to early calving period	5.56	1.6,19.4	0.007*
	Late calving period compared to early calving period	0.42	0.2,0.9	0.04*
	Number of calves in shared pen environment (per increase in one calf)	1.15	1.01,1.3	0.035*
Potential STEC	Potential STEC colonised (RAMS) calf on same farm visit	4.83	1.5,15.6	<0.009*
	Potential STEC colonised (RAMS) cow on same farm visit	3.35	1.9,5.9	<0.0001*
	Calving period:			<0.0001†
	Middle calving period compared to early calving period	2.1	1.5,2.9	<0.0001*
	Late calving period compared to early calving period	0.8	0.6,1.1	0.13

*Significant variable (p<0.05)
† Likelihood ratio test p-value of variable as a whole
'Top7' STEC random effects variance: Farm (0.48)
Potential STEC random effects variance: Farm (0.06)

4.5.3   Factors associated with contamination of calf hides at processing plants and pre-intervention calf carcasses

Risk factors for 'Top 7' STEC and potential STEC positive prevalence were determined for calf hides and pre-intervention calf carcasses at processing plants. Contamination of calf hides post-slaughter with 'Top 7' STEC was associated with a colonised (RAMS) cow on farm during the farm visit (OR=9.4), and increased visual contamination of the hide (faecal score 2 compared to faecal score 1, OR =3.02) (Table 4-16). The significance of increased visual contamination as a risk factor for calf hide contamination at the processing plant should be interpreted with caution, as the variable as a whole was deemed to not be significant using a likelihood-ratio test (p=0.14). The proportion of contaminated calf hides in calf pens on farms was

associated with increased potential STEC calf hide contamination at the processing

plant (Table 4-16).

Table 4-16: Generalised linear model of 'Top 7' STEC and potential STEC prevalence for the hide of a calf at the processing plant

Outcome variable	Factor	OR	95% CI	p-value
'Top 7' STEC	'Top 7' STEC colonised (RAMS) cow on the same farm visit	9.40	2.1,43.0	0.0008*
	Plant hide faecal score:			0.14†
	Plant hide faecal score is 2 compared to 1	3.02	1.1,8.3	0.03*
	Plant hide faecal score is 3 compared to 1	3.07	0.8,11.6	0.09
	Plant hide faecal score is 4 compared to 1	3.77	0.6,22.3	0.1
Potential STEC	Proportion of potential STEC positive calf hides in calf pen on the same day as the processing plant sampling (per 10% increase in prevalence)	1.15	1.03,1.28	0.03

*Significant variable (p<0.05)
† Likelihood ratio test p-value of variable as a whole
'Top 7' STEC random effects variance: Farm (0.21)
Potential STEC random effects variance: Farm (0.02)

Evaluation of risk factors for pre-intervention calf carcasses at the processing plant indicated increased number of farms visited by the bobby calf transport truck was associated with positive 'Top 7' STEC results (OR=1.09, per increase in one farm visited by transport truck). Visibly increased faecal contamination of hides at the processing plant, as well as an increased proportion of calves with contaminated hides on the processing line, was associated with an increased risk of potential STEC contamination of carcasses (Table 4-17).

Table 4-17: Generalised linear model of 'Top 7' STEC and potential STEC prevalence for the pre-intervention calf carcass at the processing plant

Outcome variable	Factor	OR	95% CI	p-value
'Top 7' STEC	Number of farms visited by calf truck	1.09	1.00,1.20	0.049*
	'Top 7' STEC positive calf hide sample on farm on the same farm visit	3.16	0.93,10.74	0.065#
	Number of calves in calf truck	0.99	0.97,1.00	0.065#
Potential STEC	Proportion of potential STEC positive calf hides from calves transported on the same transport truck (per 10% increase in prevalence)	1.20	1.05, 1.4	0.006
	Plant hide faecal score			0.01†
	Plant hide faecal score is 2 compared to 1	2.09	0.98,4.45	0.055
	Plant hide faecal score is 3 compared to 1	3.72	1.2, 11.2	0.02*
	Plant hide faecal score is 4 compared to 1	9.63	1.6, 56.4	0.01*

*Significant variable (p<0.05)
#Confounding factor left in model
† Likelihood ratio test p-value of variable as a whole
'Top 7' STEC random Effects Variance: Farm (Variance = 0.0)
Potential STEC random Effects Variance: Farm (Variance = 0.0)

4.5.4 Bacterial isolation

Individual bacterial isolates (n=101), up to four from a single sample, were retrieved from 40 frozen glycerol enrichment broths. The overall success of bacterial isolation of a detected 'Top 7' STEC serogroup from a sample was 19.6% (40/204), however some of these were non-pathogenic, therefore the overall success of bacterial isolation of STEC was 14.2% (29/204) (Table 4-18).

Table 4-18: Detection of 'Top 7' STEC by NeoSEEK and success of retrieval of bacterial isolates

Serogroup	Samples detected as 'Top 7' STEC by NeoSEEK	Isolate recovered from sample*	STEC isolate recovered*	Overall STEC recovery
O157	13	4/13 (30.8%)	4/4 (100%)	4/13 (30.8%)
O26	56	25/56 (44.6%)	24/25 (96%)	24 (42.9%)
O45	15	1/15 (6.7%)	0/1 (0%)	0%
O103	70	4/70 (5.7%)	0/4 (0%)	0%
O111	5	0	0	0%
O145	45	6/45 (13.3%)	0/6 (0%)	0%

*At least one isolate was recovered from the frozen enrichment broth

4.5.5   Whole genome sequencing (WGS) analyses of bacterial isolates

One 'Top 7' STEC serogroup isolate recovered from each sample from the six farms were whole genome sequenced (n=40). The subsequent analyses focused on the 25 isolates belonging to serogroup O26, which was the most prevalent and widely distributed serogroup retrieved from all farms (n=6) and several sources (n=8). This allowed for an analysis of the population structure of similar STEC within and between farms.

The population genetic analysis was conducted using three measures of genetic variation: the core genome (based on single nucleotide polymorphisms between core genes); the accessory genome (presence or absence of genes not present in all isolates); and virulence factor genes. Two potential sources of variation were considered: farm and isolation source (Table 4-19).

Table 4-19: PERMANOVA analysis of core genome, accessory genome, and virulence factors of O26 isolates (n=25) by farm and isolation source

	Factor evaluated	Df	Psuedo-F	p-value
Core genome	Farm (n=6)	4	17.17	0.0001
(n=1,974 SNPs)	Isolation source (n=8)	6	1.05	0.42
Accessory genome	Farm (n=6)	4	10.64	0.0001
(n=2,265 genes)	Isolation source (n=8)	6	0.996	0.47
Virulence genes	Farm (n=6)	4	19.53	0.0004
(n=28 genes)	Isolation source (n=8)	6	0.97	0.40

Variation for core genome: farm (83.9%), residual (16.1%)
Variation for accessory genome: farm (75.6%), residual (24.45)
Variation for virulence genes: farm (86.7%), residual (13.3%)

Farm was the strongest determinant of genetic variation; isolates recovered from the same farm were much more similar than isolates compared between farms. 'Farm' explained 75-87% of the genetic variation, whereas the sample type/isolation source explained very little of the variation once farm had been taken into consideration. This means that isolates from multiple sources on the same farm displayed a high degree of genetic similarity, consistent with widespread dissemination of clonal strains across multiple environments within the same farm.

The population structure described in Table 4-19 was further visualised using dendrograms. Phylogenetic trees of the core and accessory genomes for serogroup O26 (Figure 4-1) and other 'Top 7' STEC serogroups (Figure 4-2) were annotated with antibiotic resistance gene classes and virulence factor genes identified in isolates. Hierarchical cluster analysis was performed for the serogroup O26 isolates with farm and isolation (Figure 4-3). Figure 4-1 and Figure 4-3 show clear differentiation of serogroup O26 *E. coli* based on farm but not isolation source, consistent with the PERMANOVA analysis summarised in Table 4-19.
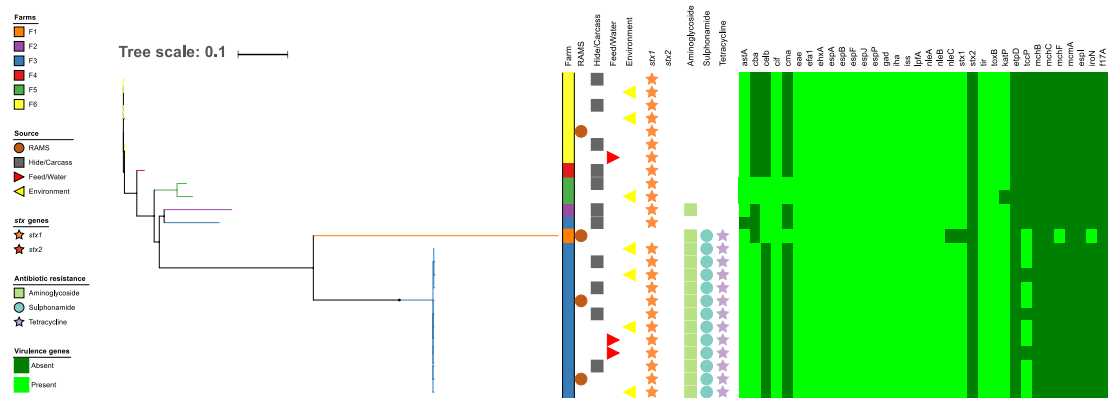
a: core genome
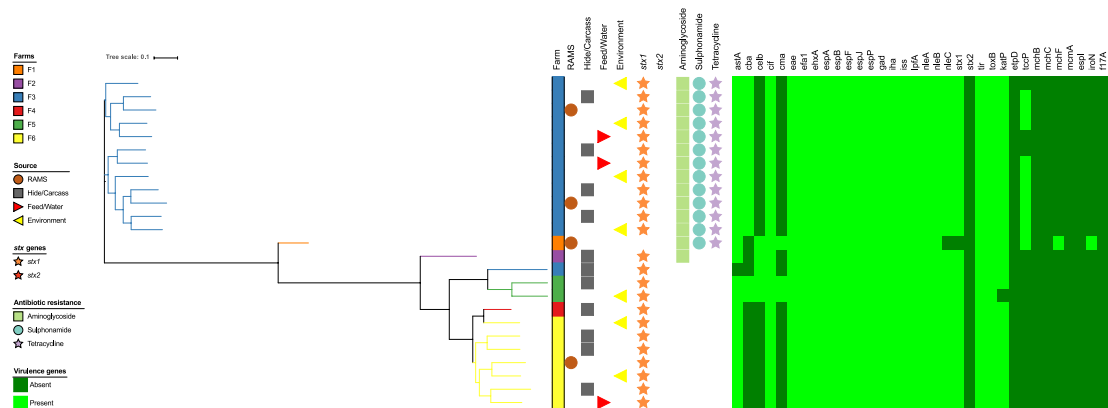
b: accessory genome

Figure 4-1: RAXml phylogenetic tree of *E. coli* serogroup O26 core (*a*) and accessory (*b*) genomes annotated by farm, source, antibiotic resistance gene class, and virulence genes
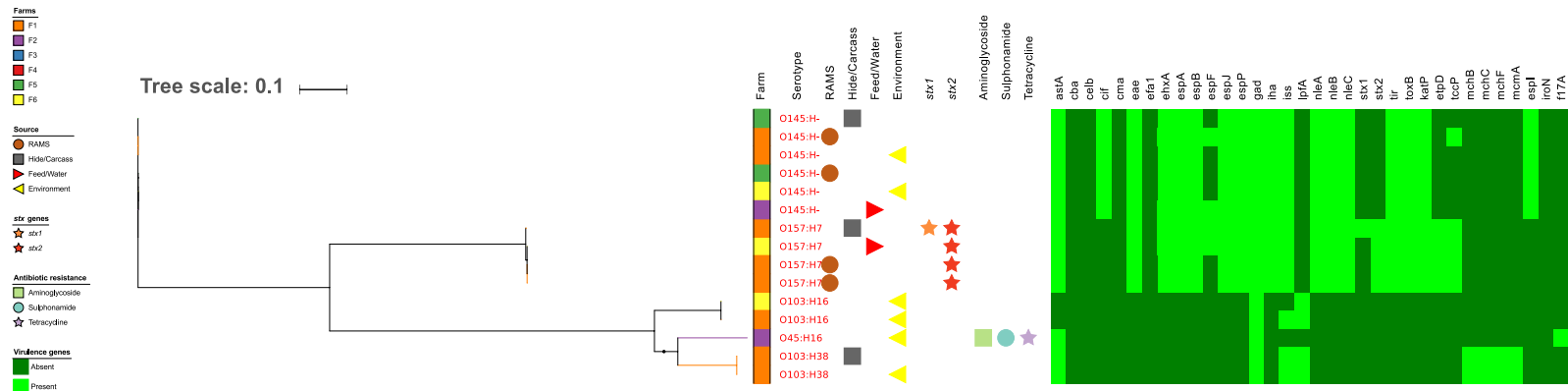
a: core genome
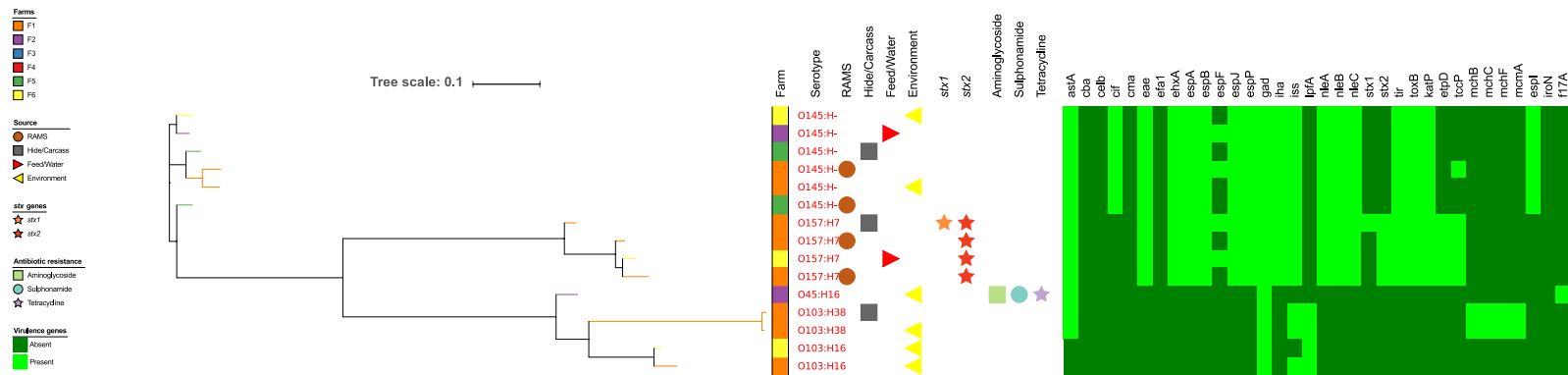
b: accessory genome

Figure 4-2: RAXml phylogenetic tree of non-O26 serogroup *E. coli* core (*a*) and accessory (*b*) genomes annotated by farm, source, antibiotic resistance gene class, and virulence genes
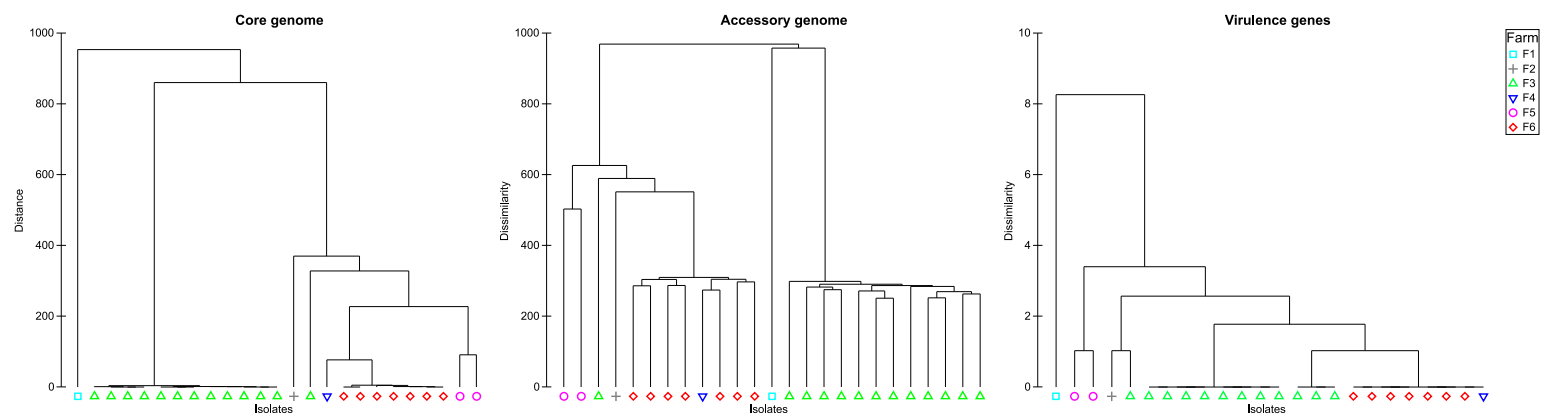
a: farm (n=6)
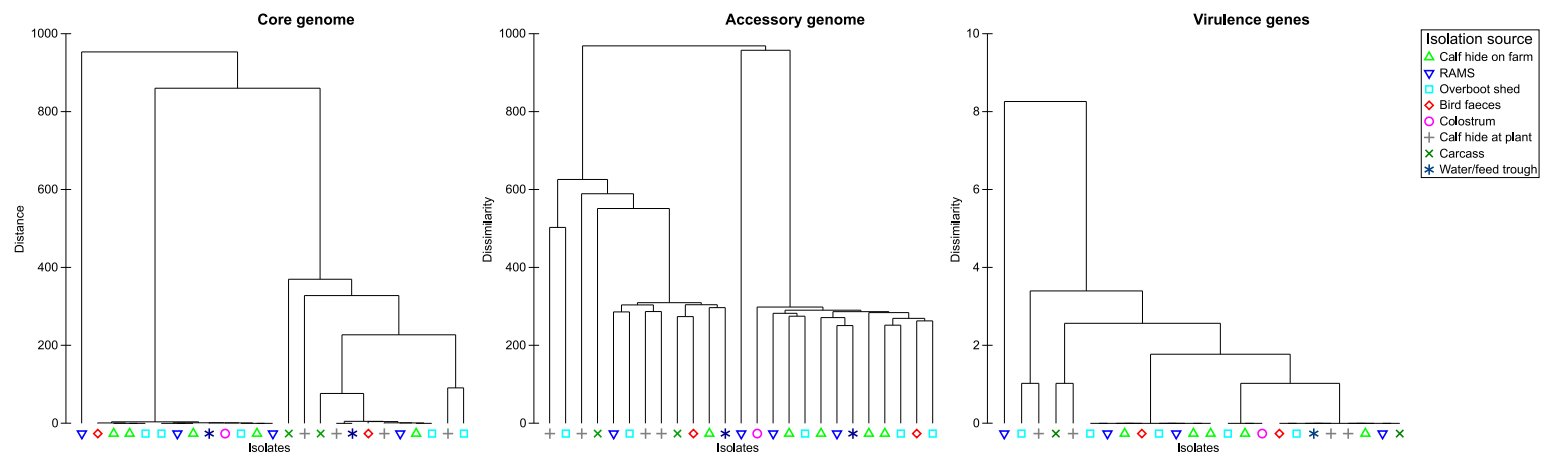
b: isolation source (n=8)

Figure 4-3: Hierarchical cluster analysis of serogroup O26 by farm (*a*) and source (*b*) for core genome, accessory genome, and virulence genes

4.6 Discussion

The aim of this study was to identify factors influencing the carriage and contamination of calves by STEC on dairy farms, and how these relate to post-slaughter veal carcass contamination, in order to identify potential interventions that would decrease carcass contamination at processing plants.

4.6.1 Prevalence and transmission routes of STEC in environmental and animal samples on farm

'Top 7' STEC and potential STEC were detected in every animal, feed, and environmental source in this study (n=17). The actual rectal carriage (RAMS) of 'Top 7' STEC by calves and cows was relatively low (2.7-4.7%) when compared to several environmental and feed sources on farm, such as the calf pen floor (26.0%).

Under experimental conditions, a study from Kulow et al. evaluated 4 to 5 month old steers that were orally inoculated with STEC O157 in a controlled pen environment; the inoculated strain was detected on hides, pen floors, water troughs, and pen walls as well as in RAMS of inoculated steers (58). Other studies have evaluated STEC in the farm environment, with STEC isolated from 17.1% of samples in an evaluation of animals, water, and wildlife samples on 19 farms in the USA (154), to isolation of STEC O26 and STEC O157 from only 1-2% of milk filters in Italy.

In this study 17% of bird faecal samples screened were positive for 'Top 7' STEC. European starlings (*Sternus vulgaris*) can shed STEC O157 with transmission to calves within three days during experimental conditions (168). Bird droppings sampled from two farms 32.5 km apart found the same restriction endonuclease digestion (REDP) subtype of O157:H7, indicating birds as a possible vector (153). Since faecal samples were not directly sampled from birds in the calf pen environment in our study, the

results should be interpreted with caution, as dust present in calf pens has been shown in another study to be STEC positive (59).

Evaluation of factors associated with 'Top 7' STEC and potential STEC calf carriage (RAMS) and calf hide contamination on farm indicated calf to calf, environment to calf, and cow to calf were viable routes of transmission and contamination.

Positive 'Top 7' STEC carriage by calves was associated with the proportion of contaminated calf hides in a shared pen environment on the same day of sampling (OR= 2.01, per 10% increase in prevalence), and not surprisingly, hide contamination in the calf pen was strongly associated with the proportion of calves colonised (RAMS) in that pen on the same day of sampling (OR=4.85, per 10% increase in RAMS positive prevalence). This finding, while intuitively obvious, indicates that active shedding of 'Top STEC' by calves in calf pens is closely linked to hide contamination, which may eventually lead to veal carcass contamination. Hide contamination of calves was also associated with increased shedding of *E. coli* O157 (p=0.002, n=52 heifers) in a similar study in Australia (225). Another study in Australia also found that detection of a non-O157 STEC serogroup in a faecal sample using the NeoSEEK assay was associated with positive prevalence in a hide sample in culled dairy cattle (201).

Similar to the findings in a cross-sectional study of carriage of 'Top 7' STEC by very young calves (**Chapter 3**), increasing numbers of calves in a pen increases the risk of 'Top 7' STEC hide contamination (OR=1.14, increase of one calf in the calf pen). The density of calves in a pen (number of calves divided by the area of the pen) was not significant, suggesting that the biological amplification of 'Top 7' STEC by each animal is more important than decreased living area. Increased transmission with increasing

numbers of animals is biologically plausible and consistent with frequency rather than density-dependent transmission. A mixed effect of direct transmission (such as calves nuzzling each other's contaminated hides) and environmental contamination of pen floors could lead to increased risk of STEC colonization in the shared pen environment. Controlled studies indicated that calf-to-calf transmission was only associated with 20% of serogroup O26 infection, with environmental exposure a more predominant factor (157). Modelling of STEC transmission dynamics revealed 40% of infections were attributable to calf-to-calf transmission, with 60% from environmental sources (158). Increased faecal contamination of calf pens was associated with increased STEC carriage and hide contamination of calves in a controlled infection experiment (159). Factors affecting the pen, such as the previously identified increased risk of 'Top 7' STEC colonisation of calves with higher humidity in the pen (OR=1.09, 1% increase in relative humidity)(**Chapter 3**), may be involved in this pathway by allowing proliferation and environmental persistence of STEC bacteria.

Cow related factors (colonised (RAMS) cow on same farm visit, positive milk filter the same farm visit, positive colostrum sample on the same farm visit) were associated with increased risk of potential STEC in calf RAMS samples and hides of calves on farm (Table 4-14, Table 4-15). These findings suggest that cows that have recently calved can transmit potential STEC to calves, either due to direct transmission via suckling of dams' udders shortly after birth, or by administration of contaminated colostrum in the calf pens.

While no significant difference was seen in 'Top 7' STEC prevalence between calving periods, analysis of calf RAMS and calf hide samples taken on farm indicate a higher

risk (OR=5.6) during the middle, or peak, calving period compared to the early calving period (Table 4-15). A similar increased risk was noted for potential STEC (Table 4-15, OR=2.1). This could indicate that increased risk of calf hide contamination exists during the peak of the calving season, when a large number of calves are being born and living in shared pen environments.

Whole genome sequencing (WGS) analysis of forty bacterial isolates revealed isolates from multiple sources on the same farm displayed a high degree of genetic similarity, consistent with widespread dissemination of clonal strains across multiple environments within the same farm (Figure 4-1). One clear example of clonal transmission was seen in Farm 3 (F3), where twelve STEC O26 isolates from eight animal and environmental sources (Figure 4-1, Figure 4-3), were isolated in 2015 over three periods of the calving season. This provides evidence of persistence and widespread dissemination of specific strains of STEC in the dairy farm environment. In a study in the USA, similar PFGE profiles were seen year to year in the same feedlot, and there was also little to no evidence of transmission of organisms between pens in each feedlot, indicating highly localised transmission dynamics (144). In New Zealand, PFGE analysis of *E. coli* serogroup O26 bacteria found that isolates from the same farm were more clustered than from other farms (177). Overall, the molecular evidence of STEC transmission dairy farms indicates several direct and indirect transmission routes for STEC colonisation and hide contamination of calves.

4.6.2 Prevalence and transmission routes of STEC on hide and pre-intervention veal carcasses at processing plants

A significant increase was noted between 'Top 7' STEC positive calf hides on farm (10.3%) and at the processing plant (37.5%), indicating significant contamination

occurred during calf transport and lairage. This finding has been noted in another study where STEC O157 hide prevalence increased from 50.3% to 94.4% between the time cattle (n=286) were loaded for transport and stunned at the processing plant (208).

Post-slaughter, 12.7% of pre-intervention calf carcass samples were positive for the 'Top 7' STEC, and 39.0% tested positive for potential STEC. Other research has found that the prevalence of STEC contamination of calf hides and carcasses may vary dependent on plant and interventions adopted. A U.S. Department of Agriculture study of young veal calves at processing plants isolated STEC O157 bacteria from 20% of hides and 7% of pre-intervention carcasses (202). This study also utilized the NeoSEEK assay and identified 90% of hides and 68% of carcasses positive for non-O157 STEC; 39% of those detections were confirmed by culture (202).

'Top 7' STEC on calf hides at the processing plant was associated with colostrum cow STEC colonisation (RAMS) on the same day of sampling (OR=9.4), and increased visual contamination of calf hides at the processing plant (faecal score 2 compared to 1, OR=3.02; Table 4-16). An Australian study found that greater concentrations of STEC O157 on hides were correlated to greater concentrations of STEC O157 on pre-evisceration carcasses (206). Similar to previous findings on calf contamination and carriage of STEC on farms, the presence of an actively shedding cow may simply be indicative of widespread dissemination of STEC on the dairy farm.

'Top' 7 STEC contamination of pre-intervention carcass samples was associated with an increasing number of farms visited by the transport truck (OR=1.09, increase of one farm; Table 4-16). These findings indicate the importance of hide contamination on

farm and cross-contamination of calves during transport to eventual contamination of pre-intervention carcasses. For every farm visited by the bobby transport truck, the chances of an individual calf being contaminated increase; a single calf may have the potential to cross-contaminate many other calves during transport and lairage (184). A positive association was found between cattle carcasses that were positive for *E. coli* O157:H7 and transportation in a truckload which contained at least one high shedding (greater than 10^4 CFU/g of faeces) cow (230). Analysis in our study indicated that neither the number of calves in the truck, nor the duration of transport of calves, were found to be significantly associated with contamination of pre-intervention carcass samples. This suggests that an individual calf from a single farm can have a profound impact on hide and pre-intervention carcass contamination at processing plants.

Two sequenced bacterial isolates provided examples highlighting the risk of cross-contamination of hides of calves during transport and lairage. One STEC O26 from a calf hide at the processing plant from Farm 3 (F3), as well as an STEC O157 from a pre-intervention carcass from Farm 1 (F1), both showed clear genetic differences from the predominant strains on their respective farms (Figure 4-1, Figure 4-3). Although these strains could have been present at low levels on the farm of origin, this supports previous New Zealand research that STEC from another calf on another farm directly or indirectly contaminated the hide and carcass of these calves (184). Other research found that only 29% of STEC O157 isolates collected from cattle at a processing plant matched isolates collected on farm using PFGE, indicating cross-contamination of carcasses from other sources (208). This finding provides further evidence to the risk of transport and lairage to increased hide contamination, as well as the risk of

contamination with increasing numbers of farms visited by bobby calf transport trucks.

4.6.3 Limitations

Due to the depth and breadth of sample collection and processing (up to 96 samples a day), this study was limited by the number of farms evaluated (n=6). However, the focused assessment on several farms allowed for an in-depth evaluation of transmission and contamination pathways from farms to processing plants.

Selection bias existed in the 'Top 7' STEC results from NeoSEEK, as most samples that had previously been screened as potential STEC using RT-PCR were eligible for NeoSEEK testing. This strategy was due to resource limitations for the NeoSEEK testing of all samples (n=2580). However, our method used random selection of pre-screened samples, as well as the inclusion of a cohort of bobby calves, to help identify potential transmission routes and risk factors for STEC colonisation and contamination.

Our advanced molecular detection methods (NeoSEEK, RT-PCR) only allowed for prevalence estimates of STEC, rather than the concentration of bacteria in the samples. Understanding concentration of bacteria may be useful for chemical interventions to decrease STEC contamination. While this may be a limitation, these types of molecular methods are used to screen and confirm the presence of 'Top 7' STEC in beef trim for the export market; these methods are directly relevant to the New Zealand meat industry.

Finally, our analysis of the effect of transport and lairage of bobby calves to processing plants was limited by not sampling bobby calves from other farms that were

transported with our cohort calves. However, our investigation identified increased cross-contamination during transport and lairage, as well as the increasing risk of cross-contamination with each farm visited; this would be an important factor to fully evaluate in further studies.

4.7 Conclusion

The key driver for colonisation of very young calves appears to be a combination of dam-to-calf, calf-to-calf, and environment-to-calf factors. Several dam-related pathways, including cow colonisation (RAMS) and contamination of colostrum and milk filters, strongly indicate cows are part of the transmission cycle. The contamination of calf hides, while indicative of shedding of 'Top 7' STEC within the pen, may also act as a transmission route, due to calves nuzzling behaviour with other calves. Our genomic analyses support the conclusion that cows, calves, the environment, and feed sources are contaminated or colonised by the same strains of STEC, indicating that multiple transmission pathways are in action.

Transport and lairage led to significant increases in the prevalence and genomic diversity of 'Top 7' STEC on calf hides at the plant, suggesting cross-contamination of hides had occurred. Visually detectable contamination of hides, as well as contamination of calf hides on farms increased the risk of eventual pre-intervention carcass contamination. The increase in farms visited by the transport truck increased the risk for pre-intervention carcass contamination. This would suggest that calf hide contamination or calf colonisation with STEC from one farm could lead to increased contamination of calf hides and pre-intervention carcasses of calves from other farms.

Due to the large number of potential transmission routes identified in this study, preventing exposure of very young calves to STEC on dairy farms is likely to be difficult to achieve. Even within the first three days of life, 10% of calves had 'Top 7' STEC hide contamination and some were already colonised and shedding STEC.

Higher numbers of calves in shared pens has been identified both in this study and a larger cross-sectional study (**Chapter 3**) as a risk factor for STEC carriage, suggesting that transmission between calves could be reduced by decreasing pen-occupancy rates. Fewer calves in pens would result in decreased amplification of STEC in the environment. Calf pen faecal contamination could lead to hide contamination on the ventral surface of the calf; the ventral surface is where opening cuts are made in the hide during processing, and faecal contamination in this area may pose a higher risk for contamination of veal carcasses. Random clinical trials should be designed to evaluate the influence of decreasing pen-occupancy of calves on the contamination and colonisation of veal calves on farms.

Reduced contamination of calf hides may lead to decreased transmission of STEC on farms, as well as a decreased opportunity for initial contamination of carcasses during slaughter and dressing. Decreasing persistence of STEC in the calf pen environment, as well as on transport trucks and in lairage, may further decrease the level of contamination. Sanitizers and local disinfection could be applied, but enlisting farmers to participate may be impractical; several opportunities for chemical interventions exist during transport and lairage: loading into a transport truck, unloading from a transport truck, and while in lairage.

As opposed to interventions to try and minimise the level of contamination associated with veal calves presented for slaughter and dressing, the value of meat hygiene training at cattle processing plants cannot be overlooked. An assessment of current processing plant practices in the Netherlands revealed half the plants had structural and procedural inadequacies that led to a higher level of microbial contamination of carcasses (209). This study led to significant improvements in hygiene practices. A USDA study in veal processing plants found that after three processors improved their harvest process as recommended by the FSIS, there were significant drops of almost 50% in the contamination levels of hides compared to carcasses (202). In 2016, the Meat Industry Association in New Zealand worked together with the Ministry of Primary Industries to introduce nine initiatives aimed at reducing the risk of veal carcass contamination with 'Top 7' STEC, including hosting workshops targeted at senior operators, supervisors, technical staff and on-site verification staff (14). Continued educational efforts at meat processing plants in New Zealand may reduce the risk of veal carcass contamination.

These results suggest that 'Top 7' STEC is likely to be maintained to some degree in the farm environment throughout the year, and there are a number of risk factors that have the potential to increase the level of colonisation of young calves on dairy farms, as well as hide and pre-intervention carcass contamination during transport and lairage. This research has provided evidence that is useful for risk mitigation of 'Top 7' STEC on farms and at processing plants, and can be used for future studies and application of effective control strategies.

4.8 Acknowledgements

MASSEY UNIVERSITY
GRADUATE RESEARCH SCHOOL

STATEMENT OF CONTRIBUTION
TO DOCTORAL THESIS CONTAINING PUBLICATIONS

(To appear at the end of each thesis chapter/section/appendix submitted as an article/paper or collected as an appendix at the end of the thesis)

We, the candidate and the candidate's Principal Supervisor, certify that all co authors have consented to their work being included in the thesis and they have accepted the candidate's contribution as indicated below in the *Statement of Originality*.

Name of Candidate: Andrew Springer Browne

Name/Title of Principal Supervisor: Distinguished Professor Nigel French

Name of Published Research Output and full reference:
Browne, AS, Biggs, PJ, Cookson, AL, Wilkinson, D, Bloomfield, S, Midwinter, AC, Marshall, JC, Benschop, J, Rogers, L, Hranac, R, Withers, H, Hathaway, S, George, T, Jaros, P, Irshad, H, Fong, Y, Dufour, M, Kariki, N, Winkleman, T. and French, NP. A global genomic examination of Shiga toxin-producing Escherichia coli (STEC) serogroup O26 and non-toxigenic variants from multiple source. In preparation for Emerging Infectious Diseases.
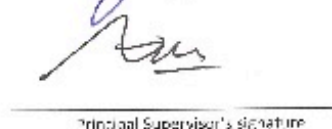
In which Chapter is the Published Work: 5
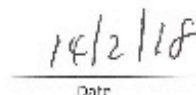
Please indicate either:

- The percentage of the Published Work that was contributed by the candidate:

and / or

- Describe the contribution that the candidate has made to the Published Work:
AS Browne designed the study with input from all supervisors.
AS Browne performed all the laboratory work with assistance from L Rogers.
Other coauthors are cited for assistance with genomic analysis (Wilkinson, Bloomfield), data analysis (Hranac), laboratory methods (Fong), and provision of bacterial isolates for the study (George, Jaros, Irshad, Dufour, Kariki). AS Browne drafted the manuscript and created all figures, with input from supervisors and coauthors.

Candidate's Signature

13/2/18
Date

Principal Supervisor's signature

14/2/18
Date

Tanto monta cortar como desatar

-Spanish proverb

5 A global genomic examination of Shiga-toxin producing *Escherichia coli* (STEC) serogroup O26 and non-toxigenic variants from multiple sources

5.1 Abstract

Shiga toxin-producing *Escherichia coli* (STEC) serogroup O26 are a public health risk, capable of causing haemorrhagic diarrhoea and life-threatening kidney failure. For this study, human and bovine *E. coli* O26 isolates from New Zealand (n=152) were whole genome sequenced and compared with 252 publicly available *E. coli* O26 genomes from 14 other countries. Evolutionary analysis indicated *E. coli* serogroup O26 have an open genome (a Heap's Law coefficient of 0.35), with 2,718 core genes, 8,904 accessory genes, and 9,777 singleton genes detected for a pangenome size of 21,399 genes. Phylogenetic analyses revealed multiple strains have circulated globally and are present in many countries. PERMANOVA analysis of the pangenome composition indicated that genomic variability among isolates was mostly explained by multilocus sequence type (32.5%), country of origin (18%), and *stx* profile (6%). Isolation source (<0.01%) was not a significant predictor of variability. Eight classes of antibiotic resistance genes were detected among all isolates analysed, with aminoglycoside (30%) and sulphonamide (23%) the most commonly detected. Time of most recent common ancestor (TMRCA) estimates revealed key periods of introduction of serogroup O26 sequence type 21 (ST-21) into New Zealand between the

144

1920's and 1980's, along with introduction of non-virulent O26 ST-29 strains during the early 2000's. New Zealand has monophyletic clades that indicate several points of introduction in the past, followed by independent evolution. When compared to another island nation, Japan, New Zealand showed remarkably few instances of introduction and comparisons of historical introductions of live cattle into these countries indicate cattle importation may heavily influence the rate of incursion of novel strains into countries. The emergence of highly pathogenic *stx2* strains in Europe and Japan, which were not identified in New Zealand, may be due to a combination of international trade in live cattle as well as local evolution. Our assessment shows that whole genome sequencing provides valuable insight into the pathways of zoonotic disease transmission between countries.

5.2   Introduction

Shiga toxin-producing *Escherichia coli* (STEC) are a public health risk, capable of causing haemorrhagic diarrhoea and life-threatening kidney failure, particularly in children (5). STEC are primarily transmitted via the faecal-oral route, and ruminants, predominantly cattle, are important reservoirs of this zoonotic pathogen (6).

While much initial research effort focused on STEC serotype O157:H7 as the main STEC pathogen involved in haemolytic uremic syndrome (HUS) (16), STEC serogroup O26, particularly serotype O26:H11, has become an increasingly common cause of human disease. STEC O26 is the second most frequently detected serogroup causing STEC illness in New Zealand (11), the USA (231), and Europe (10). Recent STEC O26 outbreaks were associated with contaminated flour (106) and food consumed at a restaurant chain in the USA (2), and dairy products from Italy and Romania (232). The

emergence of highly pathogenic strains that harbour the *stx2* toxin gene, particularly in Europe, has led to an increase in HUS related to the O26 serogroup (46, 47).

Whole genome sequencing (WGS) has become an essential part of food safety public health investigations by offering high resolution identification of related bacterial isolates, helping to direct source attribution investigations and interventions (233). Cooperative efforts such as GenomeTrakr (234), a laboratory network established by the U.S. Food & Drug Administration (FDA) that includes the Americas, Europe, and Australia, are strengthening the global WGS network for human pathogens. The large amount of sequence data produced by such initiatives provides an opportunity to interpret the evolution and transmission of various organisms across national boundaries.

New Zealand is a geographically isolated island nation that offers a unique opportunity to evaluate the effects of importation and biosecurity measures on control and transmission of zoonotic diseases (7). New Zealand has a relatively high incidence of notified human disease caused by STEC compared to other countries where the disease is notifiable, with 8.9 STEC cases per 100,000 population reported in 2016 (210). A case-control study in New Zealand identified contact with animal manure and the presence of cattle in the local area as significant risk factors for human infection (12). Following this case-control study, a 2014 national cross-sectional study of young dairy calves found STEC O26 in 7.2% of sampled animals using a culture-independent PCR MALDI-TOF assay (**Chapter 3**). The purpose of this study was to undertake comparative genomics of *E. coli* serogroup O26 isolates from human clinical cases, cattle in New Zealand, and genomes of bacterial isolates from non-New Zealand

sources, examining the genetic diversity and population structure, evolution, time of most recent common ancestor (TMRCA), virulence genes, and antibiotic resistance genes. These data can be used to infer the likely importation, transmission, and evolution of STEC within New Zealand, which can inform risk management decisions with the regard to movement of reservoir animals, as well as potential interventions for public health.

5.3   Methods

This research received Massey University Ethics Notification Number 4000016530. This project has been evaluated by peer review and judged to be low risk. Phylogenetic figures were created using the iTOL (Interactive Tree of Life) software (221), and further amended using Inkscape open source software version 0.92.2 (https://inkscape.org).

4.8.5   New Zealand bacterial isolates: selection, DNA and library preparation, and sequencing

Random stratified selection was performed by year, region, farm, and source, to select 152 serogroup O26 bacterial isolates from New Zealand human (n=32) and bovine sources (n=120), from 1985 to 2016. Human isolates were obtained from the New Zealand Reference Culture Collection (The Institute of Environmental Science and Research, Wallaceville, New Zealand), and bovine isolates from the Hopkirk Research Institute at Massey University. DNA extraction was performed on a single colony picked from Columbia Horse Blood Agar (Fort Richard Laboratories, Auckland, New Zealand) using the QIAamp® DNA MiniKit (Qiagen, Hilden, Germany), and sequencing libraries were prepared using the Nextera XT DNA Library Preparation Kit (Illumina, San Diego, United States). Prepared libraries were submitted to New Zealand Genomics Limited (University of Otago, Dunedin, New Zealand), who

performed sequencing using Illumina MiSeq 2 x 250bp PE and Illumina HiSeq 2 x 125 bp PE v4.

Processed reads are publicly available on the NCBI Sequence Read Archive (SRA) under BioProject ID PRJNA396667; metadata are stored under the BioSample accession numbers SAMN07430747 to SAMN07430900 inclusive (Appendix Table 5).

4.8.6 Selection and retrieval of publicly available *E. coli* serogroup O26 raw sequence data

Due to quality threshold considerations for downstream analyses, only bacterial isolates with raw sequencing data were selected for inclusion in this study. Publicly available *E. coli* serogroup O26 sequences were obtained by examining literature in the PubMed database for "O26" and "genome", and searching public databases, including NCBI SRA metadata and the EMBL-EBI European Nucleotide Archive. Isolate selection was stratified by country, year, and isolation source; up to four isolates were selected randomly from the same country, year, and isolation source. For the purposes of this study, raw beef samples taken from trim or carcasses were classified as "bovine", while the "food" classification indicated non-meat samples (e.g. spinach, flour). All potential sequences for this study were selected prior to January 2018. A complete list of the 252 publicly available serogroup O26 sequences selected for this study is available in Appendix Table 6, with BioProject, BioSample, SRA run number, and study strain name listed therein.

4.8.7 Assembly, annotation, and initial analyses of whole genome sequence data

Whole genome sequence data was evaluated for sequence contamination using Kraken (235), assembled, and annotated using the Nullarbor pipeline in the "accurate" mode (218). Virulence and resistance genes were identified using ABRicate (https://github.com/tseemann/abricate). Virulence and resistance genes were defined

as present in a genome when greater than 60% gene coverage could be identified by this method. ECTyper was used to identify somatic (O) and flagellar (H) antigens (O:H serotype) (https://github.com/phac-nml/ecoli_serotyping). Identified attributes and metadata for all genomes are available in Appendix Table 7.

Pangenome analysis was performed with the FindMyFriends package in the R environment (236). This method groups genes into orthologous clusters by implementing the cd-hit clustering algorithm (237), followed by a cluster-refinement based on k-mer similarity. The HierarchicalSets package (238) is then used to examine pangenome composition, estimating the similarity of isolates based on the number of shared (core), and characteristic (accessory/pan) genes. Gamma heterogeneity, defined as the ratio of the number of core genes (intersect) to the number of pan genes (union) is calculated for each group of genomes, and isolates are hierarchically clustered in order to minimise total heterogeneity, producing a dendrogram representation of genomic similarity. Functional annotation of clusters of orthologous groups (COGs) created by FindMyFriends was performed with eggNOG-Mapper (239). RaxML maximum-likelihood trees were generated from the concatenated alignment of all core genes that were output by FindMyFriends (219). A dissimilarity matrix was created with the virulence gene output, based on the presence and absence of virulence genes between pairs of isolates, and used to create neighbour-joining trees. We performed hierarchical clustering on the virulence gene presence/absence matrix using the "hclust" package in R (222), and compared it to a core genome phylogeny of all 404 genomes from Snippy version 3.0 (https://github.com/tseemann/snippy).

A distance matrix of the HierarchicalSets, as well as a dissimilarity matrix of the virulence genes (n=192), were evaluated with PERMANOVA (PRIMER-E, Quest Research Limited, Auckland, New Zealand) with sequence type, country, isolation source, and *stx* profile as independent factors.

4.8.8 Single nucleotide polymorphism (SNP) alignment and time of most recent common ancestor (TMRCA) analyses

DECIPHER was used to create a core gene alignment from the FindMyFriends package (240). Two core alignments were performed for 2 sequence types: ST-21 (n=345 isolates), and ST-29 (n=48 isolates). Recombinant regions and identical isolates were removed using Gubbins 2.3.1 (241), resulting in 344 isolates (ST-21) and 48 isolates (ST-29) in the final analysis.

The time of most recent common ancestor (TMRCA) was determined using BEAST 2 (242). Model evaluation of a combination of substitution, clock, and population models was performed using a method-of-moments-estimator (243) and evaluation of log files using Tracer v1.6.1 led to a preferred model selection. General Time Reversible (GTR) substitution models were used to estimate TMRCAs with a Coalescent Extended Bayesian Skyline model and relaxed molecular clock (244). TMRCA analysis was calibrated by tip dates (ST-21: 1947-2017; ST-29: 1952-2017); decimal dates were rounded to the middle of the month or year if an exact date was not available within the month or year. Effective sample size (ESS) exceeded 100 for all models evaluated. To determine the substitution rate for each sequence type, we multiplied the substitution rate estimated by BEAST 2 by the number of analysed core SNPs and then divided the product by the mean genome size of the isolates analysed. Maximum clade credibility trees were created using TreeAnnotator v2.4.7 with a 10% burn-in (242).

4.8.9 Cattle importation data

New Zealand cattle importation data were combined from a previous publication of historical importations of cattle into New Zealand (245), and Food and Agriculture Organization (FAO) data from 1961 to 2013 (246). Live cattle imports into Japan from 1961 to 2013 were obtained from FAO data to allow for a comparison to New Zealand live cattle imports (246).

5.4 Results

The 404 *E. coli* serogroup O26 genomes analysed in this study are summarized in Table 5-1. Most isolates were obtained from New Zealand (n=152), Japan (n=94), and the USA (n=79). The majority of isolates were ST-21 (n=345) and ST-29 (n=48), multiple *stx* gene profiles were represented (*stx1*, *stx2*, *stx1* & *stx2*, and no *stx*), and the source of isolates fell into four groups: "human", "bovine", "food", and "other animal".

Table 5-1: Summary of *E. coli* serogroup O26 isolates (n=404) by country (n=15)

Country	*stx* profile				Sequence type			Source			
	stx1	*stx2*	*stx1& stx2*	No stx	ST-21	ST-29	Other-ST*	Human	Bovine	Food	Other animal
Australia (n=1)	1				1			1			
Belgium (n=24)	20	1	2	1	20	1	3	16	8		
Continental Europe* (n=21)	3	13	2	3	6	13	2	19	2		
Japan (n=94)	70	8	11	5	88	5	1	77	16		1
New Zealand (n=152)	104			48	136	16		32	120		
Other North America# (n=4)	3			1	3		1	2	1	1	
United Kingdom (n=29)	10	7	8	4	25	3	1	28	1		
United States of America (n=79)	60	9	5	5	66	10	3	45	27	4	3
Total (n=404)	271	38	28	67	345	48	11	220	174	5	5

*Denmark (n=1), France (n=9), Germany (n=6), Italy (n=1), Norway (n=2), Poland (n=1), Switzerland (n=1)
#Canada (n=3), Mexico (n=1)

4.8.10 Evolutionary dynamics of *E. coli* serogroup O26

Gene clustering analyses identified 2,718 core genes, 8,904 accessory genes, and 9,777 singleton genes, for a pangenome size of 21,399 genes. Core genes are present in all

genomes, accessory genes are present in more than one but not all genomes, and singleton genes are only present in a single genome. A plot of the number of core, accessory, and singleton gene groups with the addition of each additional O26 isolate indicates an open genome (Figure 5-1), meaning the addition of new bacterial isolates leads to the addition of new genes into the pangenome. The optimal fit of the Heap's Law coefficient to these data was 0.35, further supporting a classification as an open genome (Figure 5-2). The number of genes identified in each genome varied from 4,388 to 5,855, with an average of 5,305 genes per genome. When placed in a bar graph with sequence type, country, isolation source, and *stx* profile, bacterial isolates that contained no *stx* gene had lower gene numbers, but no other clear patterns were visible (Figure 5-3).
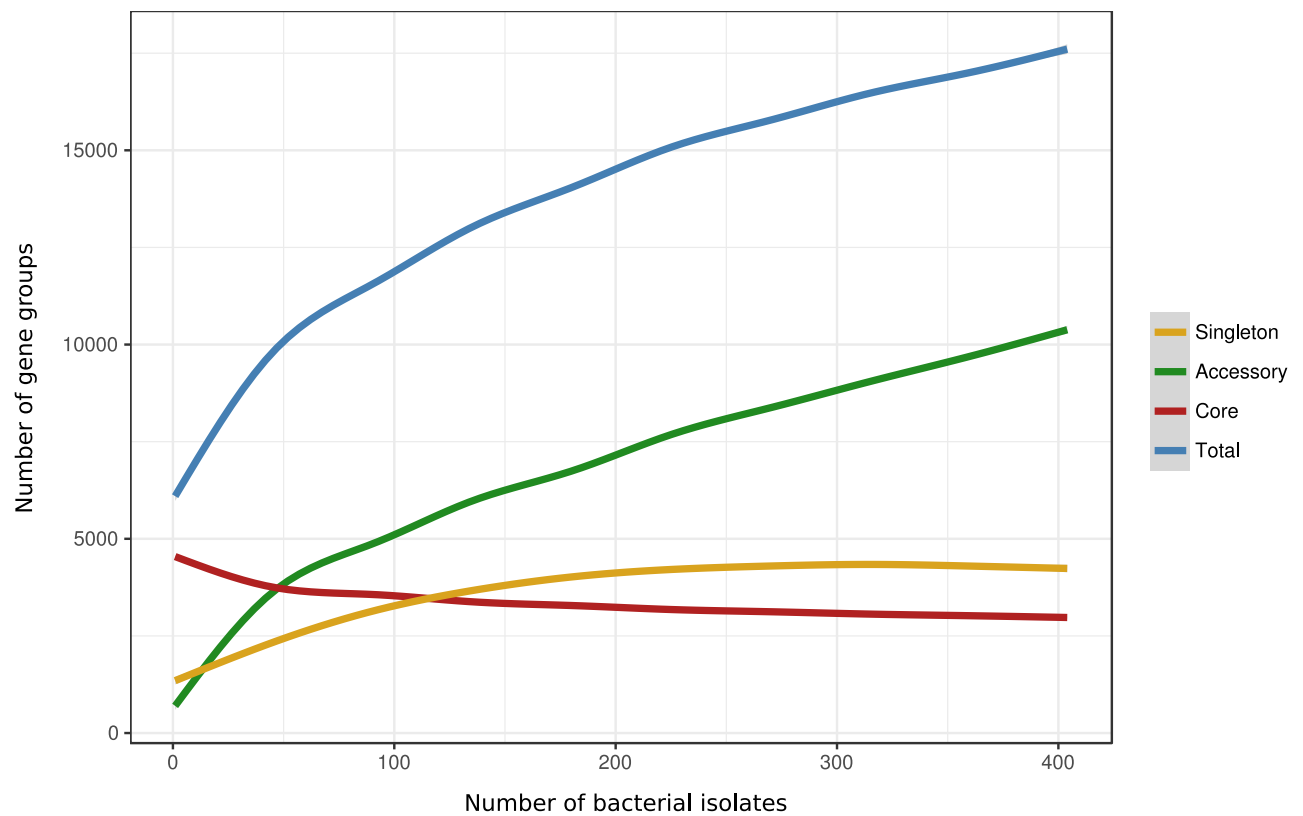
Figure 5-1: Number of gene groups per serogroup O26 bacterial isolate by core, accessory, and singleton gene group counts
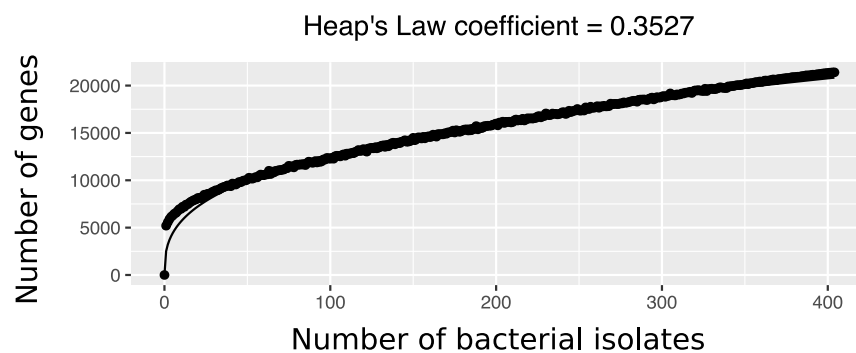
Figure 5-2: Heap's Law coefficient of serogroup O26 bacterial isolates (n=404), by number of new genes per bacterial isolate
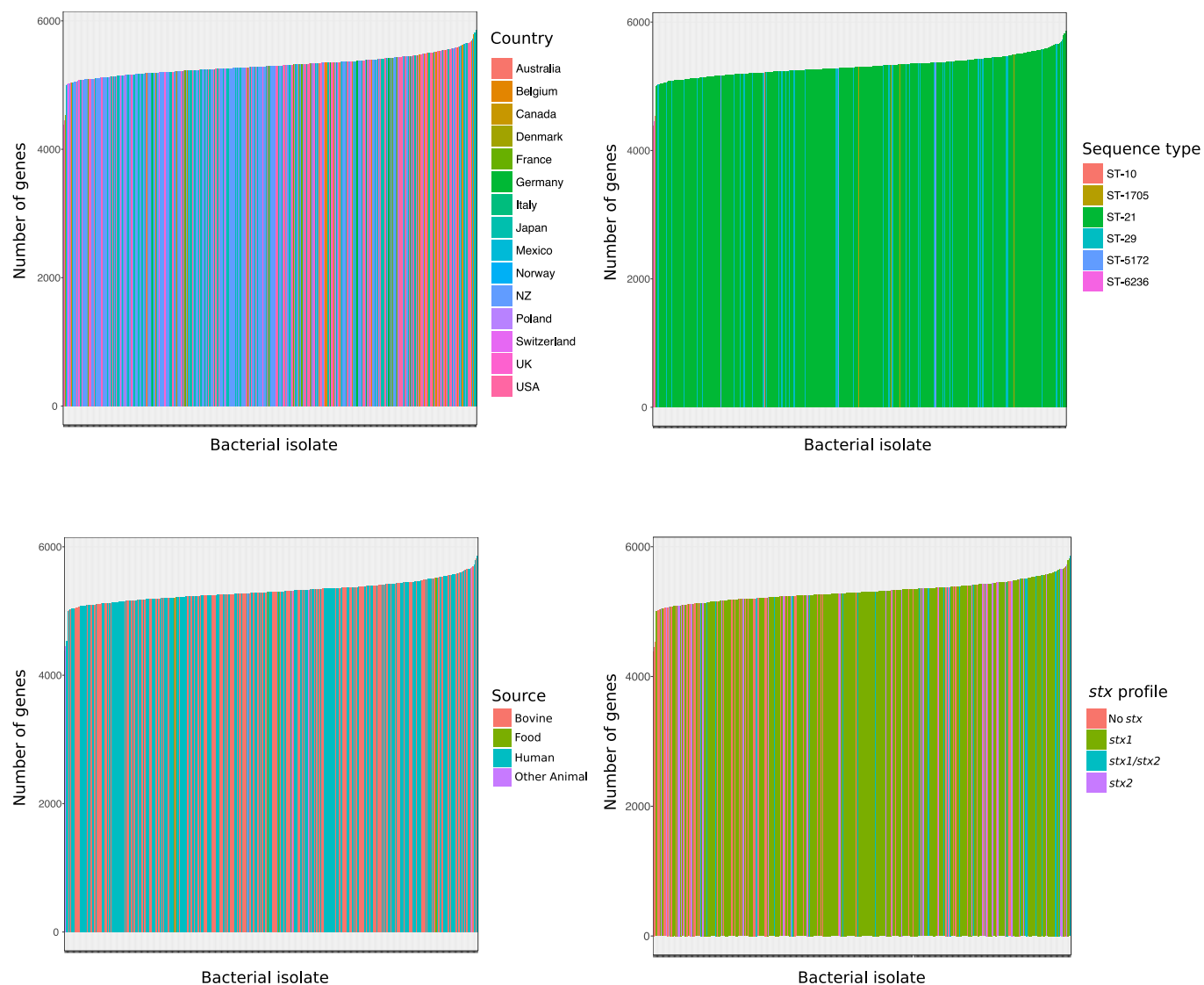
Figure 5-3: Number of genes detected in serogroup O26 bacterial isolates (n=404) by country, sequence type, source, and *stx* profile

Hierarchical set clustering outputs, with shared gene groups present, are shown in Figure 5-4, with clades "A" through "E" marked for comparisons between figures; Appendix Figure 2 includes the real branch lengths of the phylogenetic tree. The green blocks in the "Intersection" portion of Figure 5-4 indicate the number of core genes shared by that group; these indicate that clades "A" through "D" each share approximately 4,000 core genes. The "Union" portion of Figure 5-4 represents the pangenome relatedness between bacterial isolates. A summary of the functional

annotation gene groups is represented in Figure 5-5, for all genomes as well as specific

clades. Proportions of gene functional groups were visually similar between the clades,

indicating that no particular clade had a preference for more genes that perform a

particular function. A pangenome hierarchical set tree is annotated with country,

sequence type, isolation source, and antibiotic resistance gene class (Figure 5-6).

Multiple strains have circulated globally and are present in many countries, but New

Zealand has evidence of monophyletic clades. Clades "A" and "D" had only ST-21

isolates, while clades "B" and "C" had a mixture of ST-21, ST-29, and other multilocus

sequence types. Bovine and human isolation sources were interspersed in the same

clades. Distinct combinations of resistance genes were present in isolates from

different countries.
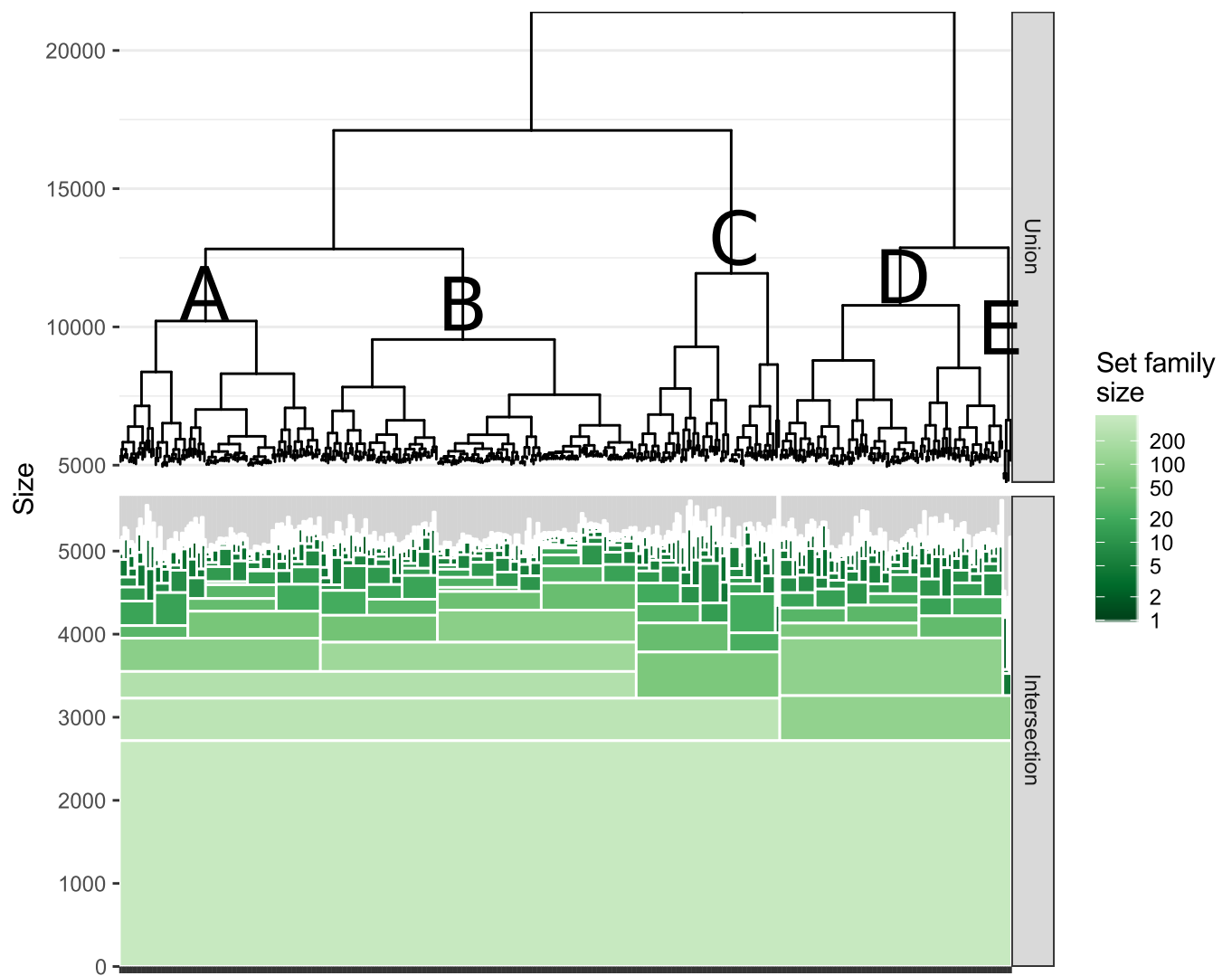
Figure 5-4: Hierarchical set analysis of *E. coli* serogroup O26 isolates (n=404), with a hierarchical set tree and shared gene groups visualized in green
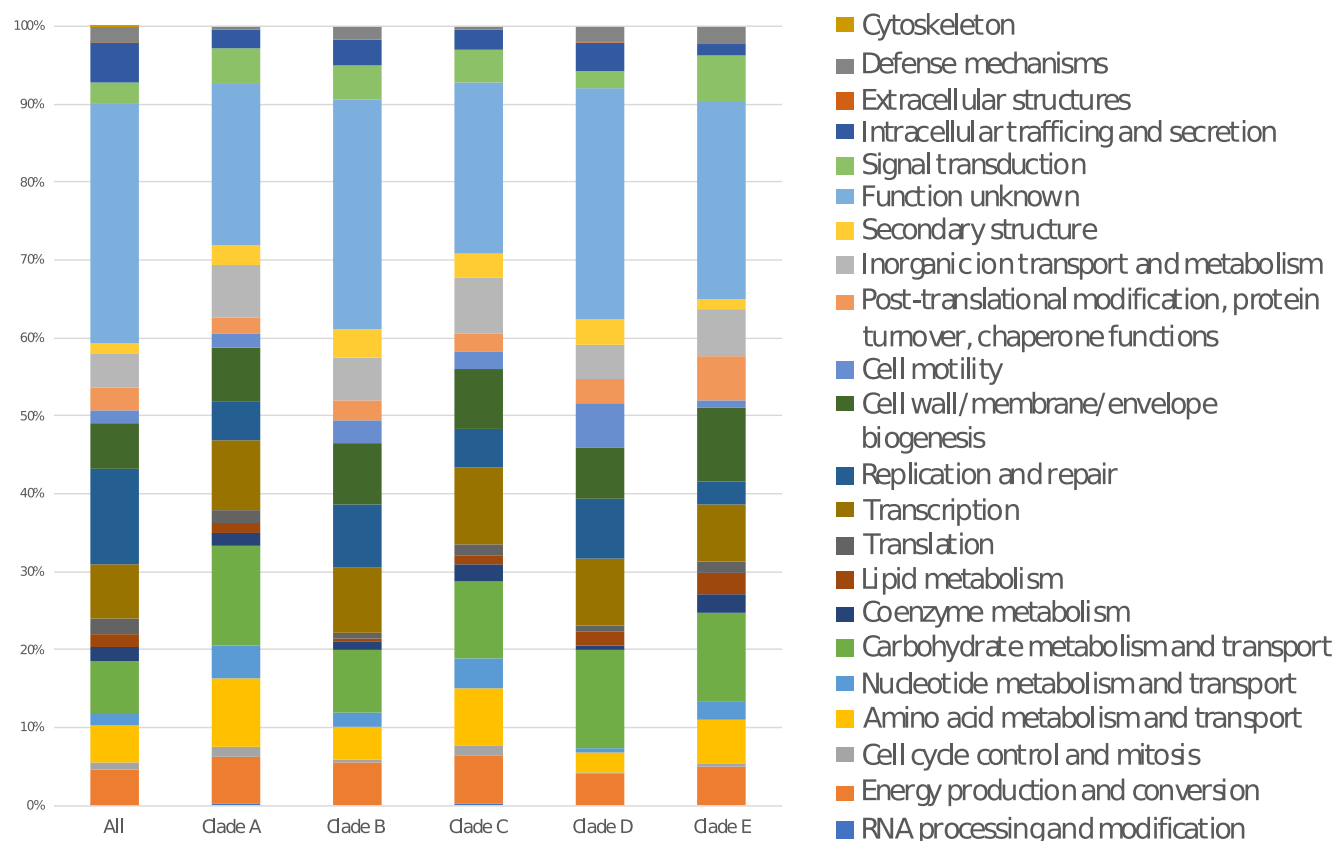
Figure 5-5: Functional annotation classes for gene groups per labelled clade (A through E) and all bacterial isolates (n=404)

Legend:
- Cytoskeleton
- Defense mechanisms
- Extracellular structures
- Intracellular trafficing and secretion
- Signal transduction
- Function unknown
- Secondary structure
- Inorganic ion transport and metabolism
- Post-translational modification, protein turnover, chaperone functions
- Cell motility
- Cell wall/membrane/envelope biogenesis
- Replication and repair
- Transcription
- Translation
- Lipid metabolism
- Coenzyme metabolism
- Carbohydrate metabolism and transport
- Nucleotide metabolism and transport
- Amino acid metabolism and transport
- Cell cycle control and mitosis
- Energy production and conversion
- RNA processing and modification

Country of origin
- New Zealand
- Japan
- United States of America
- Belgium
- United Kingdom
- Other Continental Europe
- Other North America
- Australia

Sequence type
- ST-21
- ST-29
- ST-other

Shiga toxin
- *stx1*
- *stx2*

Source
- Human
- Bovine
- Other Animal
- Food

Antibiotic resistance
- Aminoglycoside
- Beta-lactam
- Macrolide
- Phenicol
- Quinolone
- Sulphonamide
- Tetracycline
- Trimethoprim

Figure 5-6: Hierarchical set tree of pangenome elements of *E. coli* serogroup O26 isolates (n=404), annotated by country, sequence type (ST), *stx* profile, source, and antibiotic resistance gene class. The clades as defined in previous figures above are annotated.

4.8.11   Pathogenicity of serogroup O26 *E. coli*

Figure 5-7 illustrates a neighbour-joining tree based on a distance matrix of the presence and absence of virulence genes detected (n=192); a real branch length figure is also available (Appendix Figure 3). Of particular interest are the large numbers of New Zealand isolates with identical virulence profiles from human and bovine sources,

while a large clade from Japan, the USA, and Belgium also have identical profiles. A hierarchical cluster analysis of virulence genes loosely cluster the genes into 3 clades, separated by the presence of *stx1*, *stx2*, and *eae*, indicating that if these genes were present in a bacterial isolate, other particular virulence genes were also likely to be present (Figure 5-8). The clusters of virulence genes support the presence of potential pathogenicity islands, such as the Locus of Enterocyte Effacement pathogenicity island (LEE), indicated by the genes clustering with *eae* (e.g. *esp*, *sep*). To evaluate the importance of virulence genes to overall evolution and differentiation between isolates, we generated a hierarchical tree of virulence genes compared to a core gene alignment generated with Snippy (Figure 5-9). This comparison indicates several virulence gene clusters are predictive for the overall phylogenetic Snippy tree. For instance, the presence of *gsp* and *fae* genes, involved with the general secretory pathway (*gsp*) and bacterial adhesion to the intestinal villi (*fae*), defined a group of ST-29 isolates from New Zealand, Belgium, Switzerland, Italy, the UK, and France.
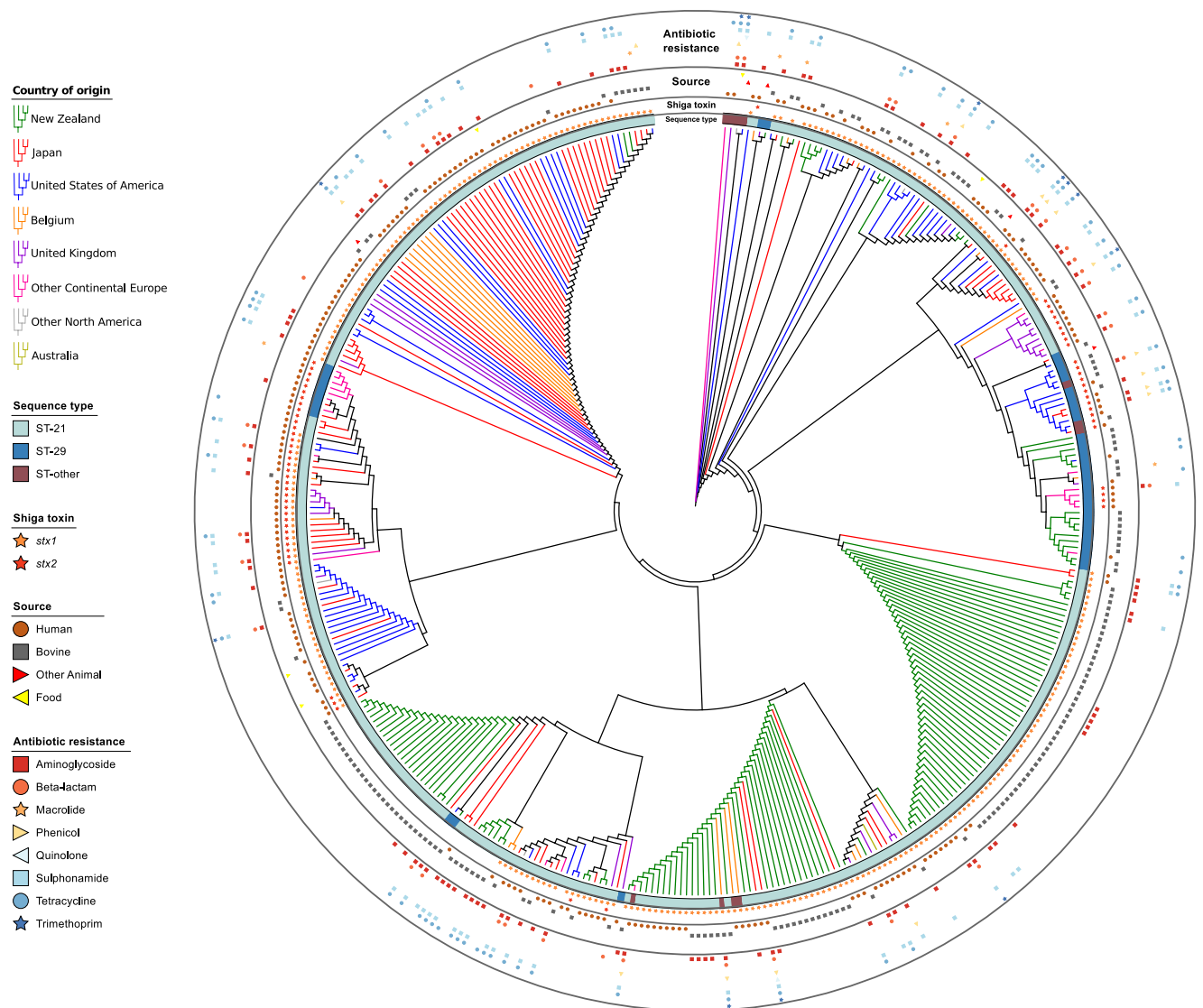
Figure 5-7: Neighbour-joining tree of virulence genes (n=192) of serogroup O26 isolates (n=404) annotated with country, sequence type, *stx* profile, source, and antibiotic resistance gene class
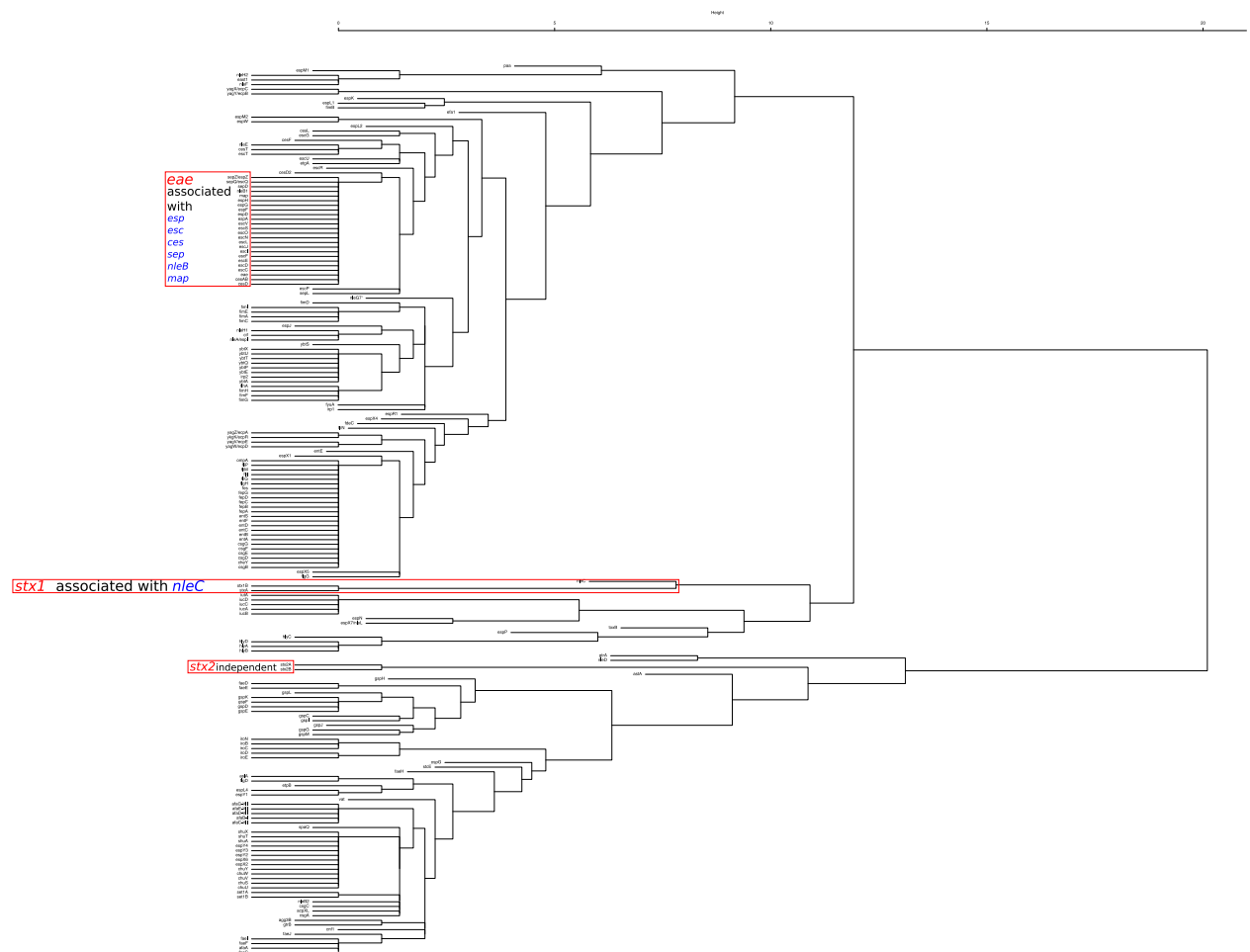
Figure 5-8: Hierarchical cluster tree of dissimilarity matrix of virulence genes detected (n=192) with *stx1, stx2, eae* and associated virulence genes highlighted
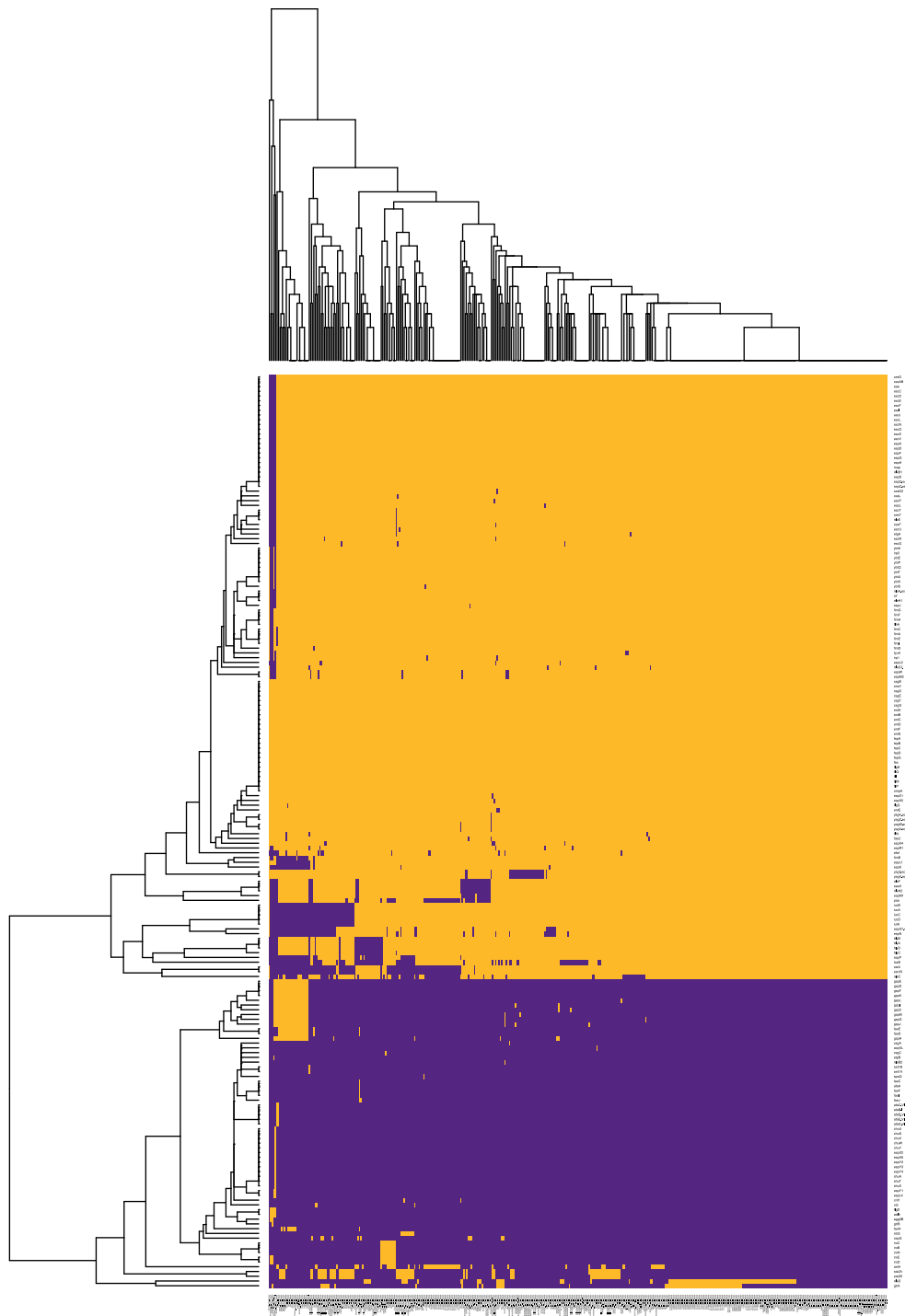
Figure 5-9: Heatmap (gold=present; purple=absent) of hierarchical cluster analysis of virulence genes (x-axis, n=192) compared with a SNP core gene alignment (y-axis, n=404)

4.8.12 Evolution of predictors of genetic variability

PERMANOVA analysis for the pangenome and virulence genes revealed variability among isolates was mostly predicted by sequence type (33%, 84%), country of origin (18%, 2%), and *stx* profile (6%) respectively, while isolation source was not a significant factor (Table 5-2).

Table 5-2: PERMANOVA analysis of pangenome genes and virulence genes by sequence type (ST), country, isolation source, and *stx* profile

Dataset	Variable	df	Pseudo-F	p-value	Component of variation
Pangenome (n=21,399 genes)	Sequence type	5	9.3	0.0001	33%
	Country	14	8.3	0.0001	18%
	Isolation source	3	1.1	0.358	<0.01%
	stx profile	3	6.9	0.0001	6%
Virulence genes (n=192)	Sequence type	5	119.3	0.0001	83.7%
	Country	14	5.4	0.01	1.9%
	Isolation source	3	2.8	0.07	0.3%
	stx profile	3	35.3	0.0001	6.2%

Residual variation: pangenome (43%), virulence genes (7.8%)

4.8.13 Time of most recent common ancestor (TMRCA) analysis and inferred global importation and transmission of *E. coli* O26

A core gene alignment of serogroup O26 ST-21 isolates (n=344) generated 9,702 SNPs, while ST-29 isolates (n=48) generated 4,686 SNPs. Gubbins (241) identified 324 areas of recombination in ST-21 and 277 areas of recombination in ST-29 alignments (Figure 5-10, Figure 5-11). The time of most recent common ancestor (TMRCA) estimates for ST-21 and ST-29 isolates are shown in Figure 5-12 and Figure 5-13 respectively, with important convergence dates annotated with a 95% highest posterior density (HPD) interval; detailed figures with posterior probabilities and each isolate identified are

available in Appendix Figure 5 and Appendix Figure 6 for the ST-21 and ST-29 isolates respectively.

In Figure 5-12, four New Zealand ST-21 monophyletic clades indicate TMRCA estimates between the 1920's to 1990's. Individual New Zealand monophyletic clades show evidence of importation from European sources (1958 to 1982 95% HPD interval) or USA sources (1971 to 1992 95% HPD interval). Paraphyletic clades are visible from European sources, and in particular from USA and Japan isolates, which create a panmectic community, indicating frequent transmission between these countries. The TMRCA for all ST-21 isolates was in the late-18th Century (1717 to 1822 95% HPD interval). In Figure 5-13, two New Zealand ST-29 monophyletic clades show TMRCA estimates from the 1970's to the early 21st century. TMRCA estimates of clades of individual countries indicate various dates from the mid 19th century onwards. The TMRCA for all ST-29 isolates was older than the ST-21 isolates, at 1427 to 1696 (95% HPD interval). In both Figure 5-12 and Figure 5-13, there is no evidence of transmission of New Zealand strains to the other countries evaluated in this study. The calculated substitution rate for ST-21 was 1.4×10^{-7} substitutions per site per year (1.1×10^{-7} - 1.7×10^{-7} 95% CI) and the substitution rate for ST-29 isolates was 3.2×10^{-7} substitutions per site per year (2.3×10^{-7} - 3.9×10^{-7} 95% CI).
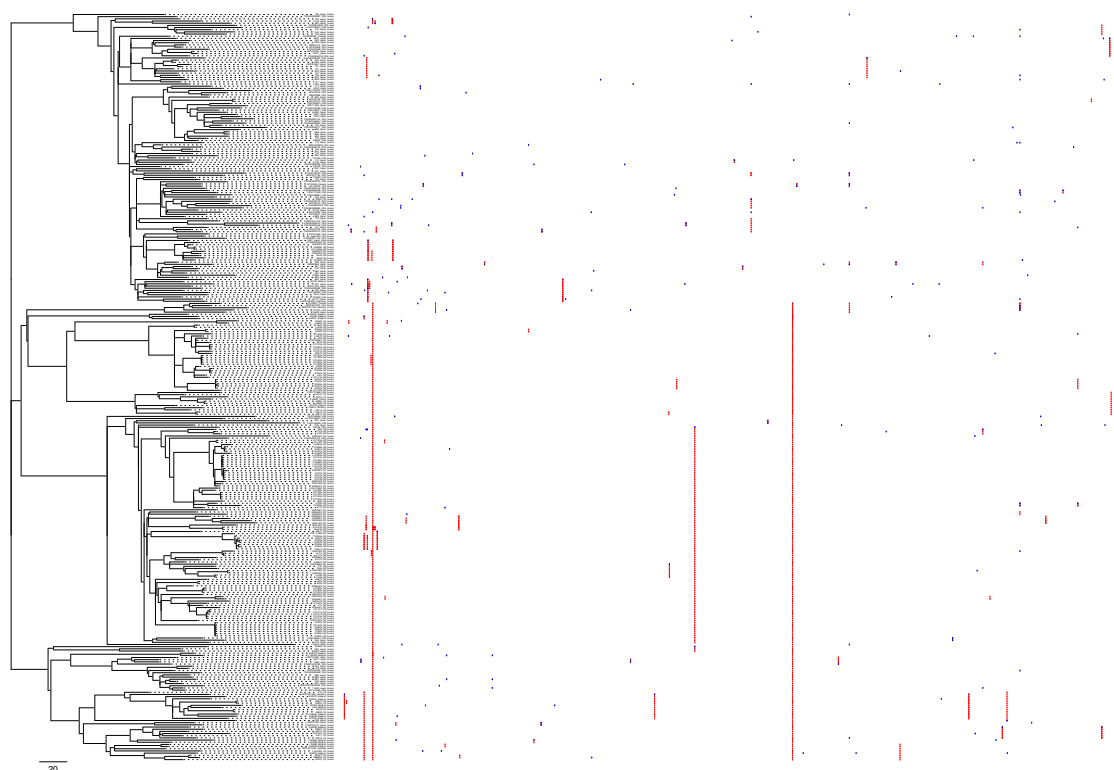
Figure 5-10: Areas of recombination (n=324) removed during Gubbins (241) processing for *E. coli* serogroup O26 sequence type 21 (ST-21) isolates (n=344). For each isolate, blocks representing the regions identified as recombinations are indicated by coloured blocks. Blue blocks are unique to a single isolate while red blocks are shared by multiple isolates. The horizontal position of the blocks represents their position in the core gene alignment of 3310 genes.
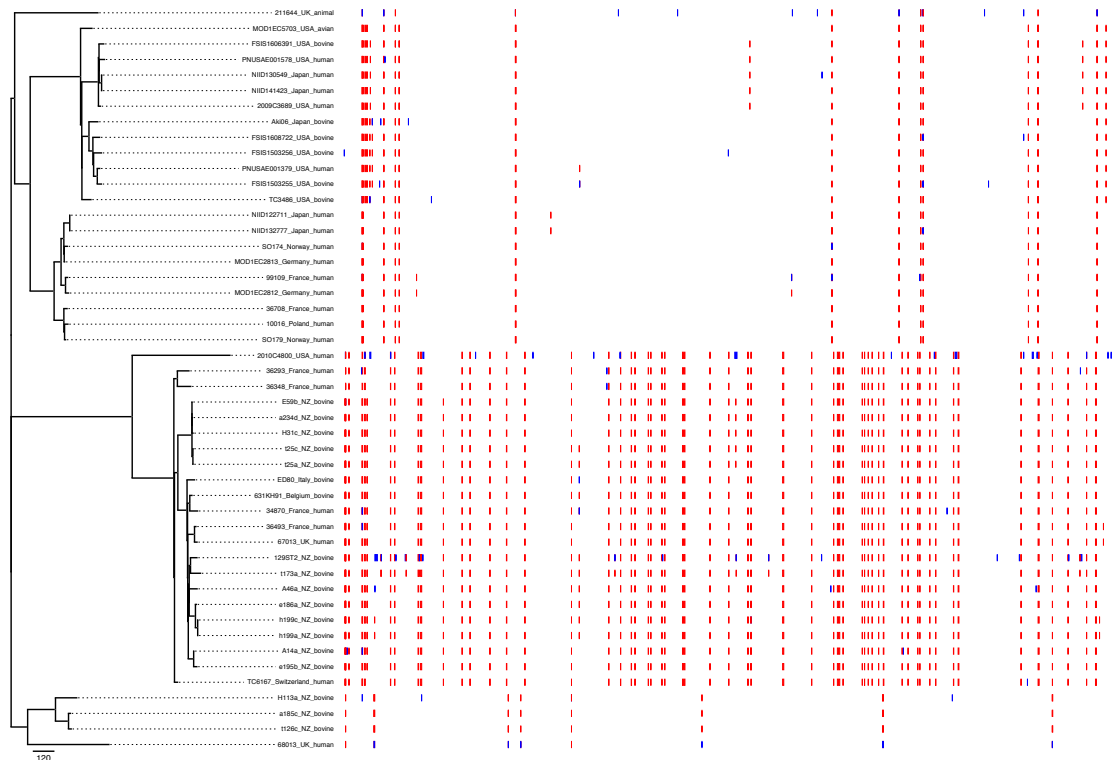
Figure 5-11: Areas of recombination (n=277) removed during Gubbins (241) processing for *E. coli* serogroup O26 sequence type 29 (ST-29) isolates (n=48). For each isolate, blocks representing the regions identified as recombinations are indicated by coloured blocks. Blue blocks are unique to a single isolate while red blocks are shared by multiple isolates. The horizontal position of the blocks represents their position in the core gene alignment of 3,727 genes.
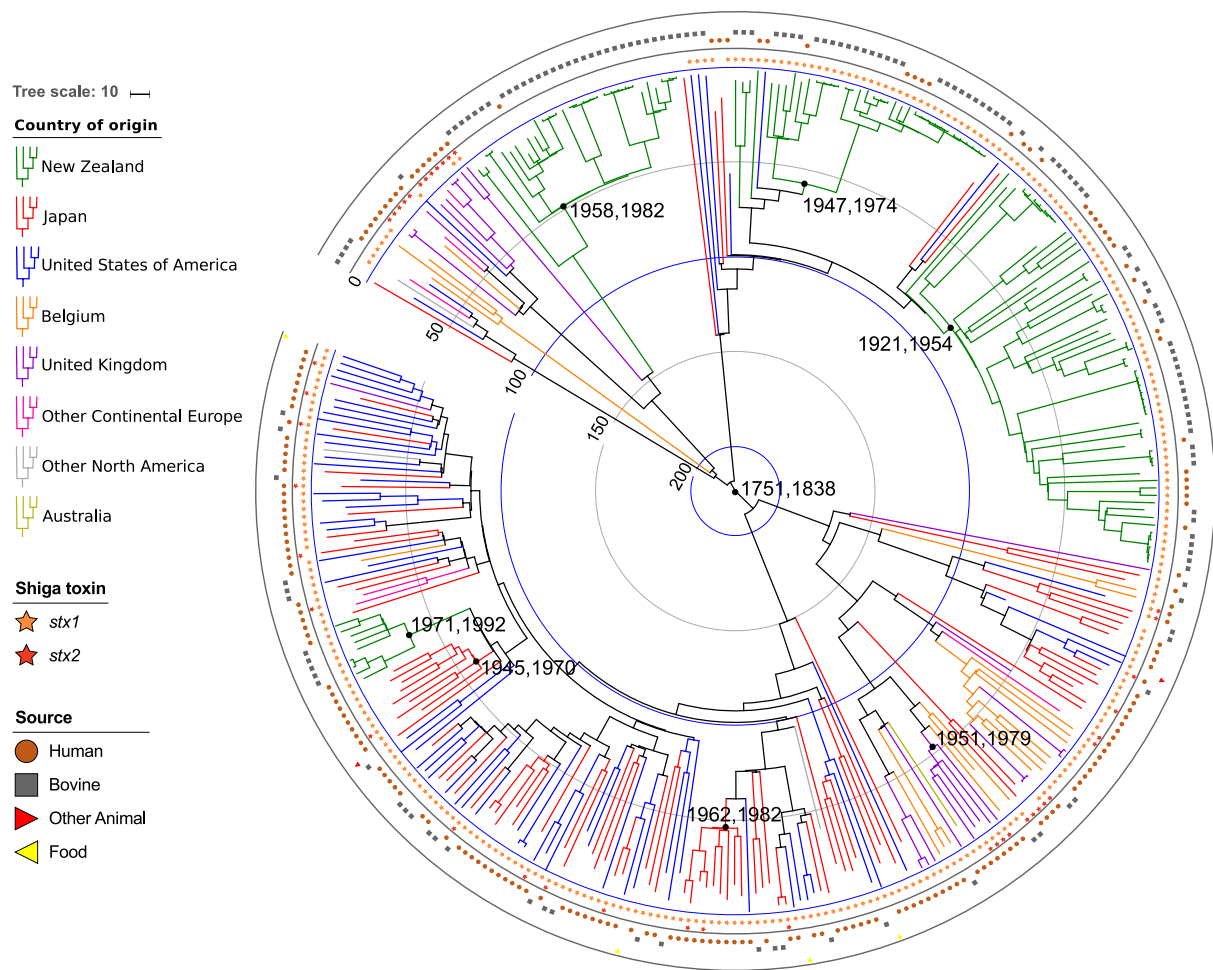
Figure 5-12: Maximum clade credibility tree of time of most recent common ancestor (TMRCA) analysis of *E. coli* serogroup O26 sequence type 21 (ST-21) isolates (n=344), annotated by country, *stx* profile, and source. Key convergence dates are annotated with 95% HPD intervals, and the concentric circles indicate prior time periods (blue, 100 years; grey 50 years) from the age of the newest isolate (2017.5 in decimal years)
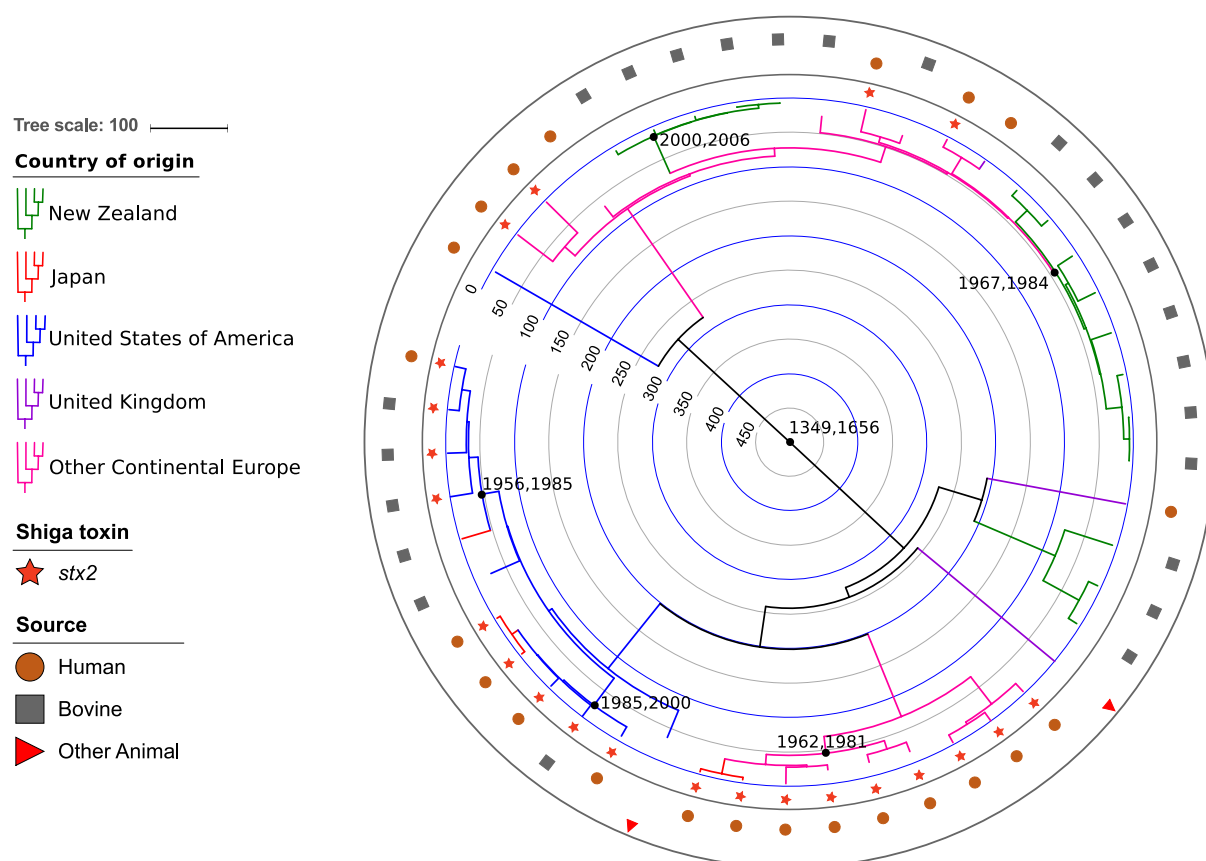
Figure 5-13: Maximum clade credibility tree of time of most recent common ancestor (TMRCA) analysis of *E. coli* serogroup O26 sequence type 29 (ST-29) isolates (n=48), annotated by country and virulence. Key convergence dates are annotated with 95% HPD intervals, and concentric circles indicate prior time periods (blue, 100 years; grey 50 years) from the age of the newest isolate (2017.0411 in decimal years)

Detection of resistance genes for eight classes of antibiotics are shown in Table 5-3, grouped by country and isolation source; a table of detected resistance gene classes for all isolates is available (Appendix Table 8). Resistance genes were detected in 252 bacterial isolates (62.4%). The most commonly detected resistance genes belonged to the aminoglycoside (29.9%), sulphonamide (23.0%), and tetracycline (19.3%) classes of

antibiotics. New Zealand and the USA had low prevalence of antibiotic resistance genes for all classes of antibiotics analysed. Human isolates had a higher prevalence of antibiotic resistance genes than bovine isolates for the beta-lactam, sulphonamide, and tetracycline classes of antibiotics.

Table 5-3: Antibiotic resistance gene detection of serogroup O26 bacterial isolates (n=404) by antibiotic class (n=8) compared with country and source

Factor evaluated		Aminoglycoside	Beta-lactam	Macrolide	Phenicol	Quinolone	Sulphonamide	Tetracycline	Trimethoprim
Country	Australia (n=1)	100%			100%		100%		
	Belgium (n=24)	67%	25%	4%	17%		67%	42%	17%
	Continental Europe (n=21)	29%	14%	10%	5%		24%	14%	
	Japan (n=94)	32%	13%	1%	6%	1%	32%	24%	3%
	New Zealand (n=152)	26%	0.7%	0.7%			12%	12%	0.7%
	Other North America (n=4)	50%	25%	25%		25%	50%	100%	
	United Kingdom (n=29)	24%	21%		4%		29%	25%	4%
	United States of America (n=79)	13%	13%		4%		4%	14%	2.5%
Source*	Human (n=220)	28%	12%	0.4%	4%	0.4%	27%	21%	3%
	Bovine (n=175)	27%	6%	3%	4%		16%	15%	2%
	Food (n=5)	20%	20%			20%	20%	20%	20%
All isolates (n=404)		29.9%	10.1%	2.0%	4.2%	0.5%	23.0%	19.3%	2.7%

*No detection of antibiotic resistance genes in Other Animals (n=5)

Figure 5-14 displays the historical importation of live cattle into New Zealand. The vast majority of cattle were imported in the 1860's, while importations increased during the 1950's to 1990's. A comparison between live importation of cattle into both New Zealand and Japan from 1961 to 2013 is shown in Figure 5-15. New Zealand imported fewer cattle than Japan for all years examined; since 1991, New Zealand has consistently imported less than one hundred live cattle a year, while Japan has imported over 10,000 a year (Figure 5-15). In contrast, the resident cattle stock numbers present in New Zealand were higher (6.4 million in 1961, to 10.2 million in 2016) when compared with Japan (3.2 million in 1961, to 3.8 million in 2016) (246).
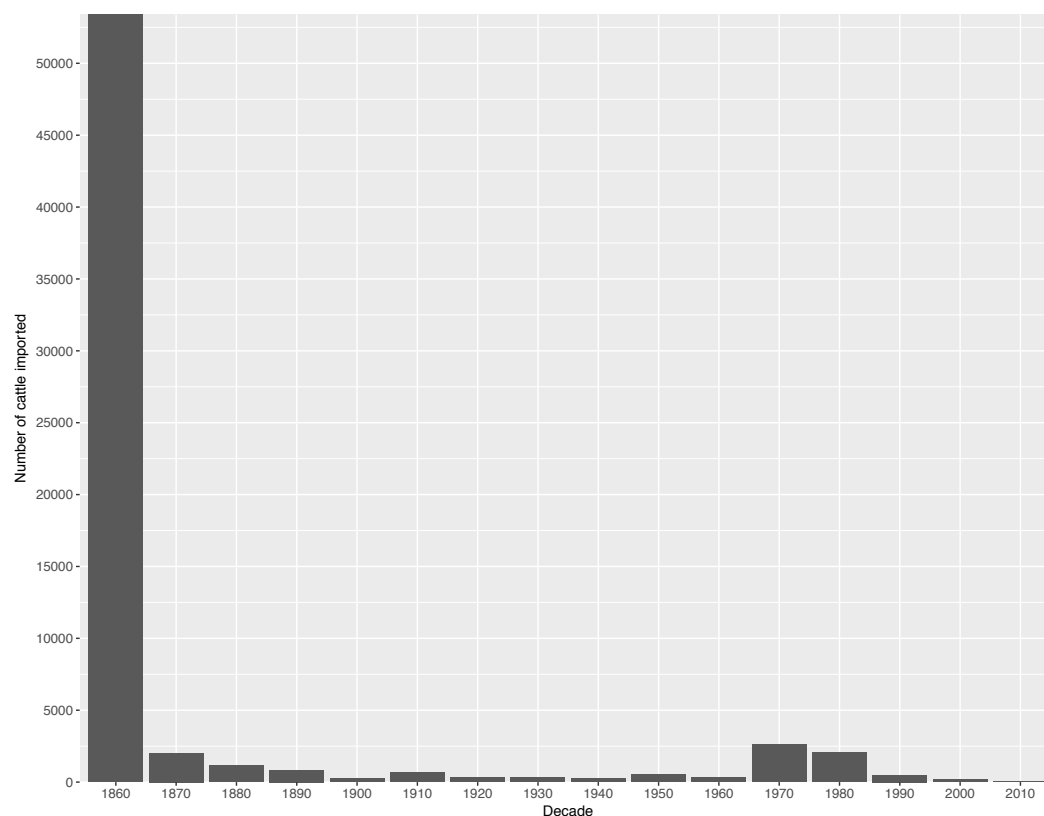
Figure 5-14: Bar graph of historical importations of live cattle into New Zealand between 1860 and 2010 by decade
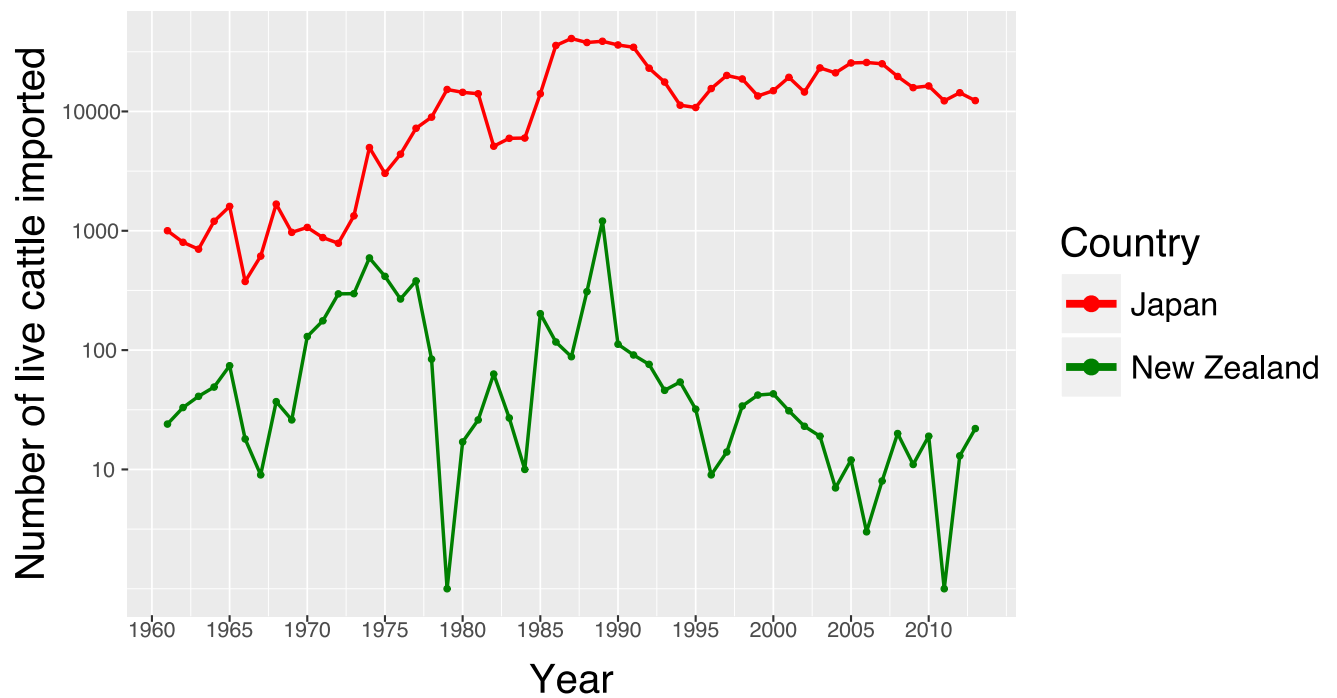
Figure 5-15: Comparison of live cattle imported ($\log_{10}$ scale) into New Zealand and Japan from 1961 to 2013

5.5 Discussion

This study used whole genome sequencing (WGS) to compare New Zealand *E. coli* serogroup O26 bacteria with isolates from around the world. Analyses revealed multiple strains are present globally, with periods of between-country transmission occurring mainly during the 20th century. Phylogenetic analysis and examination of virulence and antibiotic resistance genes suggests several strains have differentiated within New Zealand following few periods of introduction.

4.8.14 Evolutionary dynamics of *E. coli* serogroup O26

Comparative genomic analysis indicated that serogroup O26 *E. coli* had an open genome (Figure 5-1, Heap's Law coefficient 0.35). Evolution based on recombination and horizontal gene transfer is predicted by the Public Goods Hypothesis for the evolution of life on earth (43). In contrast to a Tree of Life evolutionary model (44),

the Public Goods Hypothesis proposes that the horizontal exchange of widely available DNA sequences, notably genes, are the primary driver for local bacterial evolution (43). Applying this hypothesis to the evolution of serogroup O26 *E. coli* implies continued adaptation of the bacterial strains in local geographical environments as they continue to acquire and share genes. Ultimately, this could lead to the emergence of new pathogenic lineages of STEC.

Gene counts (Figure 5-3) and functional annotation (Figure 5-5) did not reveal any obvious patterns based on the factors evaluated in this study (sequence type, country, isolation source), with the exception of fewer genes present in isolates that did not contain a *stx* gene. Given the evidence that serogroup O26 *E. coli* has an open genome, it is possible bacteria are constantly acquiring and removing genes that perform different functions. While this process may lead to increased survival, it is also possible that the process is random and there is a lack of selection pressure for isolates to obtain genes with specific functions (e.g. defence mechanisms).

A previous analysis of 53 *E. coli* genomes from various serogroups and sources estimated a core genome of 1,472 genes and an accessory genome of 13,296 genes (19). This study of serogroup O26 *E. coli* found a core genome of 2,718 core genes, 8,904 accessory genes, and 9,777 singleton genes, for a pangenome size of 21,399 genes. Although it may seem remarkable that a selection of *E. coli* from a single serogroup would have a greater number of genes detected than various *E. coli* (19), this difference can be attributed to the open genome, high level of genomic plasticity of *E. coli*, and the larger sample size in this study (404 genomes compared to 53 genomes).

One example of ongoing STEC evolution is the emergence of serogroup O26 ST-29 *E. coli* in Europe. Historically, STEC O26 harboured the less pathogenic *stx1* gene, and most of the pathogenic STEC O26 were part of the ST-21 sequence type. Beginning in the late 1990's, a new strain of STEC O26 ST-29 with the more virulent *stx2* emerged in Germany (46), and more recently in France (47), and Switzerland (48). This strain is more pathogenic, leading to an increased incidence of HUS and illness in adults. Controlled experiments revealed that *stx2* can be acquired and lost in serogroup O26 isolates, leading to dynamic changes in pathogenicity (49). The emergence of a highly virulent ST-29 STEC O26 clade has been identified in Japan since 2012, and genomic analysis revealed it is unique from the European variant; this provides another example of evolution of this serogroup (50).

4.8.15 Between and within-country differentiations

The hierarchical set pangenome tree illustrated clades of New Zealand isolates that were monophyletic, while most foreign isolates were paraphyletic, interspersed with isolates from other countries (Figure 5-6). Pangenome and virulence gene PERMANOVA analysis (Table 5-2) indicated that variability was best explained by multilocus subtype, country of origin, and *stx* profile. The virulence gene PERMANOVA analysis suggests that key genetic differences of strains between countries are due to acquisition or loss of virulence genes, suggesting convergent evolution may have contributed to adaptation in different ecosystems (i.e. homoplasy). The lack of isolation source as a significant factor for pangenome or virulence gene analysis (Table 5-2) suggests that serogroup O26 isolates present in humans, cattle, food, and other animals are not genetically differentiated and transmissions of this organism occurs frequently. From an evolutionary perspective,

the presence of a non-toxigenic clade of ST-21 *E. coli* from New Zealand (Figure 5-6, Figure 5-7) indicates that these bacteria have not simply lost the *stx* bearing bacteriophage *in vitro* during laboratory isolation, but have a fixed evolutionary pathway that does not contain the *stx* gene.

Whole genome sequencing analysis of O157:H7 showed similar findings to our results, where the within-country host diversity was unique, compared to the diversity between sources (45, 247). Molecular typing of STEC O157 from New Zealand, Australia, and the USA indicated strong population differentiation and a unique genotype distribution for each country (45). The same study revealed patterns of local spatial clustering of genotypes in New Zealand (45). Multiple occurrences of international transmission, followed by local genetic evolution and differentiation were also identified in another study of O157:H7 (247). Genes that are classified as virulent for humans are important for the intestinal colonization of cattle (248, 249), therefore the contrasting virulence attributes associated with isolates from different countries may indicate separate niche adaptation and advantages for colonization of local host populations.

4.8.16   Pathogenicity of STEC O26

A wide array of virulence genes (n=192) were detected in the analysed bacterial isolates (n=404). Groups of isolates with identical virulence gene profiles were frequently observed in isolates from New Zealand, Belgium, USA, and Japan (Figure 5-7).

Hierarchical clustering of virulence genes identified three clades, separated by the presence of *stx1*, *stx2*, and *eae* (Figure 5-8). These clusters of virulence genes may indicate pathogenicity islands introduced by bacteriophage (26). Acquiring new traits

in a given environment can provide fitness and increase survival of pathogens (27), for example it has been proposed that STEC are more resistant to consumption by protozoa due to the presence of Shiga toxin, offering a survival advantage to the bacteria (31).

Evidence of relatively recent acquisition of the *stx2* virulence gene within STEC O26 ST-29 is a cause of concern (47). Other research has noted the potential for cattle ST-29 strains to acquire the *stx2*-containing bacteriophage (250). The recombination analysis (Figure 5-10, Figure 5-11) showed evidence of a difference between recombination rates in ST-21 (n=7 areas) and ST-29 (n=38 areas), suggesting ST-29 may have a greater propensity for horizontal gene transfer. At the time of publication, no STEC O26 with the *stx2* virulence gene has been isolated from animals or humans in New Zealand, even though non-STEC ST-29 strains are present in cattle. In a controlled study of 4 *stx*-negative O26:H11 strains and *stx2* prophages, lysogeny could not be induced, and sensitivity of strains to the prophage varied, indicating that assimilation of *stx2* into the bacterial genome of ST-29 isolates may be a rare event (251). However, the presence of the *stx2* virulence gene in highly pathogenic STEC O26 ST-29 in other countries indicate horizontal genetic transfer of *stx2* has occurred, although this phenomenon has not been observed in New Zealand. The 2011 outbreak of STEC O104:H4 in Germany illustrated the clinical significance of the emergence of new lineages following the acquisition of virulence genes (51).

4.8.17 Antibiotic resistance profiles and evolution

The resistance profiles form distinct combinations of resistance genes in isolates from particular countries (Figure 5-6, Figure 5-7). The use of antibiotics is not usually prescribed for human STEC infections (99), however selection pressure from antibiotic

use in livestock and humans with undiagnosed diarrheal illness may influence the evolution of resistance.

A comprehensive study of over two hundred *E. coli* strains containing *eae* isolated from various livestock species found that over 65% of the strains were resistant to tetracycline, streptomycin, erythromycin, and sulfamethoxazole, with resistance to ampicillin, chloramphenicol, trimethoprim, and trimethoprim-sulfamethoxazole being less common (approximately 30%) (33). The lower prevalence of antibiotic resistance genes in New Zealand isolates in this study could be related to the lower use of antibiotics for animals compared to other countries. A recent assessment of antimicrobial use from New Zealand, as well as 26 European countries, Australia, Canada, and the USA, found that New Zealand was the third lowest user of antimicrobials for animals (mg active ingredient of antimicrobial per kg of biomass) (252). This study also reported that antimicrobial use in humans exceeded animal use by 12.8 times in New Zealand (252). The presence of antibiotic resistance genes in serogroup O26 *E. coli* may be due to a combination of human and animal antibiotic administration.

4.8.18 Inferred global importation and transmission of *E. coli* O26

The time of most recent common ancestor (TMRCA) of New Zealand O26 clades suggests several separate importations of strains that appear to coincide with cattle importation events (Figure 5-12, Figure 5-13, Figure 5-14). The presence of at least four distinguishable New Zealand ST-21 monophyletic clades (Figure 5-12), and three New Zealand ST-29 monophyletic clades (Figure 5-13), supports the view of multiple introductions, followed by independent evolution in New Zealand. Phylogenetic analysis suggests certain New Zealand clades are more associated with specific

countries (e.g. USA or Continental Europe), indicating there are likely to be transmission pathways via live animal imports.

The TMRCA of all analysed globally circulating ST-21 serogroup O26 *E. coli* was estimated to be between 1751 and 1838 (95% HPD interval) (Figure 5-12). This period is earlier than TMRCA estimates of 175 years ago for STEC O157:H7 isolated in England (253), but similar to estimates for serogroup O26 *E. coli* by Ogura et al. of 1772-1831 (95% HPD interval) (254). The TMRCA of all ST-29 strains was estimated to have been earlier, but with wide credibility intervals of 1349 to 1656 (95% HPD interval; Figure 5-13). This is consistent with the estimated origin of all ST-29 O26 in another study (1531-1663 95% HPD interval) (254). It has been hypothesized that STEC O26 ST-21 emerged from non-toxigenic attaching and effacing *E. coli* (AEEC) strains following the acquisition of key virulence genes (49). This analysis did not establish a timeline associated with important horizontal gene transfer events, but these results indicate that ST-29 is an older lineage of the *E. coli* O26 serogroup than ST-21. The estimated mutation rate for ST-21 and ST-29 isolates in the present study are similar to previous estimates for serogroup O26 ($2.8\text{-}4.3 \times 10^{-7}$) (254) and O157:H7 (253).

4.8.19 Global transmission of STEC O26 via the movement of live cattle

The Food and Agriculture Organization of the United Nations estimates that 1.4 billion cattle are kept worldwide (123). Movement of cattle between Europe, the Americas, Japan, Australia, and New Zealand increased during the late 19th century. For example, annual live cattle exports from the USA to Great Britain dramatically increased from 299 to 137,377 between 1875 and 1885, only to quickly decrease with the improvement of refrigeration and growing concerns over infectious disease, specifically pleuro-pneumonia (255).

A remarkable difference was seen between two islands nations with the most bacterial isolates analysed in this study: New Zealand (n=152) and Japan (n=94). In the hierarchical set pangenome (Figure 5-6), virulence gene (Figure 5-7), and BEAST2 ST-21 (Figure 5-12) analyses, New Zealand isolates show monophyletic clades, while Japanese isolates are paraphyletic with isolates from the USA. This difference may be explained by the historically large number of live cattle importations into Japan compared to New Zealand (Figure 5-15). Japan was once a major importer of U.S. live cattle in the second half of the 20th century, until the detection of bovine spongiform encephalopathy (BSE) led to a ban on all live cattle from the USA in 2003, with current importations coming from Australia (256, 257). Total numbers of cattle in New Zealand have been consistently twice that of cattle in Japan over the last 55 years (10.2 million cattle in New Zealand in 2016 compared to 3.8 million cattle in Japan). The relatively large number of imported live cattle in Japan could therefore explain the different population structure of serogroup O26 *E. coli* in New Zealand and Japan.

The growing utilization of artificial insemination using cryofrozen semen has allowed countries to maintain genetically diverse and viable livestock, while selecting for traits that may benefit production. An added benefit is the decreased disease risks and costs of transportation of live animals into a country. While intensification of cattle farming may be associated with increased pathogen transmission and prevalence within herds and within countries, incursion of pathogens into countries may have decreased. Our dataset allows minimal interpretation of open border areas such as the European Union, or countries in the North American Free Trade Agreement (Mexico, Canada, and the USA), but our results from New Zealand suggest the introduction of serogroup O26 bacterial strains occurred during periods of intensive cattle importation events.

STEC in cattle is a commensal bacteria and cattle shed STEC intermittently (182), therefore testing cattle before transportation is unrealistic. Our TMRCA and phylogenetic analyses suggest that minimal exchange of strains has occurred between countries in the 21st century, but continued movement of cattle across international borders will likely continue to influence the spread and genetic diversity of STEC around the world.

4.8.20 Study limitations

Limitations of our study included the lack of availability of serogroup O26 *E. coli* from other countries for analysis. Notably, more serogroup O26 *E. coli* isolates from Australia (n=1 available on public databases) would have allowed for a better comparison of the effect of importation of cattle into New Zealand, as Australia was the source of many historical cattle importations (245). Sequence data were more common from the last several years, and mostly human isolates were available, due to increased use of whole genome sequencing as an epidemiological tool. However, we randomly selected our New Zealand isolates from human and bovine isolates spanning over 30 years from a diverse geographic range in New Zealand, allowing a more thorough comparison of New Zealand strains against others worldwide.

Our focus on a particular O surface antigen (O26) to classify bacterial isolates and evaluate evolutionary and phylogenetic relationships is consistent with other studies (253, 254). However, non-STEC and STEC strains of the same serogroup are commonly of different lineages (88). Epidemiological and public health investigations of STEC are often focused on the presence of particular serogroups (e.g. the 'Top 7' STEC: O157, O26, O45, O103, O111, O121, O145), therefore this approach is relevant for both international beef trade and outbreak investigations.

5.6 Conclusion

Our results suggest worldwide dissemination of multiple strains of ST-21 and ST-29 STEC and non-toxigenic serogroup O26 lineages occurred during the 20th century. This differentiation may be explained by combinations of genetic factors, which may give an advantage for survival in the environment or colonization of reservoir animals in the host population. Close genetic similarities between multiple sources of *E. coli* serogroup O26 indicates common transmission pathways between humans, animals, and food sources.

Whole genome sequencing offers high resolution information on pathogens. While the rapid sequencing of human clinical isolates is becoming increasingly common, further sequencing of historical isolates from multiple sources will improve evolutionary and epidemiological studies; the recent submission of a large number of human and bovine bacterial isolates from Japan spanning several decades is an excellent example (254). Full utilization of the genomic information of STEC will require a coordinated international approach to sequencing, data curation, analysis, and interpretation of that data (258). While it is difficult to directly attribute transmission and emergence of STEC strains based on global historical events, interpreting evolutionary genomic data against a background of historical events can help determine the drivers of pathogen emergence and dissemination, and inform future policy.

5.7 Acknowledgments

maintain whole genome sequencing data which can be used for public health research; particular thanks to our colleagues in Japan, Ogura et al., whose dataset provided such an excellent comparison to ours.

We also wish to acknowledge the contribution of NeSI high-performance computing facilities to the results of this research, provided by the New Zealand eScience Infrastructure and funded jointly by NeSI's collaborator institutions and through the New Zealand Ministry of Business, Innovation & Employment's Research Infrastructure program.

The purpose of scientific enquiry is not to compile an inventory of factual information, nor to build up a totalitarian world picture of natural Laws in which every event that is not compulsory is forbidden.

We should think of it rather as a logically articulated structure of justifiable beliefs about nature.

Induction and Intuition in Scientific Thought, Sir Peter B. Medawar

6 General discussion and future research opportunities for Shiga toxin-producing *E. coli* (STEC) in New Zealand

6.1 Introduction

This thesis utilised molecular epidemiology, genomics, microbiology, and statistical methods to understand the national prevalence of STEC in young calves on New Zealand dairy farms (**Chapter 3**), while determining STEC transmission routes for calves on dairy farms and potential sources of cross contamination of hides and pre-intervention carcasses at processing plants (**Chapter 4**). The application of WGS analyses predicted the time of introduction of STEC O26 into New Zealand and analysed serogroup O26 *E. coli* isolates in New Zealand, in comparison to the rest of the world (**Chapter 5**).

All research questions (**Chapter 1**) were effectively addressed in each research chapter. In **Chapter 3**, a randomly stratified cross-sectional study detected 'Top 7' STEC in 20% of calves (n=1508) on 75% of dairy farms (n=102) evaluated. Each region (n=6) evaluated varied in the prevalence of each 'Top 7' STEC serogroup. WGS analyses indicated that there was no evidence of farm to farm transmission within regions, and serogroup O26 had clonal strains that were disseminated between calves on each farm. Decreasing pen-occupancy rates in calf pens was identified as a feasible and practical intervention to decrease STEC carriage of calves on dairy farms. **Chapter 3** also utilised and evaluated the PCR/MALDI-TOF method, NeoSEEK, and found a very

high specificity and good sensitivity for detection of the 'Top 7' STEC serogroups, when compared to RT-PCR methods. This method was utilised by commercial meat processors in New Zealand for the confirmation of 'Top 7' STEC in veal trim beginning in the 2016 calving season.

In **Chapter 4**, a longitudinal study of animals and the dairy farm environment identified widespread environmental and hide contamination, as well as colonisation of animals with 'Top 7' STEC, for all sample sources evaluated in the study (n=17). WGS analyses indicated that clonal strains of STEC O26 were widely disseminated through the farm environment and had colonised very young calves. Pen-occupancy rates were found to significantly increase the risk of 'Top 7' STEC contamination of hides on dairy farms. Sampling of a cohort of bobby calves (n=118), on-farm and at the processing plant, revealed that transport and lairage were significantly increasing hide contamination with 'Top 7' STEC, and WGS analyses revealed calf hides were being cross-contaminated with novel strains during transport. Increasing the number of farms visited by the bobby calf transport truck increased the risk of pre-intervention carcass contamination, therefore it is likely that the mixing of calves from different farms during transport leads to cross-contamination of calf hides, which can lead to contamination of pre-intervention carcasses at the processing plant.

In **Chapter 5**, WGS analyses of New Zealand serogroup O26 isolates (n=152), along with bacterial isolates from fourteen other countries (n=252), revealed that New Zealand has experienced few introductions of the O26 serogroup into New Zealand compared to other countries. Time of most recent common ancestor analysis (TMRCA) predicted key periods of introduction in the 20th and early 21st centuries.

New Zealand had low prevalence of antibiotic resistance genes detected compared to other countries, possibly due to lower use of antibiotics for animals in this country. When compared to another island nation, Japan, which showed paraphyletic clades with the USA, New Zealand had monophyletic clades of serogroup O26, and this finding may be due to the low numbers of live cattle importations into New Zealand. A highly pathogenic STEC O26 ST-29 strain, that has the *stx2* gene, is present in Europe and Japan but has never been detected in New Zealand.

6.2   Potential impact of thesis findings

4.8.21   Market and trade impact

This thesis addressed three key objectives of a MPI and the Strategic Directions Group (SDG) of the Meat Industry Association (MIA) work programme for STEC: assessment of practical and cost-efficient laboratory methods for detection and confirmation of STEC, understanding the epidemiology of STEC in animal populations and impacts on beef carcasses, and assessment of risk factors and interventions that can be applied at the farm level and pre-slaughter to reduce beef contamination. This research provided data that spanned the supply chain gap between determining STEC prevalence on farm, and regulatory screening of veal performed at processing plants.

This study supports the objectives of New Zealand Foreign Affairs & Trade's strategies for NZ Inc.: to strengthen New Zealand's economic, political, and security relationships with international partners. STEC are a trade concern, particularly with the USA, which receives approximately 50% of New Zealand beef exports (14). The application of directed scientific studies provides evidence that can be used to inform policy and market access decisions.

This research also provided data that contributed toward the implementation of the NeoSEEK assay for STEC confirmation for veal carcasses in New Zealand. Before the implementation of NeoSEEK, STEC screening and confirmation required two to three weeks to complete; after implementation, STEC screening and confirmation was completed in three days. The marked decrease in holding time for potentially contaminated veal has led to significant cost savings for New Zealand beef processors.

4.8.22 Public health impact

While a previous study had estimated the prevalence of STEC O157 and STEC O26 in New Zealand by sampling cattle and calves at processing plants (127), this thesis provided additional knowledge of STEC prevalence: 1) How common is STEC in New Zealand calves? and 2) How widespread is STEC contamination in the farm environment? STEC is commonly found on New Zealand dairy farms. From dissemination of STEC across the dairy farm environment, to approximately 75% farm level and 20% calf level colonisation of STEC in young dairy calves, STEC is a widespread coloniser and contaminant.

Understanding the prevalence of STEC on New Zealand dairy farms is directly relevant for informing decisions for public health interventions. These data can be used to effectively communicate the risk of STEC transmission, particularly to children, in the dairy farm environment, and inform future public health interventions to decrease the incidence of STEC in New Zealand.

6.3   Discussion of potential opportunities for future research

4.8.23 On farm interventions to decrease STEC prevalence

This thesis provided evidence that increasing the number of calves in a shared pen environment both increased the risk of STEC colonisation in young calves (**Chapter**

3), and STEC hide contamination of calves on dairy farms (**Chapter 4**). These findings indicate that frequency-dependent transmission is important for STEC in calf pens, while density-dependent transmission may not have a significant impact. The absolute number of calves in a shared environment is linked to the biological amplification of STEC bacteria, as each animal is colonised and sheds the bacteria in its faeces.

Decreasing the number of calves in a shared pen area is a viable intervention that could be implemented on dairy farms. For instance, 49% of farmers in the national prevalence study (**Chapter 3**) already isolate bobby calves in a separate pen from replacement calves. Asking dairy farmers to limit the number of bobby calves in a pen is a realistic request, which may have a marked effect on calf level STEC prevalence. While farmers may have limited pen space, placing a simple barrier inside a bobby calf shed could help decrease transmission. Ideally, bobby calves would be isolated from each other in an outdoor calf hutch, which is a method used in the USA to decrease disease transmission (particularly respiratory disease); this practice may have animal welfare applications and therefore may not be appropriate in New Zealand. Future research could investigate the effect of frequency-dependent transmission of STEC for calves in a controlled-environment or field study.

4.8.24 Improvement of bacterial isolation methods using culture media

The disparity in STEC detection between molecular methods and bacterial isolation was described in the literature review (**Chapter 2**) and the national prevalence study (**Chapter 3**). Similar to other studies, it was difficult to isolate 'Top 7' STEC that were detected in this study, with the exception of serogroup O157 and O26. A recently submitted Massey University Master of Science thesis evaluated the metabolic characteristics of serogroup O145 from New Zealand using both *in vitro* metabolic

profiles (Omnilog®, Biolog, Hayward, USA) and WGS (R Collis, Metabolic characteristics and genomic epidemiology of *Escherichia coli* serogroup O145). Approaches such as this may offer improved culture-based detection methods for analysis and differentiation of non-O157 STEC serogroups that can be coupled with molecular methods for improved detection and isolation.

However, the differentiation of STEC on agar media based on metabolic characteristics, in order to differentiate both serogroup and pathogenicity (*stx* gene present), may not be possible. Given the acquisition of pathogenic genes can occur through horizontal transmission, there is little reason to expect that fixed metabolic traits will be correlated with pathotypes.

Bacterial isolation using agar media is vital for WGS library preparation of individual bacteria, and valuable for epidemiological source attribution studies. However, this thesis showed that the use of a culture independent diagnostic test (NeoSEEK) is valuable for epidemiological studies. Overall, future STEC research can utilise highly sensitive and specific CIDTs while continuing to collect bacterial isolates for other elements of epidemiological studies.

4.8.25 Knowledge and public health risk communication of STEC

Public knowledge of STEC is lacking. Ongoing research data from Taranaki, the second largest dairy region in New Zealand, found that eleven out of twelve families whose young children were diagnosed with an STEC infection had no prior knowledge of the disease (J Jarman, Taranaki Health Board, unpub. data). During field sampling for the national prevalence study (**Chapter 3**), only two farmers (out of 102

interviewed) that participated in the study were aware of STEC: one large animal veterinarian, and one former microbiologist (AS Browne, personal observation).

Efforts to lower the incidence of STEC in New Zealand should focus on educating the public, particularly families in rural environments, of the risk of STEC infection for their very young children. Several studies have identified the benefits of handwashing and hygiene to lower risk of STEC infection in humans (227). Furthermore, this thesis has reinforced that the calf pen environment, in particular, has a high prevalence of contamination. It is reasonable to suspect that the biphasic peaks in incidence of STEC notification in New Zealand are related to the spring calving season, and increased recreational swimming in the summer (8), meaning avoidance of STEC exposure during the calving season would decrease the incidence of STEC during this period.

Health awareness campaigns for the general public, or directed more specifically at dairy farmers and their families, may decrease STEC exposure and decrease the incidence of STEC in New Zealand (259). These public health communications should be designed as scientific research studies, to evaluate the effectiveness of the communication and impact on decreasing the incidence of STEC. Public communications should focus on 3 talking points: recognizing the clinical signs of STEC infection in children (bloody diarrhoea, no urine production), avoiding young children having unsupervised contact with young calves, and the importance of hand-washing. It was common for dairy farmers to spend time with their children feeding calves after milking (AS Browne, personal observation during field work over three Spring calving seasons); it is important for farmers to share time with their children while having positive experiences with cattle. Public health recommendations should

avoid absolutist statements (e.g. "Don't allow children to have any contact with calves"), and focus instead on communicating the increased risk of unsupervised contact and the importance of hygiene to prevent infection. Farmers work in dangerous environments every day, and they are capable of understanding that certain activities can increase risk.

Efforts to communicate the risk of STEC in farm environments should involve relevant interest groups like DairyNZ, Rural Women New Zealand, Federated Farmers, Young Farmers, and the Māori-owned Miraka dairy cooperative. The District Health Boards and Starship Children's Hospital are also valuable resources, for both public health communication and to understand how those who fall ill to STEC perceive their situation. These groups may be able to communicate issues in a personal way that is trusted by their audience, in comparison to large government entities that may be perceived as being overcautious.

Proper feedback following scientific studies is also vital. Every farmer that was involved in studies described in this thesis (n=102) was contacted by phone to convey the results of the STEC testing and provided with further advice to avoid STEC infection. Farmers are incredibly generous and accommodating to participate in scientific studies without financial or material gain, and it is imperative that researchers take time to communicate their results back to them in a timely manner.

4.8.26 Raw milk consumption

Although not identified as a risk factor in a prospective case-control study (12), raw milk consumption has been linked to several STEC outbreaks in New Zealand. For

instance, consumption of raw milk by children on a school visit to a dairy farm in

Timaru led to five infections, with two children hospitalised (260).

The recent events surrounding two 2016 STEC outbreaks in Whangarei, linked to the

raw milk producer, Wholy Moo, provide an excellent example of the complexity of the

raw milk industry (261). Two outbreaks of STEC O157:H7 occurred in eleven people

(children and adults) who had consumed raw milk from Wholy Moo. The owner, Chris

Lethbridge, stated that he closed his business after compliance costs were expensive

and prohibitive (261). Mr. Lethbridge stated: "We all get sick but just because they

were drinking raw milk means they [raw milk suppliers] get picked on. MPI think they

have proven [a link between the illnesses and his product] but I can't see how they

have" (261). He also was 99.9 percent sure that no *E. coli* was present in his milk

because, "We can prove that we were grade free in the milk we supplied Fonterra all

this current season" (261). These statements highlight two important factors with

respect to raw milk producers and consumers: 1) any regulation or scientific evidence

of the risk of STEC infection with raw milk consumption is likely to be perceived as

prejudicial or biased, and 2) producers and consumer may perceive themselves more

knowledgeable about infectious disease than they are; simple laboratory testing of

coliform contamination is not indicative of absence of a significant pathogen like

STEC or *Campylobacter* (e.g. grade free simply indicates the sample was below a

certain CFU/mL threshold).

Despite good hygiene practices on dairy farms, complete prevention of faecal

contamination of raw milk is very unlikely. Although not focused on STEC

contamination of raw milk, this thesis provided evidence of 'Top 7' STEC

contamination in milk filters and colostrum on dairy farms (**Chapter 4**). Targeted

public outreach should stress the risk of raw milk consumption for very young

children, who are at the biggest risk for severe STEC infection. Similar to previous

comments regarding public health communication, the message may be better suited

to be distributed by special interest groups (e.g. DairyNZ, Rural Women). Cross-

sectional studies to evaluate risk factors for STEC contaminated raw milk sold by dairy

farms may help identify improved farm practices, as well as the prevalence of STEC in

raw milk throughout New Zealand.

4.8.27 Use of whole genome sequencing (WGS) technology

This thesis generated WGS data from a large number of 'Top 7' STEC isolates from

humans and cattle that were stratified over time in many regions of New Zealand. The

WGS processing of a large number of bacterial isolates led to decreased costs and

increased capability for library preparation at ^{m}EpiLab, as well as the implementation

of technology to streamline assembly, annotation, and analysis of genomes. While the

WGS data produced was evaluated from a broad epidemiological viewpoint for this

thesis, these data can be further evaluated for future investigative research, such as

phenotype-genotype association studies. These data have also been entered into the

public record and will be released as manuscripts are published, which will increase

the profile of STEC research in New Zealand for the international science community.

The more recent development and emergence of nanopore technology, such as the

Oxford Nanopore MinION genome sequencing instrument, shows great promise in

the future of whole genome sequencing. Specific benefits include portability, such as

its field use in 2015 during the Ebola outbreak in Guinea (262), and its ability to

produce long read length, up to the point of creating an accurate and complete

bacterial genome *de novo* (263). *ᵐ*EpiLab recently acquired this technology, and the Massey Genome Service also has personal expertise in the use of MinION. As WGS becomes less expensive and, this technology may be used for field investigations, such as STEC outbreak investigations, as well as applied research.

Overall, it is vital that research that utilises WGS is hypothesis driven, and not simply descriptive. This thesis approached WGS data with specific research questions (**Chapter 1**), which led to specific discoveries regarding STEC and non-toxigenic variants of O26 in New Zealand.

4.8.28 Waterway contamination

The New Zealand dairy and beef industries have been under scrutiny for the impact of cattle faecal contamination in natural waterways (264). Recreational swimming was also previously recognized as a risk factor for human STEC infection in New Zealand (12). Further studies linking water sources to STEC infections in people are needed to identify high risk areas that can be targeted for interventions. Clinicians at the Starship Children's Hospital Paediatric Nephrology service in Auckland identified high rain events and potential private bore contamination as a significant cause of HUS in hospitalised children (W Wong, personal communication). Significant rain events can lead to faecal contamination of rivers, as well as private bores, and increased scrutiny of water sources during these times may help decrease the risk of STEC transmission. For instance, heavy rains were linked to the 2016 Havelock North *Campylobacter* outbreak, which caused 5,500 people to become ill with campylobacteriosis due to a contaminated water bore (265). The One Health Aotearoa initiative (https://onehealth.org.nz/about-us/), a collaboration between academic, industry, and government partners, is approaching infectious disease from an animal, human, and

environmental perspective; this One Health approach would be beneficial for addressing STEC contamination of water sources.

6.4   Conclusion

This research benefited from synergistic cooperation between academic, government, and industry partners in order to attain common goals. Aspects of the thesis, from direct communication with government and industry stakeholders, to access to meat processing facilities, to cutting edge genomic analytic approaches, would not have been possible without the cooperative efforts of government, academic, and industry partners. Opening communication pathways by visiting and discussing STEC with the Taranaki Health Board and clinicians at the Starship Children's Hospital are good examples of broadening academic research to effectively communicate results and begin to influence policy to decrease the incidence of STEC in New Zealand. Hopefully, the successful completion of this thesis to answer industry-driven research questions will encourage future research collaborations.

This thesis has provided novel data and interpretation for the prevalence and transmission of the 'Top 7' STEC of dairy calves in New Zealand on a farm and national level, while applying genomic analyses to interpret New Zealand STEC on an international level. Future STEC research in New Zealand should incorporate molecular epidemiology, genomics, statistical modelling, sociology, and market analysis to decrease the incidence of STEC in the New Zealand human population and safeguard the trade of New Zealand beef around the world.

7   Bibliography

1.   Bielaszewska M, Mellmann A, Zhang W, Köck R, Fruth A, Bauwens A, Peters G, Karch H. 2011. Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. Lancet Infect Dis 11:671–6.

2.   Centers for Disease Control and Prevention. 2016. *Escherichia coli* O26 infections linked to Chipotle Mexican Grill restaurants (Final update). https://www.cdc.gov/ecoli/2015/o26-11-15/.

3.   Rivas M, Sosa-Estani S, Rangel J, Caletti MG, Vallés P, Roldán CD, Balbi L, Marsano de Mollar MC, Amoedo D, Miliwebsky E, Chinen I, Hoekstra RM, Mead P, Griffin PM. 2008. Risk factors for sporadic Shiga toxin-producing *Escherichia coli* infections in children, Argentina. Emerg Infect Dis 14:763–71.

4.   Majowicz SE, Scallan E, Jones-Bitton A, Sargeant JM, Stapleton J, Angulo FJ, Yeung DH, Kirk MD. 2014. Global incidence of human Shiga toxin–producing *Escherichia coli* infections and deaths: A systematic review and knowledge synthesis. Foodborne Pathog Dis 11:447–455.

5.   Rosales A, Hofer J, Zimmerhackl L-B, Jungraithmayr TC, Riedl M, Giner T, Strasak A, Orth-Höller D, Würzner R, Karch H. 2012. Need for long-term follow-up in enterohemorrhagic *Escherichia coli*-associated hemolytic uremic syndrome due to late-emerging sequelae. Clin Infect Dis 54:1413–21.

6.   Duffy G, Burgess CM, Bolton DJ. 2014. A review of factors that affect transmission and survival of verocytotoxigenic *Escherichia coli* in the European farm to fork beef chain. Meat Sci 97:375–383.

7.   Crump JA, Murdoch DR, Baker MG. 2001. Emerging infectious diseases in an island ecosystem: The New Zealand perspective. Emerg Infect Dis 7:767–772.

8.   The Institute of Environmental Science and Research Ltd. 2017. Notifiable diseases in New Zealand: Annual report 2016. https://surv.esr.cri.nz/PDF_surveillance/AnnualRpt/AnnualSurv/2016/2016Annu

alNDReportFinal.pdf.

9.   Marder EP, Cieslak PR, Cronquist AB, Dunn J, Lathrop S, Rabatsky-Ehr T, Ryan
        P, Smith K, Tobin-D'Angelo M, Vugia DJ, Zansky S, Holt KG, Wolpert BJ, Lynch
        M, Tauxe R, Geissler AL. 2017. Incidence and trends of infections with pathogens
        transmitted commonly through food and the effect of increasing use of culture-
        independent diagnostic tests on surveillance — Foodborne Diseases Active
        Surveillance Network, 10 U.S. Sites, 2013–2016. MMWR Morb Mortal Wkly Rep
        66:397–403.

10.   European Centre for Disease Prevention and Control. 2017. Disease data from
        ECDC Surveillance Atlas - *Escherichia coli* infection.
        https://ecdc.europa.eu/en/escherichia-coli-ecoli/surveillance/atlas.

11.   The Institute of Environmental Science and Research Ltd. 2016. Surveillance
        report: Notifiable diseases in New Zealand 2015. www.surv.esr.cri.nz.

12.   Jaros P, Cookson AL, Campbell DM, Besser TE, Shringi S, Mackereth GF, Lim E,
        Lopez L, Dufour M, Marshall JC, Baker MG, Hathaway S, Prattley DJ, French NP.
        2013. A prospective case-control and molecular epidemiological study of human
        cases of Shiga toxin-producing *Escherichia coli* in New Zealand. BMC Infect Dis
        13:450.

13.   Food Safety and Inspection Service. 2012. Shiga toxin-producing *Escherichia coli*
        in certain raw beef products. Fed Regist 77:31975–31981.

14.   Meat Industry Assocation. 2017. 2017 Annual Report.
        https://www.mia.co.nz/assets/Annual-Reports/MIA-Annual-report-2017-
        online.pdf.

15.   Riley LW, Remis RS, Helgerson SD, McGee HB, Wells JG, Davis BR, Hebert RJ,
        Olcott ES, Johnson LM, Hargrett NT, Blake PA, Cohen ML. 1983. Hemorrhagic
        colitis associated with a rare *Escherichia coli* serotype. N Engl J Med 308:681–
        685.

16.   Centers for Disease Control and Prevention (CDC). 1993. Update: multistate
        outbreak of *Escherichia coli* O157:H7 infections from hamburgers--western

United States, 1992-1993. MMWR Morb Mortal Wkly Rep 42:258–63.

17.   The Institute of Environmental Science and Research Ltd. 2016. Public Health Surveillance Report. New Zeal Public Heal Surveill Rep 14:1-A3.

18.   Brooks JT, Sowers EG, Wells JG, Greene KD, Griffin PM, Hoekstra RM, Strockbine NA. 2005. Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983-2002. J Infect Dis 192:1422–9.

19.   Lukjancenko O, Wassenaar TM, Ussery DW. 2010. Comparison of 61 Sequenced *Escherichia coli* genomes. Microb Ecol 60:708–720.

20.   Wirth T, Falush D, Lan R, Colles F, Mensa P, Wieler LH, Karch H, Reeves PR, Maiden MCJ, Ochman H, Achtman M. 2006. Sex and virulence in *Escherichia coli*: An evolutionary perspective. Mol Microbiol 60:1136–1151.

21.   Nataro JP, Kaper JB. 1998. Diarrheagenic *Escherichia coli*. Clin Microbiol Rev 11:142–201.

22.   Herschleb J, Ananiev G, Schwartz DC. 2007. Pulsed-field gel electrophoresis. Nat Protoc 2:677–684.

23.   Domingues S, da Silva GJ, Nielsen KM. 2012. Integrons: Vehicles and pathways for horizontal dissemination in bacteria. Mob Genet Elements 2:211–223.

24.   Delannoy S, Mariani-Kurkdjian P, Webb HE, Bonacorsi S, Fach P. 2017. The mobilome; A major contributor to *Escherichia coli* stx2-Positive O26: H11 strains intra-serotype diversity. Front Microbiol 8:1–17.

25.   Casadesús J, Low D. 2006. Epigenetic gene regulation in the bacterial world. Microbiol Mol Biol Rev 70:830–56.

26.   Croxen M, Finlay B. 2010. Molecular mechanisms of *Escherichia coli* pathogenicity. Nat Rev Microbiol 8:26–38.

27.   Shames SR, Auweter SD, Finlay BB. 2009. Co-evolution and exploitation of host cell signaling pathways by bacterial pathogens. Int J Biochem Cell Biol 41:380–389.

28.   Bergan J, Dyve Lingelem AB, Simm R, Skotland T, Sandvig K. 2012. Shiga toxins.

Toxicon 60:1085–1107.

29.   Law D. 2000. The history and evolution of *Escherichia coli* O157 and other Shiga toxin-producing *E. coli*. World J Microbiol Biotechnol 16:701–709.

30.   Imamovic L, Jofre J, Schmidt H, Serra-Moreno R, Muniesa M. 2009. Phage-mediated Shiga toxin 2 gene transfer in food and water. Appl Environ Microbiol 75:1764–8.

31.   Mauro SA, Opalko H, Lindsay K, Colon MP, Koudelka GB. 2013. The microcosm mediates the persistence of Shiga toxin-producing *Escherichia coli* in freshwater ecosystems. Appl Environ Microbiol 79:4821–4828.

32.   Colello R, Etcheverría AI, Di Conza JA, Gutkind GO, Padola NL. 2015. Antibiotic resistance and integrons in Shiga toxin-producing *Escherichia coli* (STEC). Brazilian J Microbiol 46:1–5.

33.   Medina A, Horcajo P, Jurado S, De R, Fuente L, Ruiz-Santa-Quiteria JA, Domínguez-Bernal G, Orden JA. 2011. Phenotypic and genotypic characterization of antimicrobial resistance in enterohemorrhagic *Escherichia coli* and atypical enteropathogenic *E. coli* strains from ruminants. J Vet Diagn Invest 23:91–95.

34.   Johnson KE, Thorpe CM, Sears CL, Thorpe CM, Sears CL. 2006. The emerging clinical importance of non-O157 Shiga toxin-producing *Escherichia coli*. Clin Infect Dis 43:1587–95.

35.   Ray PE, Liu XH. 2001. Pathogenesis of Shiga toxin-induced hemolytic uremic syndrome. Pediatr Nephrol 16:823–839.

36.   Fuller CA, Pellino CA, Flagler MJ, Strasser JE, Weiss AA. 2011. Shiga toxin subtypes display dramatic differences in potency. Infect Immun 79:1329–1337.

37.   Mayer CL, Leibowitz CS, Kurosawa S, Stearns-Kurosawa DJ. 2012. Shiga toxins and the pathophysiology of hemolytic uremic syndrome in humans and animals. Toxins (Basel) 4:1261–87.

38.   Pruimboom-Brees IM, Morgan TW, Ackermann MR, Nystrom ED, Samuel JE, Cornick NA, Moon HW. 2000. Cattle lack vascular receptors for *Escherichia coli*

O157:H7 Shiga toxins. Proc Natl Acad Sci 97:10325–10329.

39.   Grauke LJ, Kudva IT, Yoon JW, Hunt CW, Williams CJ, Hovde CJ. 2002. Gastrointestinal tract location of *Escherichia coli* O157:H7 in ruminants. Appl Environ Microbiol.

40.   Döpfer D, Geue L, Schares S, Mintel B, Hoffmann B, Fischer EAJ. 2012. Dynamics of Shiga-toxin producing *Escherichia coli* (STEC) and their virulence factors in cattle. Prev Vet Med 103:22–30.

41.   Reid SD, Herbelin CJ, Bumbaugh AC, Selander RK, Whittam TS. 2000. Parallel evolution of virulence in pathogenic *Escherichia coli*. Nature 406:64–67.

42.   Lorenz SC, Monday SR, Hoffmann M, Fischer M, Kase JA. 2016. Plasmids from Shiga toxin-producing *Escherichia coli* of rare enterohemolysin (ehxA) subtypes reveal pathogenicity potential and display a novel evolutionary path. Appl Environ Microbiol 82:6367–6377.

43.   McInerney JO, Pisani D, Bapteste E, O'Connell MJ. 2011. The public goods hypothesis for the evolution of life on Earth. Biol Direct 6:41.

44.   Darwin C. 1859. On the origin of species. London: John Murray.

45.   Jaros P, Cookson A, Compbell D, Duncan G, Prattley D, Carter P, Besser T, Shringi S, Hathaway S, Marshall J, French N. 2014. Geographic divergence of bovine and human Shiga toxin-producing *Escherichia coli* O157:H7 genotypes, New Zealand. Emerg Infect Dis 20:1980–1989.

46.   Zhang WL, Bielaszewska M, Liesegang A, Tschäpe H, Schmidt H, Bitzan M, Karch H. 2000. Molecular characteristics and epidemiological significance of Shiga toxin-producing *Escherichia coli* O26 strains. J Clin Microbiol 38:2134–40.

47.   Delannoy S, Mariani-Kurkdjian P, Bonacorsi S, Liguori S, Fach P. 2015. Characteristics of emerging human-pathogenic *Escherichia coli* O26: H11 strains isolated in France between 2010 and 2013 and carrying the *stx* 2d gene only. J Clin Microbiol 53:486–492.

48.   Zweifel C, Cernela N, Stephan R. 2013. Detection of the emerging Shiga toxin-producing *Escherichia coli* O26:H11/H- sequence type 29 (ST29) clone in human

patients and healthy cattle in Switzerland. Appl Environ Microbiol 79:5411–3.

49. Bielaszewska M, Prager R, Köck R, Mellmann A, Zhang W, Tschäpe H, Tarr PI, Karch H. 2007. Shiga toxin gene loss and transfer *in vitro* and *in vivo* during enterohemorrhagic *Escherichia coli* O26 infection in humans. Appl Environ Microbiol 73:3144–50.

50. Ishijima N, Lee K-I, Kuwahara T, Nakayama-Imaohji H, Yoneda S, Iguchi A, Ogura Y, Hayashi T, Ohnishi M, Iyoda S. 2017. Identification of a new virulent clade in enterohemorrhagic *Escherichia coli* O26:H11/H- Sequence Type 29. Sci Rep 7:43136.

51. Rasko DA, Webster DR, Sahl JW, Bashir A, Boisen N, Scheutz F, Paxinos EE, Sebra R, Chin C-S, Iliopoulos D, Klammer A, Peluso P, Lee L, Kislyuk AO, Bullard J, Kasarskis A, Wang S, Eid J, Rank D, Redman JC, Steyert SR, Frimodt-Møller J, Struve C, Petersen AM, Krogfelt KA, Nataro JP, Schadt EE, Waldor MK. 2011. Origins of the *E. coli* strain causing an outbreak of Hemolytic–Uremic Syndrome in Germany. N Engl J Med 365:709–717.

52. USDA. 2014. FSIS Procedure for the use of *Escherichia coli* O157:H7 screening tests for meat products and carcass and environmental sponges. MLB 5A04.

53. USDA. 2014. Detection and isolation of non-O157 Shiga toxin-producing *Escherichia coli* (STEC) from meat products and carcass and environmental sponges. MLG 5B05.

54. MAF. 2011. United States of America: Technical procedures for monitoring Escherichia coli O157:H7 programme in bulk manufacturing bobby veal.

55. Huszczynski G, Gauthier M, Mohajer S, Gill A, Blais B. 2013. Method for the detection of priority Shiga toxin–producing *Escherichia coli* in Beef Trim. J Food Prot 76:1689–1696.

56. Rice DH, Sheng HQ, Wynia SA, Hovde CJ. 2003. Rectoanal mucosal swab culture is more sensitive than fecal culture and distinguishes *Escherichia coli* O157:H7-colonized cattle and those transiently shedding the same organism. J Clin Microbiol 41:4924–9.

57.   Agga GE, Arthur TM, Hinkley S, Bosilevac JM. 2017. Evaluation of rectoanal mucosal swab sampling for molecular detection of enterohemorrhagic *Escherichia coli* in Beef Cattle. J Food Prot 80:661–667.

58.   Kulow MJ, Gonzales TK, Pertzborn KM, Dahm J, Miller BA, Park D, Gautam R, Kaspar CW, Ivanek R, Dopfer D. 2012. Differences in colonization and shedding patterns after oral challenge of cattle with three *Escherichia coli* O157:H7 strains. Appl Environ Microbiol 78:8045–8055.

59.   Widgren S, Eriksson E, Aspan A, Emanuelson U, Alenius S, Lindberg A. 2013. Environmental sampling for evaluating verotoxigenic *Escherichia coli* O157. J Vet Diagnostic Investig 25:189–198.

60.   Eisenberg T, Wolter W, Lenz M, Schlez K, Zschöck M. 2013. Boot swabs to collect environmental samples from common locations in dairy herds for *Mycobacterium avium ssp. paratuberculosis* (MAP) detection. J Dairy Res 80:485–489.

61.   Pearce MC, Jenkins C, Vali L, Smith AW, Knight HI, Cheasty T, Smith HR, Gunn GJ, Woolhouse MEJ, Amyes SGB, Frankel G. 2004. Temporal shedding patterns and virulence factors of *Escherichia coli* serogroups O26, O103, O111, O145, and O157 in a cohort of beef calves and their dams. Appl Environ Microbiol 70:1708–16.

62.   Paton A, Paton J. 1998. Detection and characterization of Shiga toxigenic *Escherichia coli* by using multiplex PCR assays for *stx1, stx2, eaeA,* enterohemorrhagic *E . coli hlyA , rfb O111 , and rfb O157*. J Clin Microbiol 36:598–602.

63.   Paddock Z, Shi X, Bai J, Nagaraja TG. 2012. Applicability of a multiplex PCR to detect O26, O45, O103, O111, O121, O145, and O157 serogroups of *Escherichia coli* in cattle feces. Vet Microbiol 156:381–388.

64.   Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT. 2009. The MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments. Clin Chem 55:611–622.

65. Nielsen EM, Andersen MT. 2003. Detection and characterization of verocytotoxin-producing *Escherichia coli* by automated 5' nuclease PCR assay. J Clin Microbiol 41:2884–2893.

66. Anklam KS, Kanankege KST, Gonzales TK, Kaspar CW, Döpfer D. 2012. Rapid and reliable detection of Shiga toxin–producing *Escherichia coli* by real-time multiplex PCR. J Food Prot 75:643–650.

67. Kanki M, Seto K, Harada T, Yonogi S, Kumeda Y. 2011. Comparison of four enrichment broths for the detection of non-O157 Shiga-toxin-producing *Escherichia coli* O91, O103, O111, O119, O121, O145 and O165 from pure culture and food samples. Lett Appl Microbiol 53:167–173.

68. Vimont A, Delignette-Muller ML, Vernozy-Rozand C. 2007. Supplementation of enrichment broths by novobiocin for detecting Shiga toxin-producing *Escherichia coli* from food: A controversial use. Lett Appl Microbiol 44:326–331.

69. Verstraete K, De Zutter L, Messens W, Herman L, Heyndrickx M, De Reu K. 2010. Effect of the enrichment time and immunomagnetic separation on the detection of Shiga toxin-producing *Escherichia coli* O26, O103, O111, O145 and sorbitol positive O157 from artificially inoculated cattle faeces. Vet Microbiol 145:106–112.

70. Ogden ID, Hepburn NF, MacRae M. 2001. The optimization of isolation media used in immunomagnetic separation methods for the detection of *Escherichia coli* O157 in foods. J Appl Microbiol 91:373–379.

71. Wang F, Yang Q, Kase JA, Meng J, Clotilde LM, Lin A, Ge B. 2013. Current trends in detecting non-O157 Shiga toxin–producing *Escherichia coli* in food. Foodborne Pathog Dis 10:665–677.

72. Ferens WA, Hovde CJ. 2011. *Escherichia coli* O157:H7: animal reservoir and sources of human infection. Foodborne Pathog Dis 8:465–87.

73. Sánchez S, Martínez R, Rey J, García A, Blanco J, Blanco M, Blanco J, Mora A, Herrera-León S, Echeita A, Alonso J. 2010. Pheno-genotypic characterisation of *Escherichia coli* O157:H7 isolates from domestic and wild ruminants. Vet

Microbiol 142:445–449.

74. Hiramatsu R, Matsumoto M, Miwa Y, Suzuki Y, Saito M, Miyazaki Y. 2002. Characterization of Shiga toxin-producing *Escherichia coli* O26 strains and establishment of selective isolation media for these strains. J Clin Microbiol 40:922–925.

75. Possé B, De Zutter L, Heyndrickx M, Herman L. 2008. Novel differential and confirmation plating media for Shiga toxin-producing *Escherichia coli* serotypes O26, O103, O111, O145 and sorbitol-positive and -negative O157. FEMS Microbiol Lett 282:124–131.

76. Possé B, De Zutter L, Heyndrickx M, Herman L. 2008. Quantitative isolation efficiency of O26, O103, O111, O145 and O157 STEC serotypes from artificially contaminated food and cattle faeces samples using a new isolation protocol. J Appl Microbiol 105:227–235.

77. Hirvonen JJ, Siitonen A, Kaukoranta S-S. 2012. Usability and performance of CHROMagar STEC medium in detection of Shiga toxin-producing *Escherichia coli* strains. J Clin Microbiol 50:3586–3590.

78. Gill A, Huszczynski G, Gauthier M, Blais B. 2014. Evaluation of eight agar media for the isolation of Shiga toxin—producing *Escherichia coli*. J Microbiol Methods 96:6–11.

79. Brusa V, Piñeyro PE, Galli L, Linares LH, Ortega EE, Padola NL, Leotta GA. 2016. Isolation of Shiga toxin-producing *Escherichia coli* from ground beef using multiple combinations of enrichment broths and selective agars. Foodborne Pathog Dis 13:163–170.

80. Conrad CC, Stanford K, McAllister TA, Thomas J, Reuter T. 2014. Further development of sample preparation and detection methods for O157 and the top 6 non-O157 STEC serogroups in cattle feces. J Microbiol Methods 105:22–30.

81. Stanford K, Johnson RP, Alexander TW, McAllister TA, Reuter T, Fazil A. 2016. Influence of season and feedlot location on prevalence and virulence factors of seven serogroups of *Escherichia coli* in feces of Western-Canadian slaughter

cattle. PLoS One 11:e0159866.

82. Mellor GE, Fegan N, Duffy LL, Millan KEMC, Jordan D, Barlow RS. 2016. National survey of Shiga toxin–producing *Escherichia coli* Australian beef cattle feces. J Food Prot 79:1868–1874.

83. Irshad H, Cookson A, Hotter G. 2012. Epidemiology of Shiga toxin-producing *Escherichia coli* O157 in very young calves in the North Island of New Zealand. N Z Vet J 60:37–41.

84. Hegde N V., Cote R, Jayarao BM, Muldoon M, Lindpaintner K, Kapur V, DebRoy C. 2012. Detection of the top six non-O157 Shiga toxin–producing *Escherichia coli* O groups by ELISA. Foodborne Pathog Dis 9:1044–1048.

85. Leonard SR, Mammel MK, Lacher DW, Elkins CA. 2015. Application of metagenomic sequencing to food safety: Detection of Shiga toxin-producing *Escherichia coli* on fresh bagged spinach. Appl Environ Microbiol 81:8183–8191.

86. Christner M, Trusch M, Rohde H, Kwiatkowski M, Schlüter H, Wolters M, Aepfelbacher M, Hentschke M. 2014. Rapid MALDI-TOF mass spectrometry strain typing during a large outbreak of Shiga-toxigenic *Escherichia coli*. PLoS One 9:e101924.

87. Novais Â, Sousa C, de Dios Caballero J, Fernandez-Olmos A, Lopes J, Ramos H, Coque TM, Cantón R, Peixe L. 2014. MALDI-TOF mass spectrometry as a tool for the discrimination of high-risk *Escherichia coli* clones from phylogenetic groups B2 (ST131) and D (ST69, ST405, ST393). Eur J Clin Microbiol Infect Dis 33:1391–1399.

88. Norman KN, Strockbine NA, Bono JL. 2012. Association of nucleotide polymorphisms within the O-antigen gene cluster of *Escherichia coli* O26, O45, O103, O111, O121, and O145 with serogroups and genetic subtypes. Appl Environ Microbiol 78:6689–703.

89. Shea S, Kubota KA, Maguire H, Gladbach S, Woron A, Atkinson-Dunn R, Couturier MR, Miller MB. 2017. Clinical microbiology laboratories' adoption of culture-independent diagnostic tests is a threat to foodborne-disease

surveillance in the United States. J Clin Microbiol 55:10–19.

90. Langley G, Besser J, Iwamoto M, Lessa FC, Cronquist A, Skoff TH, Chaves S, Boxrud D, Pinner RW, Harrison LH. 2015. Effect of culture-independent diagnostic tests on future emerging infections program surveillance. Emerg Infect Dis 21:1582–8.

91. Iwamoto M, Huang JY, Cronquist AB, Medus C, Hurd S, Zansky S, Dunn J, Woron AM, Oosmanally N, Griffin PM, Besser J, Henao OL. 2015. Bacterial enteric infections detected by culture-independent diagnostic tests — FoodNet, United States, 2012–2014. Morb Mortal Wkly Rep 252–257.

92. Bielaszewska M, Zhang W, Tarr PI, Sonntag A-K, Karch H. 2005. Molecular profiling and phenotype analysis of *Escherichia coli* O26:H11 and O26:NM: Secular and geographic consistency of enterohemorrhagic and enteropathogenic isolates. J Clin Microbiol 43:4225–4228.

93. Pollock KGJ, Locking ME, Beattie TJ, Maxwell H, Ramage I, Hughes D, Cowieson J, Allison L, Hanson M, Cowden JM. 2010. Sorbitol-fermenting *Escherichia coli* O157, Scotland. Emerg Infect Dis 16:881–2.

94. Wright J, Fraser D, Baker M. 1993. *Escherichia coli* O157:H7 infection: First New Zealand case report. Commun Dis New Zeal Mon Rep 113–116.

95. Hussein HS, Sakuma T. 2005. Prevalence of Shiga toxin-producing *Escherichia coli* in dairy cattle and their products. J Dairy Sci 88:450–465.

96. Gould LH, Mody RK, Ong KL, Clogher P, Cronquist AB, Garman KN, Lathrop S, Medus C, Spina NL, Webb TH, White PL, Wymore K, Gierke RE, Mahon BE, Griffin PM. 2013. Increased recognition of Non-O157 Shiga Toxin–producing *Escherichia coli* infections in the United States During 2000–2010: Epidemiologic features and comparison with *E. coli* O157 infections. Foodborne Pathog Dis 10:453–460.

97. Manning SD, Motiwala AS, Springman AC, Qi W, Lacher DW, Ouellette LM, Mladonicky JM, Somsel P, Rudrik JT, Dietrich SE, Zhang W, Swaminathan B, Alland D, Whittam TS. 2008. Variation in virulence among clades of *Escherichia*

coli O157:H7 associated with disease outbreaks. Proc Natl Acad Sci U S A 105:4868–73.

98. Matussek A, Jernberg C, Einemo I-M, Monecke S, Ehricht R, Engelmann I, Löfgren S, Mernelius S. 2017. Genetic makeup of Shiga toxin-producing *Escherichia coli* in relation to clinical symptoms and duration of shedding: a microarray analysis of isolates from Swedish children. Eur J Clin Microbiol Infect Dis 36:1433–1441.

99. Wong CS, Mooney JC, Brandt JR, Staples AO, Jelacic S, Boster DR, Watkins SL, Tarr PI. 2012. Risk factors for the hemolytic uremic syndrome in children infected with *Escherichia coli* O157:H7: A multivariable analysis. Clin Infect Dis 55:33–41.

100. Gerber A, Karch H, Allerberger F, Verweyen HM, Zimmerhackl LB. 2002. Clinical course and the role of Shiga toxin–producing *Escherichia coli* infection in the Hemolytic-Uremic Syndrome in pediatric patients, 1997–2000, in Germany and Austria: A prospective study. J Infect Dis 186:493–500.

101. Slutsker L, Ries AA, Maloney K, Wells JG, Greene KD, Griffin PM. 1998. A nationwide case-control study of *Escherichia coli* O157:H7 infection in the United States. J Infect Dis 177:962–6.

102. Stephan R, Schumacher S, Corti S, Krause G, Danuser J, Beutin L. 2008. Prevalence and characteristics of Shiga toxin-producing *Escherichia coli* in Swiss raw milk cheeses collected at producer level. J Dairy Sci 91:2561–2565.

103. Elhadidy M, Mohammed MA. 2013. Shiga toxin-producing *Escherichia coli* from raw milk cheese in Egypt: prevalence, molecular characterization and survival to stress conditions. Lett Appl Microbiol 56:120–127.

104. European Food Saftey Authority. 2016. Multi-country outbreak of Shiga toxin-producing *Escherichia coli* infection associated with haemolytic uraemic syndrome. Eur Cent Dis Prev Control EN-1017.

105. Buvens G, Possé B, De Schrijver K, De Zutter L, Lauwers S, Piérard D. 2011. Virulence profiling and quantification of Verocytotoxin-producing *Escherichia*

coli O145:H28 and O26:H11 isolated during an ice cream–related Hemolytic Uremic Syndrome outbreak. Foodborne Pathog Dis 8:421–426.

106. Centers for Disease Control and Prevention. 2016. Multistate outbreak of Shiga toxin-producing *Escherichia coli* infections linked to flour (Final update). https://www.cdc.gov/ecoli/2016/o121-06-16/index.html.

107. Centers for Disease Control and Prevention. 2016. Multistate outbreaks of Shiga toxin-producing *Escherichia coli* infections linked to Chipotle Mexican Grill restaurants (Final update). https://www.cdc.gov/ecoli/2015/o26-11-15/index.html.

108. Jenkins C, Dallman TJ, Launders N, Willis C, Byrne L, Jorgensen F, Eppinger M, Adak GK, Aird H, Elviss N, Grant KA, Morgan D, McLauchlin J. 2015. Public health investigation of two outbreaks of Shiga toxin-producing *Escherichia coli* O157 associated with consumption of watercress. Appl Environ Microbiol 81:3946–52.

109. Buchholz U, Bernard H, Werber D, Böhmer MM, Remschmidt C, Wilking H, Deleré Y, an der Heiden M, Adlhoch C, Dreesman J, Ehlers J, Ethelberg S, Faber M, Frank C, Fricke G, Greiner M, Höhle M, Ivarsson S, Jark U, Kirchner M, Koch J, Krause G, Luber P, Rosner B, Stark K, Kühne M. 2011. German outbreak of *Escherichia coli* O104:H4 associated with sprouts. N Engl J Med 365:1763–1770.

110. Angulo FJ, Steinmuller N, Demma L, Bender JB, Eidson M, Angulo FJ. 2006. Outbreaks of enteric disease associated with animal contact: Not just a foodborne problem anymore. Clin Infect Dis 43:1596–1602.

111. O'Brien SJ, Adak GK, Gilham C. 2001. Contact with farming environment as a major risk factor for Shiga toxin (Vero cytotoxin)-producing *Escherichia coli* O157 infection in humans. Emerg Infect Dis 7:1049–1051.

112. Smith KE, Stenzel SA, Bender JB, Wagstrom E, Soderlund D, Leano FT, Taylor CM, Belle-Isle PA, Danila R. 2004. Outbreaks of enteric infections caused by multiple pathogens associated with calves at a farm day camp. Pediatr Infect Dis J 23:1098–1104.

113. Haus-Cheymol R, Espie E, Che D, Vaillant V, De Valk H, Desenclos JC. 2006. Association between indicators of cattle density and incidence of paediatric haemolytic – uraemic syndrome (HUS) in children under 15 years of age in France between 1996 and 2001: an ecological study. Epidemiol Infect 134:712.

114. Crump JA, Sulka AC, Langer AJ, Schaben C, Crielly AS, Gage R, Baysinger M, Moll M, Withers G, Toney DM, Hunter SB, Hoekstra RM, Wong SK, Griffin PM, Van Gilder TJ. 2002. An outbreak of *Escherichia coli* O157:H7 infections among visitors to a dairy farm. N Engl J Med 347:555–560.

115. Wilson JB, Clarke RC, Renwick SA, Rahn V, Johnson RP, Karmali MA, Lior H, Alves D, Gyles CL, Sandhu KS, McEwen SA, Spika JS, Woolhouse MEJ, Reid SWJ. 1996. Vero cytotoxigenic *Escherichia coli* infection in dairy farm families. J Infect Dis 174:1021–1027.

116. Óhaiseadha C, Hynds PD, Fallon UB, O'Dwyer J. 2017. A geostatistical investigation of agricultural and infrastructural risk factors associated with primary verotoxigenic *E. coli* (VTEC) infection in the Republic of Ireland, 2008–2013. Epidemiol Infect 145:95–105.

117. Mead PS, Griffin PM. 1998. *Escherichia coli* O157:H7. Lancet 352:1207–12.

118. Clogher P, Hurd S, Hoefer D, Hadler JL, Pasutti L, Cosgrove S, Segler S, Tobin-D'angelo M, Nicholson C, Booth H, Garman K, Mody RK, Gould LH. 2012. Assessment of physician knowledge and practices concerning Shiga toxin–producing *Escherichia coli* infection and enteric illness, 2009, foodborne diseases active surveillance network (FoodNet). Clin Infect Dis 54:S446-52.

119. Scallan E, Mahon BE. 2012. Foodborne diseases active surveillance network (FoodNet) in 2012: a foundation for food safety in the United States. Clin Infect Dis 54:S381-4.

120. Gow SP, Waldner CL. 2009. Antimicrobial resistance and virulence factors *stx1*, *stx2*, and *eae* in generic *Escherichia coli* isolates from calves in Western Canadian cow-calf herds. Microb Drug Resist 15:61–67.

121. Cho S, Fossler CP, Diez-Gonzalez F, Wells SJ, Hedberg CW, Kaneene JB, Ruegg

PL, Warnick LD, Bender JB. 2007. Antimicrobial susceptibility of Shiga toxin-producing *Escherichia coli* isolated from organic dairy farms, conventional dairy farms, and county fairs in Minnesota. Foodborne Pathog Dis 4:178–186.

122. Buvens G, Bogaerts P, Glupczynski Y, Lauwers S, Piérard D. 2010. Antimicrobial resistance testing of verocytotoxin-producing *Escherichia coli* and first description of TEM-52 extended-spectrum β-lactamase in serogroup O26. Antimicrob Agents Chemother 54:4907–9.

123. Food and Agriculture Organization of the United Nations. 2014. Statistical yearbook of the Food and Agricultural Organization for the United Nations. http://www.fao.org/docrep/017/i3138e/i3138e07.pdf.

124. Islam MZ, Musekiwa A, Islam K, Ahmed S, Chowdhury S, Ahad A, Biswas PK. 2014. Regional variation in the prevalence of *E. coli* O157 in cattle: A meta-analysis and meta-regression. PLoS One 9:e93299.

125. Lupindu AM. 2018. Epidemiology of Shiga toxin-producing *Escherichia coli* O157:H7 in Africa. South African J Infect Dis 33:24–30.

126. Cookson A, Taylor S, Bennett J, Thomson-Carter F, Attwood G. 2006. Serotypes and analysis of distribution of Shiga toxin-producing *Escherichia coli* from cattle and sheep in the lower North Island, New Zealand. N Z Vet J 54:78–84.

127. Jaros P, Cookson AL, Reynolds A, Prattley DJ, Campbell DM, Hathaway S, French NP. 2016. Nationwide prevalence and risk factors for faecal carriage of *Escherichia coli* O157 and O26 in very young calves and adult cattle at slaughter in New Zealand. Epidemiol Infect 144:1736–1747.

128. Kolenda R, Burdukiewicz M, Schierack P. 2015. A systematic review and meta-analysis of the epidemiology of pathogenic *Escherichia coli* of calves and the role of calves as reservoirs for human pathogenic *E. coli*. Front Cell Infect Microbiol 5:23.

129. Shahrani M, Dehkordi FS, Momtaz H. 2014. Characterization of *Escherichia coli* virulence genes, pathotypes and antibiotic resistance properties in diarrheic calves in Iran. Biol Res 47:28.

130. Cobbold R, Desmarchelier P. 2000. A longitudinal study of Shiga-toxigenic *Escherichia coli* (STEC) prevalence in three Australian dairy herds. Vet Microbiol 71:125–137.

131. Breum SØ, Boel J. 2010. Prevalence of *Escherichia coli* O157 and verocytotoxin producing *E. coli* (VTEC) on Danish beef carcasses. Int J Food Microbiol 141:90–96.

132. Cho S, Bender JB, Diez-Gonzalez F, Fossler CP, Hedberg CW, Kaneene JB, Ruegg PL, Warnick LD, Wells SJ. 2006. Prevalence and characterization of *Escherichia coli O157* isolates from Minnesota dairy farms and county fairs. J Food Prot 69:252–259.

133. Ferreira MRA, Filho EGF, Pinto JFN, Dias M, Moreira CN. 2014. Isolation, prevalence, and risk factors for infection by Shiga toxin-producing *Escherichia coli* (STEC) in dairy cattle. Trop Anim Health Prod 46:635–639.

134. Heuvelink AE, van den Biggelaar FL, de Boer E, Herbes RG, Melchers WJ, Huis in 't Veld JH, Monnens LA. 1998. Isolation and characterization of verocytotoxin-producing *Escherichia coli* O157 strains from Dutch cattle and sheep. J Clin Microbiol 36:878–82.

135. Fernández D, Sanz ME, Parma AE, Padola NL, Waltner-Toews D, Gyles CL, Fiorilli G, Sosa-Estani S, Kincaid J, Rangel J, Griffin PM, Lucchesi P, Parma AE, Blanco J. 2012. Characterization of Shiga toxin-producing *Escherichia coli* isolated from newborn, milk-fed, and growing calves in Argentina. J Dairy Sci 95:5340–3.

136. Cho S, Fossler CP, Diez-Gonzalez F, Wells SJ, Hedberg CW, Kaneene JB, Ruegg PL, Warnick LD, Bender JB. 2009. Cattle-level risk factors associated with fecal shedding of Shiga toxin-encoding bacteria on dairy farms, Minnesota, USA. Can J Vet Res 73:151–6.

137. Cobbaut K, Berkvens D, Houf K, De Deken R, De Zutter AL. 2009. *Escherichia coli* O157 prevalence in different cattle farm types and identification of potential risk factors. J Food Prot 72:1848–1853.

138. Wilson JB, Renwick SA, Clarke RC, Rahn K, Alves D, Johnson RP, Ellis AG, McEwen SA, Karmali MA, Lior H, Spika J, Futter RJ, Kay ACS, Willshaw GA, Cheasty T. 1998. Risk factors for infection with verocytotoxigenic *Escherichia coli* in cattle on Ontario dairy farms. Prev Vet Med 34:227–236.

139. Nielsen EM, Tegtmeier C, Andersen HJ, Grønbaek C, Andersen JS, Nielsen EM. 2002. Influence of age, sex and herd characteristics on the occurrence of verocytotoxin-producing *Escherichia coli* O157 in Danish dairy farms. Vet Microbiol 88:245–257.

140. Cho S, Fossler CP, Diez-Gonzalez F, Wells SJ, Hedberg CW, Kaneene JB, Ruegg PL, Warnick LD, Bender JB. 2013. Herd-level risk factors associated with fecal shedding of Shiga toxin-encoding bacteria on dairy farms in Minnesota, USA. Can Vet J = La Rev Vet Can 54:693–7.

141. Ellis-Iversen J, Cook AJC, Smith RP, Pritchard GC, Nielen M. 2009. Temporal patterns and risk factors for *Escherichia coli* O157 and *Campylobacter spp.* in young cattle. J Food Prot 72:490–496.

142. Gunn GJ, McKendrick IJ, Ternent HE, Thomson-Carter F, Foster G, Synge BA. 2007. An investigation of factors associated with the prevalence of verocytotoxin producing *Escherichia coli* O157 shedding in Scottish beef cattle. Vet J 174:554–564.

143. Kuhnert P, Dubosson CR, Roesch M, Homfeld E, Doherr MG, Blum JW. 2005. Prevalence and risk-factor analysis of Shiga toxigenic *Escherichia coli* in faecal samples of organically and conventionally farmed dairy cattle. Vet Microbiol 109:37–45.

144. LeJeune JT, Besser TE, Rice DH, Berg JL, Stilborn RP, Hancock DD. 2004. Longitudinal study of fecal shedding of *Escherichia coli* O157:H7 in feedlot cattle: Predominance and persistence of specific clonal types despite massive cattle population turnover. Appl Environ Microbiol 70:377–384.

145. Zhao L, Tyler PJ, Starnes J, Rankins D, McCaskey TA, Wang L. 2014. Evaluation of the effects of weaning diets on *Escherichia coli* O157 shedding, body weight, and fecal bacterial communities in beef calves. Foodborne Pathog Dis 11:55–60.

146. Venegas-Vargas C, Henderson S, Khare A, Mosci RE, Lehnert JD, Singh P, Ouellette L, Norby B, Funk JA, Rust S, Bartlett P, Grooms D, Manning SD. 2016. Factors associated with Shiga toxin-producing *Escherichia coli* shedding in dairy and beef cattle. Appl Environ Microbiol AEM.00829-16.

147. Edrington TS, Carter BH, Farrow RL, Islas A, Hagevoort GR, Friend TH, Callaway TR, Anderson RC, Nisbet DJ. 2011. Influence of weaning on fecal shedding of pathogenic bacteria in dairy calves. Foodborne Pathog Dis 8:395–401.

148. Hancock DD, Besser TE, Rice DH, Herriott DE, Tarr PI. 1997. A longitudinal study of *Escherichia coli* O157 in fourteen cattle herds. Epidemiol Infect 118:193–5.

149. Conedera G, Chapman PA, Marangon S, Tisato E, Dalvit P, Zuin A. 2001. A field survey of *Escherichia coli* O157 ecology on a cattle farm in Italy. Int J Food Microbiol 66:85–93.

150. Mir RA, Weppelmann TA, Elzo M, Ahn S, Driver JD, Jeong KC. 2016. Colonization of beef cattle by Shiga toxin-producing *Escherichia coli* during the first year of life: A cohort study. PLoS One 11:e0148518.

151. Polifroni R, Etcheverría AI, Sanz ME, Cepeda RE, Krüger A, Lucchesi PMA, Fernández D, Parma AE, Padola NL. 2012. Molecular characterization of Shiga toxin-producing *Escherichia coli* isolated from the environment of a dairy farm. Curr Microbiol 65:337–343.

152. Fremaux B, Raynaud S, Beutin L, Rozand CV. 2006. Dissemination and persistence of Shiga toxin-producing *Escherichia coli* (STEC) strains on French dairy farms. Vet Microbiol 117:180–191.

153. Wetzel AN, LeJeune JT. 2006. Clonal dissemination of Escherichia coli O157:H7 subtypes among dairy farms in Northeast Ohio. Appl Environ Microbiol 72:2621–2626.

154. Renter DG, Morris JG, Sargeant JM, Hungerford LL, Berezowski J, Ngo T, Williams K, Acheson DWK. 2005. Prevalence, risk factors, O serogroups, and

virulence profiles of Shiga toxin–producing bacteria from cattle production environments. J Food Prot 68:1556–1565.

155. Lambertini E, Karns JS, Van Kessel JAS, Cao H, Schukken YH, Wolfgang DR, Smith JM, Pradhan AK. 2015. Dynamics of *Escherichia coli* virulence factors in dairy herds and farm environments in a longitudinal study in the United States. Appl Environ Microbiol 81:4477–4488.

156. Shere JA, Bartlett KJ, Kaspar CW. 1998. Longitudinal study of *Escherichia coli* O157:H7 dissemination on four dairy farms in Wisconsin. Appl Environ Microbiol 64:1390–9.

157. Liu W-C, Shaw DJ, Matthews L, Hoyle D V, Pearce MC, Yates CM, Low JC, Amyes SGB, Gunn GJ, Woolhouse MEJ. 2007. Modelling the epidemiology and transmission of Verocytotoxin-producing *Escherichia coli* serogroups O26 and O103 in two different calf cohorts. Epidemiol Infect 135:1316–23.

158. Liu WC, Jenkins C, Shaw DJ, Matthews L, Pearce MC, Low JC, Gunn GJ, Smith HR, Frankel G, Woolhouse MEJ. 2005. Modelling the epidemiology of Verocytotoxin-producing *Escherichia coli* serogroups in young calves. Epidemiol Infect 133:449–58.

159. Gautam R, Kulow M, Park D, Gonzales TK, Dahm J, Shiroda M, Stasic AJ, Dopfer D, Kaspar CW, Ivanek R. 2015. Transmission of *Escherichia coli* O157:H7 in cattle is influenced by the level of environmental contamination. Epidemiol Infect 143:274–287.

160. Arthur TM, Keen JE, Bosilevac JM, Brichta-Harhay DM, Kalchayanand N, Shackelford SD, Wheeler TL, Nou X, Koohmaraie M. 2009. Longitudinal study of *Escherichia coli* O157:H7 in a beef cattle feedlot and role of high-level shedders in hide contamination. Appl Environ Microbiol 75:6515–6523.

161. Widiasih DA, Ido N, Omoe K, Sugii S, Shinagawa K. 2004. Duration and magnitude of faecal shedding of Shiga toxin-producing *Escherichia coli* from naturally infected cattle. Epidemiol Infect 132:67–75.

162. Gannon VPJ, Graham TA, King R, Michel P, Read S, Ziebell K, Johnson RP. 2002.

Escherichia coli O157:H7 infection in cows and calves in a beef cattle herd in Alberta, Canada. Epidemiol Infect 129:163–72.

163. Cray WC, Moon HW. 1995. Experimental infection of calves and adult cattle with *Escherichia coli* O157:H7. Appl Environ Microbiol 61:1586–1590.

164. Polifroni R, Etcheverría AI, Arroyo GH, Padola NL. 2014. Survival of VTEC O157 and non-O157 in water troughs and bovine feces. Rev Argent Microbiol 46:126–132.

165. Trevisani M, Mancusi R, Delle Donne G, Bacci C, Bassi L, Bonardi S. 2014. Detection of Shiga toxin (Stx)-producing *Escherichia coli* (STEC) in bovine dairy herds in Northern Italy. Int J Food Microbiol 184:45–49.

166. Cobbold R, Desmarchelier P. 2001. Characterisation and clonal relationships of Shiga-toxigenic *Escherichia coli* (STEC) isolated from Australian dairy cattle. Vet Microbiol 79:323–335.

167. Ahmad A, Nagaraja TG, Zurek L. 2007. Transmission of *Escherichia coli* O157:H7 to cattle by house flies. Prev Vet Med 80:74–81.

168. Kauffman MD, LeJeune J. 2011. European Starlings (*Sturnus vulgaris*) challenged with *Escherichia coli* O157 can carry and transmit the human pathogen to cattle. Lett Appl Microbiol 53:596–601.

169. Baltasar P, Milton S, Swecker, Jr. W, Elvinger F ˛ O, Ponder M. 2014. Shiga toxin–producing *Escherichia coli* distribution and characterization in a pasture-based cow-calf production system. J Food Prot 77:722–731.

170. Jonsson ME, Aspan A, Eriksson E, Vagsholm I. 2001. Persistence of verocytotoxin-producing *Escherichia coli* O157:H7 in calves kept on pasture and in calves kept indoors during the summer months in a Swedish dairy herd. Int J Food Microbiol 66:55–61.

171. Laegreid WW, Keen JE. 2004. Estimation of the basic reproduction ratio (R_0) for Shiga toxin-producing *Escherichia coli* O157:H7 (STEC O157) in beef calves. Epidemiol Infect 132:291–5.

172. Dunn JR, Keen JE, Vecchio R Del, Wittum TE, Thompson ARA. 2004.

Escherichia coli O157:H7 in a cohort of weaned, preconditioned range beef calves. J Food Prot 67:2391–2396.

173. Besser TE, Richards BL, Rice DH, Hancock DD. 2001. *Escherichia coli* O157:H7 infection of calves: Infectious dose and direct contact transmission. Epidemiol Infect 127:555–60.

174. Schouten J, Graat E, Frankena K, Van Zijderveld F, De Jong M. 2009. Transmission and quantification of verocytotoxin-producing *Escherichia coli* O157 in dairy cattle and calves. Epidemiol Infect 137:114–123.

175. Cobbold R, Desmarchelier P. 2002. Horizontal transmission of Shiga toxin-producing *Escherichia coli* within groups of dairy calves. Appl Environ Microbiol 68:4148–52.

176. Worley JN, Flores KA, Yang X, Chase JA, Cao G, Tang S, Meng J, Atwill ER. 2017. Prevalence and genomic characterization of *Escherichia coli* O157:H7 in cow-calf herds throughout California. Appl Environ Microbiol 83:e00734-17.

177. Irshad H, Cookson a. L, Prattley DJ, Dufour M, French NP. 2014. Distribution of *Escherichia coli* strains harbouring Shiga toxin-producing *E. coli* (STEC)-associated virulence factors (*stx1, stx2, eae, ehxA*) from very young calves in the North Island of New Zealand. Epidemiol Infect 142:2548–2558.

178. Mechie SC, Chapman PA, Siddons CA. 1997. A fifteen month study of *Escherichia coli* O157:H7 in a dairy herd. Epidemiol Infect 118:17–25.

179. Chase-Topping ME, McKendrick IJ, Pearce MC, MacDonald P, Matthews L, Halliday J, Allison L, Fenlon D, Low JC, Gunn G, Woolhouse MEJ. 2007. Risk factors for the presence of high-level shedders of *Escherichia coli* O157 on Scottish farms. J Clin Microbiol 45:1594–1603.

180. Chase-Topping M, Gally D, Low C, Matthews L, Woolhouse M. 2008. Super-shedding and the link between human infection and livestock carriage of *Escherichia coli* O157. Nat Rev Microbiol 6:904–12.

181. Williams KJ, Ward MP, Dhungyel OP. 2014. Daily variations in *Escherichia coli* O157 shedding patterns in a cohort of dairy heifers at pasture. Epidemiol Infect

1–10.

182. Robinson SE, Wright EJ, Hart CA, Bennett M, French NP. 2004. Intermittent and persistent shedding of *Escherichia coli* O157 in cohorts of naturally infected calves. J Appl Microbiol 97:1045–1053.

183. Murphy BP, McCabe E, Murphy M, Buckley JF, Crowley D, Fanning S, Duffy G. 2016. Longitudinal study of two Irish dairy herds: Low numbers of Shiga toxin-producing *Escherichia coli* O157 and O26 super-shedders identified. Front Microbiol 7:1850.

184. Jaros P, Cookson A, Reynolds A, Withers H, Clemens R, Brightwell G, Mills J, Marshall J, Prattley D, Campbell D, Hathaway S, French N. 2018. The effect of transportation and lairage on faecal shedding and carcass contamination with Escherichia coli O157 and O26 in very young calves in New Zealand. Epidemiol Infect In Review.

185. Arthur TM, Ahmed R, Chase-Topping M, Kalchayanand N, Schmidt JW, Bono JL. 2013. Characterization of *Escherichia coli* O157:H7 strains isolated from supershedding cattle. Appl Environ Microbiol 79:4294–4303.

186. Spencer SEF, Besser TE, Cobbold RN, French NP. 2015. 'Super" or just "above average"? Supershedders and the transmission of *Escherichia coli* O157:H7 among feedlot cattle.'" J R Soc Interface 12:20150446.

187. Robinson SE, Brown PE, Wright EJ, Hart CA, French NP. 2009. Quantifying within- and between-animal variation and uncertainty associated with counts of *Escherichia coli* O157 occurring in naturally infected cattle faeces. J R Soc Interface.

188. Widiasih DA, Matsuda I, Omoe K, Hu D-L, Sugii S, Shinagawa K. 2004. Passive transfer of antibodies to Shiga toxin-producing *Escherichia coli* O26, O111 and O157 antigens in neonatal calves by feeding colostrum. J Vet Med Sci 66:213–215.

189. Rabinovitz BC, Gerhardt E, Tironi Farinati C, Abdala A, Galarza R, Vilte DA, Ibarra C, Cataldi A, Mercado EC, Cataldi A, Mercado EC, Finlay BB. 2012. Vaccination of pregnant cows with EspA, EspB, γ-intimin, and Shiga toxin 2

proteins from *Escherichia coli* O157:H7 induces high levels of specific colostral antibodies that are transferred to newborn calves. J Dairy Sci 95:3318–26.

190. Vilte DA, Larzábal M, Mayr UB, Garbaccio S, Gammella M, Rabinovitz BC, Delgado F, Meikle V, Cantet RJC, Lubitz P, Lubitz W, Cataldi A, Mercado EC. 2012. A systemic vaccine based on *Escherichia coli* O157:H7 bacterial ghosts (BGs) reduces the excretion of *E. coli* O157:H7 in calves. Vet Immunol Immunopathol 146:169–176.

191. Wileman BW, Thomson DU, Olson KC, Jaeger JR, Pacheco LA, Bolte J, Burkhardt DT, Emery DA, Straub D. 2011. *Escherichia coli* O157:H7 shedding in vaccinated beef calves born to cows vaccinated prepartum with *Escherichia coli* O157:H7 SRP vaccine. J Food Prot 74:1599–1604.

192. McNeilly TN, Mitchell MC, Corbishley A, Nath M, Simmonds H, McAteer SP, Mahajan A, Low JC, Smith DGE, Huntley JF, Gally DL. 2015. Optimizing the protection of cattle against *Escherichia coli* O157:H7 colonization through immunization with different combinations of H7 Flagellin, Tir, Intimin-531 or EspA. PLoS One 10:e0128391.

193. Martorelli L, Garbaccio S, Vilte DA, Albanese AA, Mejías MP, Palermo MS, Mercado EC, Ibarra CE, Cataldi AA. 2017. Immune response in calves vaccinated with type three secretion system antigens and Shiga toxin 2B subunit of *Escherichia coli* O157:H7. PLoS One 12:e0169422.

194. Laegreid WW, Elder RO, Keen JE. 1999. Prevalence of *Escherichia coli* O157:H7 in range beef calves at weaning. Epidemiol Infect 123:291–8.

195. Sreerama S, Sanderson MW, Wilkerson M, Nagaraja TG. 2008. Impact of dexamethasone-induced immunosuppression on the duration and level of shedding of Escherichia coli O157:H7 in calves. Curr Microbiol 56:651–655.

196. Schouten JM, Graat EAM, Frankena K, Van De Giessen AW, Van Der Zwaluw WK, De Jong MCM. 2005. A longitudinal study of *Escherichia coli* O157 in cattle of a Dutch dairy farm and in the farm environment. Vet Microbiol 107:193–204.

197. Alali WQ, Sargeant JM, Nagaraja TG, Debey BM. 2004. Effect of antibiotics in

milk replacer on fecal shedding of *Escherichia coli* O157:H7 in calves 1. J Anim Sci 82:2148–2152.

198. Baines D, Erb S. 2013. Characterization of Shiga toxin-producing *Escherichia coli* infections in beef feeder calves and the effectiveness of a prebiotic in alleviating Shiga toxin-producing *Escherichia coli* infections. Ir Vet J 66:17.

199. Zhao T, Tkalcic S, Doyle MP, Harmon BG, Brown CA, Zhao P. 2003. Pathogenicity of enterohemorrhagic *Escherichia coli* in neonatal calves and evaluation of fecal shedding by treatment with probiotic *Escherichia coli*. J Food Prot 66:924–930.

200. Masana MO, Leotta GA, Castillo LL Del, D 'astek BA, Palladino PM, Galli L, Vilacoba E, Carbonari C, Rodr Guez HR, Rivas AM. 2010. Prevalence, characterization, and genotypic analysis of *Escherichia coli* O157:H7/NM from selected beef exporting abattoirs of Argentina. J Food Prot 73:649–656.

201. Stromberg ZR, Lewis GL, Aly SS, Lehenbauer TW, Bosilevac JM, Cernicchiaro N, Moxley RA. 2016. Prevalence and level of enterohemorrhagic *Escherichia coli* in culled dairy cows at harvest. J Food Prot 79:421–431.

202. Bosilovac JM, Wang R, Luedtke BE, Hinkley S, Wheeler TL, Koohmaraie M. 2017. Characterization of enterohemorrhagic *Escherichia coli* on veal hides and carcasses. J Food Prot 80:136–145.

203. Ayaz ND, Gencay YE, Erol I. 2014. Prevalence and molecular characterization of sorbitol fermenting and non-fermenting *Escherichia coli* O157:H7+/H7- isolated from cattle at slaughterhouse and slaughterhouse wastewater. Int J Food Microbiol 174:31–38.

204. Alonso S, Mora A, Blanco M, Blanco JE, Dahbi G, Ferreiro MT, López C, Alberghini L, Albonetti S, Echeita A, Trevisani M, Blanco J. 2007. Fecal carriage of *Escherichia coli O157:H7* and carcass contamination in cattle at slaughter in northern Italy. Int Microbiol 10:109–116.

205. Thomas KM, McCann MS, Collery MM, Logan A, Whyte P, McDowell DA, Duffy G. 2012. Tracking verocytotoxigenic *Escherichia coli* O157, O26, O111, O103 and

O145 in Irish cattle. Int J Food Microbiol 153:288–296.

206. Arthur TM, Bosilevac JM, Nou X, Shackelford SD, Wheeler TL, Kent MP, Jaroni D, Pauling B, Allen DM, Koohmaraie M. 2004. *Escherichia coli* O157 prevalence and enumeration of aerobic bacteria, enterobacteriaceae, and *Escherichia coli* O157 at various steps in commercial beef processing plants. J Food Prot 67:658–665.

207. Vogeleer P, Tremblay YDN, Mafu AA, Jacques M, Harel J. 2014. Life on the outside: role of biofilms in environmental persistence of Shiga-toxin producing *Escherichia coli*. Front Microbiol 5:317.

208. Arthur TM, Bosilevac JM, Brichta-Harhay DM, Guerini MN, Kalchayanand N, Shackelford SD, Wheeler TL, Koohmaraie M. 2007. Transportation and lairage environment effects on prevalence, numbers, and diversity of *Escherichia coli* O157:H7 on hides and carcasses of beef cattle at processing. J Food Prot 70:280–286.

209. Heuvelink AE, Roessink GL, Bosboom K, De Boer E, Heuvelink AE. 2001. Zero-tolerance for faecal contamination of carcasses as a tool in the control of O157 VTEC infections. Int J Food Microbiol 66:13–20.

210. The Institute of Environmental Science and Research Ltd. 2016. New Zealand public health surveillance report: December 2016. https://surv.esr.cri.nz/PDF_surveillance/NZPHSR/2016/NZPHSRDec2016.pdf.

211. Vugia DA, Cronquist J, Hadler M, Tobin-D'Angelo D, Blythe K, Smith S, Lathrop D, Morse P, Cieslak T, Jones KG, Holt JJ, Guzewich OL, Henao, Scallan E, Angulo F, Griffin P, Tauxe R. 2007. Preliminary FoodNet data on the incidence of infection with pathogens transmitted commonly through food - 10 States, 2006. Morb Mortal Wkly Rep 56:336–339.

212. Meat Industry Assocation. 2016 Annual Report. http://www.mia.co.nz/docs/annual_reports/mia_ar_2016.pdf.

213. Mellor GE, Fegan N, Duffy LL, Mcmillan KE, Jordan D, Barlow RS. 2016. National survey of Shiga toxin–producing *Escherichia coli* serotypes O26, O45, O103, O111,

O121, O145, and O157 in Australian beef cattle feces. J Food Prot 79:1868–1874.

214. Wang R, Koohmaraie M, Luedtke BE, Wheeler TL, Bosilevac JM. 2014. Effects of in-plant interventions on reduction of enterohemorrhagic *Escherichia coli* and background indicator microorganisms on veal calf hides. J Food Prot 77:745–751.

215. DairyNZ Limited. 2013. New Zealand Dairy Statistics 2012-2013. http//www.lic.co.nz/user/file/DAIRY Stat 2012-13-WEB.pdf 1–52.

216. Hosking E, Petrik D. 2013. STEC detection and identification: NeoSEEK[TM] approach to STEC detection and identification. Neogen Corp http//foodsafety.neogen.com/pdf/whitepapers/ecolistecidentification_0313.pdf.

217. Collins LM, Lanza ST. 2010. Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences. John Wiley & Sons, Inc.

218. Seemann T, Goncalves da Silva A, Bulach D, Schultz M, Kwong J, Howden B. 2017. Nullarbor. Github. https://github.com/tseemann/nullarbor.

219. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

220. Thomsen MCF, Ahrenfeldt J, Cisneros JLB, Jurtz V, Larsen MV, Hasman H, Aarestrup FM, Lund O. 2016. A bacterial analysis platform: An integrated system for analysing bacterial whole genome sequencing data for clinical diagnostics and surveillance. PLoS One 11:e0157718.

221. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. Nucleic Acids Res 44:W242-5.

222. R Core Team. 2015. R: A language and environment for statistical computing. 3.2.1. R Foundation for Statistical Computing, Vienna, Austria.

223. Liaw A, Wiener M. 2002. Classification and regression by randomForest. R news 2:18–22.

224. Lesnoff M, Lancelot R. 2011. aod: Analysis of overdispersed data. R package version 1.3.

225. Williams KJ, Ward MP, Dhungyel OP, Hall EJS. 2014. Risk factors for *Escherichia coli* O157 shedding and super-shedding by dairy heifers at pasture. Epidemiol Infect 143:1004–1015.

226. Mateparae J. 2016. Animal welfare (Calves) Regulations 2016. http://legislation.govt.nz/regulation/public/2016/0170/latest/whole.html 1–17.

227. Thomas DE, Elliott EJ. 2013. Interventions for preventing diarrhea-associated hemolytic uremic syndrome: Systematic review. BMC Public Health 13:799.

228. Food Standards Agency. 2002. Red Meat Safety & Clean Livestock. https://www.food.gov.uk/sites/default/files/multimedia/pdfs/publication/redmeatsafety.pdf.

229. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.

230. Fox JT, Renter DG, Sanderson MW, Nutsch AL, Shi X, Nagaraja TG. 2008. Associations between the presence and magnitude of *Escherichia coli* O157 in feces at harvest and contamination of preintervention beef carcasses. J Food Prot 71:1761–7.

231. National Center for Emerging and Zoonotic Infectious Diseases. 2014. National Enteric Disease Surveillance: Shiga toxin-producing *Escherichia coli* (STEC) annual report, 2014. https://www.cdc.gov/nationalsurveillance/pdfs/STEC-2014-REPORT_508c.pdf.

232. Severi E, Vial F, Peron E, Mardh O, Niskanen T, Takkinen J. 2016. Community-wide outbreaks of haemolytic uraemic syndrome associated with Shiga toxin-producing *Escherichia coli* O26 in Italy and Romania: a new challenge for the European Union. Euro Surveill 21:30420.

233. Ronholm J, Nasheri N, Petronella N, Pagotto F. 2016. Navigating microbiological food safety in the era of whole-genome sequencing. Clin Microbiol Rev 29:837–857.

234. U.S. Food & Drug Administration. Whole Genome Sequencing (WGS) Program -

GenomeTrakr Network.

https://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingPro

gramWGS/ucm363134.htm.

235. Wood DE, Salzberg SL. 2014. Kraken: Ultrafast metagenomic sequence

classification using exact alignments. Genome Biol 15:R46.

236. Pedersen T. 2016. FindMyFriends: Microbial comparative genomics in R. R

Package version 180, https//github.com/thomasp85/FindMyFriends.

237. Li W, Godzik A. 2006. Cd-hit: a fast program for clustering and comparing large

sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659.

238. Pedersen TL. 2017. Hierarchical sets: Analyzing pangenome structure through

scalable set visualizations. Bioinformatics 33:1604–1612.

239. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, von Mering C,

Bork P. 2017. Fast Genome-Wide Functional Annotation through Orthology

Assignment by eggNOG-Mapper. Mol Biol Evol 34:2115–2122.

240. Wright E. 2016. Using DECIPHER v2.0 to Analyze Big Biological Sequence Data

in R. R J 8:352–359.

241. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, Parkhill J,

Harris SR. 2015. Rapid phylogenetic analysis of large samples of recombinant

bacterial whole genome sequences using Gubbins. Nucleic Acids Res 43:e15–e15.

242. Bouckaert R, Heled J, Kühnert D, Vaughan T, Wu C-H, Xie D, Suchard MA,

Rambaut A, Drummond AJ. 2014. BEAST 2: A software platform for Bayesian

evolutionary analysis. PLoS Comput Biol 10:e1003537.

243. Baele G, Lemey P, Bedford T, Rambaut A, Suchard MA, Alekseyenko A V. 2012.

Improving the accuracy of demographic and molecular clock model comparison

while accommodating phylogenetic uncertainty. Mol Biol Evol 29:2157–2167.

244. Drummond AJ, Ho SYW, Phillips MJ, Rambaut A. 2006. Relaxed phylogenetics

and dating with confidence. PLoS Biol 4:e88.

245. Binney B, Biggs P, Carter P, Holland B, French N. 2014. Quantification of

historical livestock importation into New Zealand 1860–1979. N Z Vet J 626:309–

314.

246.   Food and Agriculture Organization of the United Nations. 2017. FAOSTAT.
        http://faostat3.fao.org/download/Q/QA/E.

247.   Strachan NJC, Rotariu O, Lopes B, MacRae M, Fairley S, Laing C, Gannon V,
        Allison LJ, Hanson MF, Dallman T, Ashton P, Franz E, van Hoek AHAM, French
        NP, George T, Biggs PJ, Forbes KJ. 2015. Whole genome sequencing
        demonstrates that geographic variation of *Escherichia coli* O157 genotypes
        dominates host association. Sci Rep 5:14145.

248.   van Diemen PM, Dziva F, Stevens MP, Wallis TS. 2005. Identification of
        enterohemorrhagic *Escherichia coli* O26:H- genes required for intestinal
        colonization in calves. Infect Immun 73:1735–43.

249.   Etcheverría AI, Padola NL. 2013. Shiga toxin-producing *Escherichia coli*: factors
        involved in virulence and cattle colonization. Virulence 4:366–72.

250.   Ison SA, Delannoy S, Bugarel M, Nagaraja TG, Renter DG, den Bakker HC,
        Nightingale KK, Fach P, Loneragan GH. 2016. Targeted amplicon sequencing for
        single-nucleotide-polymorphism genotyping of attaching and effacing
        Escherichia coli O26:H11 cattle strains via a high-throughput library preparation
        technique. Appl Environ Microbiol 82:640–649.

251.   Bonanno L, Petit M-A, Loukiadis E, Michel V, Auvray F. 2016. Heterogeneity in
        induction level, infection ability, and morphology of Shiga toxin-encoding
        phages (*stx* phages) from dairy and human Shiga toxin-producing *Escherichia
        coli* O26:H11 isolates. Appl Environ Microbiol 82:2177–86.

252.   Hillerton JE, Irvine CR, Bryan MA, Scott D, Merchant SC. 2017. Use of
        antimicrobials for animals in New Zealand, and in comparison with other
        countries. N Z Vet J 65:71–77.

253.   Dallman TJ, Ashton PM, Byrne L, Perry NT, Petrovska L, Ellis R, Allison L,
        Hanson M, Holmes A, Gunn GJ, Chase-Topping ME, Woolhouse MEJ, Grant KA,
        Gally DL, Wain J, Jenkins C. 2015. Applying phylogenomics to understand the
        emergence of Shiga-toxin-producing *Escherichia coli* O157:H7 strains causing

severe human disease in the UK. Microb Genomics 1:1–13.

254. Ogura Y, Gotoh Y, Itoh T, Sato MP, Seto K, Yoshino S, Isobe J, Etoh Y, Kurogi M, Kimata K, Maeda E, Piérard D, Kusumoto M, Akiba M, Tominaga K, Kirino Y, Kato Y, Shirahige K, Ooka T, Ishijima N, Lee K-I, Iyoda S, Mainil JG, Hayashi T. 2017. Population structure of Escherichia coli O26:H11 with recent and repeated stx2 acquisition in multiple lineages. Microb Genomics 3:1–12.

255. Zimmerman WD. 1962. Live cattle export trade between United States and Great Britain, 1868-1885. Agric Hist 36:1868–1885.

256. Mangino E. 2017. GAIN Report: Japan, increasing interest in U.S . feeder cattle. USDA Foreign Agric Serv.

257. McConnell M, Dyck J. 2010. A report from the Economic Research Service: Japan's beef market. United States Dep Agric.

258. Franz E, Delaquis P, Morabito S, Beutin L, Gobius K, Rasko DA, Bono J, French N, Osek J, Lindstedt BA, Muniesa M, Manning S, LeJeune J, Callaway T, Beatson S, Eppinger M, Dallman T, Forbes KJ, Aarts H, Pearl DL, Gannon VPJ, Laing CR, Strachan NJC. 2014. Exploiting the explosion of information associated with whole genome sequencing to tackle Shiga toxin-producing *Escherichia coli* (STEC) in global food production systems. Int J Food Microbiol 187:57–72.

259. New Zealand Food Safety Science & Research Centre. 2018. Resources: STEC - Shiga toxin-producing *E.coli* (STEC), is a nasty toxin producing bacteria, also known as VTEC (Verotoxin producing *E.coli*). http://www.nzfssrc.org.nz/resources.

260. French N, Benschop J, Marshall J. 2014. Raw milk: Is it good for you? The consumption of raw milk in New Zealand. Proc 2014 New Zeal Milk Qual Conf 75–84.

261. New Zealand Herald: The Northern Advocate. 2016. MPI: Illnesses led to halting milk sales - NZ Herald. http://www.nzherald.co.nz/northern-advocate/news/article.cfm?c_id=1503450&objectid=11602449.

262. Quick J, Loman NJ, Duraffour S, Simpson JT, Severi E, Cowley L, Bore JA,

Koundouno R, Dudas G, Mikhail A, Ouédraogo N, Afrough B, Bah A, Baum JHJ, Becker-Ziaja B, Boettcher JP, Cabeza-Cabrerizo M, Camino-Sánchez Á, Carter LL, Doerrbecker J, Enkirch T, Dorival IG-, Hetzelt N, Hinzmann J, Holm T, Kafetzopoulou LE, Koropogui M, Kosgey A, Kuisma E, Logue CH, Mazzarelli A, Meisel S, Mertens M, Michel J, Ngabo D, Nitzsche K, Pallasch E, Patrono LV, Portmann J, Repits JG, Rickett NY, Sachse A, Singethan K, Vitoriano I, Yemanaberhan RL, Zekeng EG, Racine T, Bello A, Sall AA, Faye O, Faye O, Magassouba N, Williams C V, Amburgey V, Winona L, Davis E, Gerlach J, Washington F, Monteil V, Jourdain M, Bererd M, Camara A, Somlare H, Camara A, Gerard M, Bado G, Baillet B, Delaune D, Nebie KY, Diarra A, Savane Y, Pallawo RB, Gutierrez GJ, Milhano N, Roger I, Williams CJ, Yattara F, Lewandowski K, Taylor J, Rachwal P, J. Turner D, Pollakis G, Hiscox JA, Matthews DA, Shea MKO, Johnston AM, Wilson D, Hutley E, Smit E, Di Caro A, Wölfel R, Stoecker K, Fleischmann E, Gabriel M, Weller SA, Koivogui L, Diallo B, Keïta S, Rambaut A, Formenty P, Günther S, Carroll MW. 2016. Real-time, portable genome sequencing for Ebola surveillance. Nature 530:228.

263. Loman NJ, Quick J, Simpson JT. 2015. A complete bacterial genome assembled *de novo* using only nanopore sequencing data. Nature 12:733–736.

264. Wright T. 2017. Special report: how polluted are New Zealand's rivers? Newshub http://www.newshub.co.nz/home/new-zealand/2017/02/special-report-how-polluted-are-new-zealand-s-rivers.html.

265. The Department of Internal Affairs. 2017. Government Inquiry into Havelock North Drinking Water Report. https://www.dia.govt.nz/Government-Inquiry-into-Havelock-North-Drinking-Water-Report---Part-1---Overview.

8 Appendix

Appendix Table 1: Sensitivity of NeoSEEK and reverse transcriptase PCR assays for detection of Top 7 STEC serogroups in calf faecal enrichment samples (n=1,508)

| Serogroup | Sensitivity RT-PCR | | Sensitivity NeoSEEK | |
	Mean	95% credible interval	Mean	95% credible interval
O103	0.95	0.91-0.97	0.93	0.88-0.97
O145	0.89	0.81-0.96	0.92	0.85-0.97
O157	0.93	0.87-0.98	0.79	0.69-0.90
O121	0.91	0.84-0.97	0.87	0.77-0.95
O26	0.91	0.84-0.97	0.92	0.85-0.97
O45	0.93	0.87-0.98	0.93	0.87-0.97
O111	0.89	0.77-0.96	0.89	0.89-0.96

Appendix Table 2: Specificity of NeoSEEK and reverse transcriptase PCR assays for detection of Top 7 STEC serogroups in calf faecal enrichment samples (n=1,508)

| Serogroup | Specificity RT-PCR | | Specificity NeoSEEK | |
	Mean	95% CI	Mean	95% CI
O103	0.93	0.89-0.97	0.95	0.93-0.98
O145	0.98	0.97-0.99	0.96	0.94-0.98
O157	0.96	0.94-0.99	0.98	0.97-0.99
O121	0.96	0.94-0.98	0.96	0.95-0.98
O26	0.94	0.91-0.97	0.93	0.90-0.96
O45	0.95	0.93-0.97	0.93	0.91-0.96
O111	1.00	0.99-1.00	1.00	0.99-1.00

Appendix Table 3: Location, animal, pen, and farm management factors evaluated against outcome variables of the Top 7 STEC prevalence in dairy calves

Risk Factor Level	Variable	Type	Notes	Source of measurement
Location	Region	Categorical	6 regions	
	Elevation	Continuous	Meters	GPS device
	Humidity	Continuous	Measured in open space no less than 5 meters from Pen	Thermo-Hygrometer
	Temperature	Continuous	Measured in open space no less than 5 meters from Pen	Thermo-Hygrometer
Animal	Age	Binary	Young (2 to 9 days); Old (10 to 21 days)	
	Sex	Binary	Male, Female	
	Breed	Categorical		
	Class	Categorical	Bobby, Beef, Replacement	
	FaecalScore	Ordinal	Coverage of hide with faecal matter or mud: 1 (0-25%), 2 (25-50%), 3 (50-75%), 4 (75-100%)	Visual assessment of faecal contamination on hide
Calf Pen	PenAnimalCount	Continuous	Number of calves in pen	
	PenOrientation	Categorical	Primary direction that pen faced; 8 cardinal points	GPS device

	PenHumidity	Continuous	Recorded in centre of pen	Thermo-hygrometer
	PenTemperature	Continuous	Recorded in centre of pen	Thermo-hygrometer
	PenType	Categorical	Open, closed, or conversion	
	Floor	Categorical	Flooring beneath substrate	
	Substrate	Categorical	Type of substrate used as bedding for calves	
	SubstrateCleaned	Binary	Whether substrate was cleaned during the calving season	
	SubstrateCleanedFreq	Continuous	How often pen substrate was cleaned	
	SubstrateTopped	Binary	Was new pen substrate added during the season	
	SubstrateToppedFreq	Continuous	How often new substrate was added	
	SubstrateSprayed	Binary	Was the pen sprayed with disinfectant during the season	
	SubstrateSprayedFreq	Continuous	How often pen was sprayed	
	PenFecalContamination	Ordinal	4 Groups: 0-25%, 25-50%, 50-75%, 75-100%	Visual assessment of faecal matter on pen floor
	PenAmmonia	Binary	Irritation of mucous membranes (nasal passage, eyes) to sampler (ASB)	
	TempInsideOutside	Continuous	Temperature inside pen minus temperature outside the calf housing area (~10 meters from the building)	

	HumidityInsideOutside	Continuous	Humidity inside pen minus humidity outside calf housing area (~10 meters from the building)	
Farm Management	MilkingHerdSize	Continuous	Number of cows in milking herd	
	SpringCalvingSize	Continuous	Number of cows in spring calving herd	
	DairyOnly	Binary	Only dairy cattle raised	
	BeefDairy	Binary	Beef and dairy cattle raised	
	MeatCompany	Categorical	Meat company used for veal calf processing	
	Pigs	Binary	Presence of pigs on farm	
	Deer	Binary	Presence of deer on farm	
	Sheep	Binary	Presence of sheep on farm	
	Goats	Binary	Presence of goats on farm	
	Horses	Binary	Presence of horses on farm	
	CalvesOtherFarms	Binary	Importation of calves from other farms in past two calving seasons	
	CowsOtherFarms	Binary	Importation of cows from other farms in the past two calving seasons	
	CalvingDateSpring	Date	Planned start of winter calving	
	CalvingDateAutumn	Date	Planned start of autumn calving (if applicable)	
	DaysSinceStartofCalving	Continuous	Days from start of calving to day of sampling for study	
	YearRoundCalving	Binary	Unseasonal calving management	

	VectorFlies	Binary	Flies observed in Pens while sampling	
	VectorBirdDroppings	Binary	Bird droppings observed on surfaces in Pens while sampling	
	VectorBirdNests	Binary	Bird nests observed in Pens while sampling	
	VectorsRodenticide	Binary	Farmer asked if they use rodenticide to control rodents	
	BobbysPresent	Binary	Young veal calves in Pen while sampling	
	ReplacementsPresent	Binary	Young replacement calves in Pen while sampling	
	BeefPresent	Binary	Young beef calves in Pen while sampling	
	BobbysIsolated	Binary	Young veal calves not mixed with other classes (replacement/beef)	
	ReplacementsIsolated	Binary	Replacement calves isolated from other calves	
	BeefIsolated	Binary	Beef calves isolate from other calves	
	BobbysMixedReplacements	Binary	Young veal calves and replacements mixed together	
	BobbysMixedBeef	Binary	Young veal calves and beef calves mixed together	
	ReplacementsMixedBobbys	Binary	Replacement calves and young veal calves mixed together	
	ReplacementsMixedBeef	Binary	Replacement calves and beef calves mixed together	

	BeefMixedReplacements	Binary	Beef calves and replacement calved mixed together	
	BeefMixedBobbys	Binary	Beef calves and young veal calves mixed together	
	BobbyDaysExport	Continuous	Age of young veal calves when exported for veal meat processing	
	ReplacementsDaysPastured	Continuous	Days replacement calves kept in Pens before putting out to pasture	
	BeefDaysPastured	Continuous	Days beef calves kept in Pens before putting out to pasture	
	BeefDaysExport	Continuous	Age of beef calves when exported	
	AdultsWeanlingsMixedCalves	Binary	Do adult cattle or weaned calves have any contact with calves	
	DamBirthing	Categorical	Location where Dams give birth to calves	
	DaysCalvesWithDam	Continuous	Days before calves removed from dams	
	Colostrum	Binary	Directly from dam or mixed colostrum	
	ColostrumMethod	Categorical	Orogastric tube or suckled from dam	
	MilkFresh	Binary	Fresh milk given to calves	
	MilkPowder	Binary	Powdered milk given to calves	
	MilkWaste	Binary	Antibiotic milk given to calves	
	MilkColostrum	Binary	Colostrum fed to calves	
	MilkContainer	Categorical	Teat bucket or Open bucket for feeding	

	MilkFrequency	Categorical	How many times calves fed a day	
	AdlibWater	Binary	Water available for all calves in Pens	
	ConcentrateMeal	Binary	Feed concentrates provided to calves	
	ConcentrateHayStraw	Binary	Hay or straw provided to calves	
	ConcentrateClay	Binary	Clay additive provided to calves to prevent scours	
	ConcentrateOther	Categorical	Variety of other supplements used by farmers for calves	
	ConcentrateStartFeedDay	Continuous	Day farmer started giving calves concentrates	
	ConcentratesBobby	Binary	Young veal calves received concentrate feed	
	WaterSource	Categorical	Source of water on farm	
	EffluentMethod	Categorical	Spread on farmer's pasture or shipped out	
	EffluentFrequency	Continuous	How often effluent spread on pastures	
	Grazier	Binary	Farmer ships out replacement calves to grazier, and then brings back once they grow to heifers	
	FeedPadHerdHome	Binary	Use of feed pad or herd home for adult cattle	

Appendix Table 4: Detection of 'Top 7' STEC in calves (n=1508) on farms (n=102) by number of 'Top 7' STEC serogroups detected

Farm ID	Calves with no 'Top 7' STEC detected	Calves with one 'Top 7' STEC detected	Calves with two 'Top 7' STEC detected	Calves with three 'Top 7' STEC detected
VCF1	15			
VCF2	15			
VCF3	14	1		
VCF4	10	4	1	
VCF5	13	2		
VCF6	12		3	
VCF7	13	2		
VCF8	14	1		
VCF9	14	1		
VCF10	15			
VCF11	13	2		
VCF12	15			
VCF13	13	2		
VCF14	14	1		
VCF15	9	5	1	
VCF16	4	10	1	
VCF17	12	3		
VCF18	8	4	3	
VCF19	12	2	1	
VCF20	1	9	4	1
VCF21		3	7	5
VCF22	11	2	1	1
VCF23	15			
VCF24	15			
VCF25	15			

VCF26	10	3	2	
VCF27	3	10	2	
VCF28	13	1	1	
VCF29	14	1		
VCF30	14	1		
VCF31	9	6		
VCF32	8	3	3	1
VCF33	15			
VCF34	11	3	1	
VCF35	9	3	3	
VCF36	13	2		
VCF37	8	2	4	1
VCF38	13	2		
VCF39	12	3		
VCF40	15			
VCF41	15			
VCF42	14	1		
VCF43	10	1	1	
VCF44	15			
VCF45	15			
VCF46	15			
VCF47	11	3	1	
VCF48	12	3		
VCF49	15			
VCF50	15			
VCF51	13	2		
VCF52	14		1	
VCF53	14	1		
VCF54	13	2		
VCF55	13	2		

VCF56	5	5	3	2
VCF57	8	2	5	
VCF58	10	4	1	
VCF59	9	4	2	
VCF60	15			
VCF61	14	1		
VCF62	15			
VCF63	9	4		
VCF64	5	4	6	
VCF65	14	1		
VCF66	15			
VCF67	8			
VCF68	14	1		
VCF69	13	2		
VCF70	15			
VCF71	7	8		
VCF72	11	4		
VCF73	15			
VCF74	15			
VCF75	13	2		
VCF76	11	2	1	
VCF77	11	3	1	
VCF78	15			
VCF79	15			
VCF80	11	2	2	
VCF81	9	6		
VCF82	14	1		
VCF83	14	1		
VCF84	12	3		
VCF85	11	4		

VCF86	6			
VCF87	11	4		
VCF88	13	2		
VCF89	9	3	2	1
VCF90	13	1	1	
VCF91	13	1		1
VCF92	15			
VCF93	7	6	2	
VCF94	12	3		
VCF95	11	2	2	
VCF96	8	4	3	
VCF97	13		2	
VCF98	14		1	
VCF99	11	4		
VCF100	7	7	1	
VCF101	12	3		
VCF102	11	4		
Totals	1202	217	76	13

Appendix Table 5: BioProject PRJNA396667 metadata, publicly available on the NCBI SRA archive

BioSample accession	sample_name	collection_date	geo_loc_name	Isolation source	lat_lon	genotype	host	host_age	host	host_disease	serotype
SAMN07430747	15ER0857	27-Feb-15	New Zealand: Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430748	15ER2837	28-Aug-15	New Zealand: Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430749	15ER4241	7-Dec-15	New Zealand: Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430750	16ER0672	8-Jul-05	New Zealand: Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430751	16ER1432	9-Apr-16	New Zealand: Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430752	ERL093655	31-Oct-09	New Zealand: Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430753	H13ESR01843	20-Mar-13	New Zealand: Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430754	15ER3334	26-Sep-15	New Zealand: Canterbury	human	43.75 S 171.16 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430755	A14a	17-Aug-09	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	A14	Reservoir	O26:H11
SAMN07430756	A17a	17-Aug-09	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	A17	Reservoir	O26:H11
SAMN07430757	a185c	19-Jan-11	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Cow	a185	Reservoir	O26:H11
SAMN07430758	a234d	15-Jun-11	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Cow	a234	Reservoir	O26:H11
SAMN07430759	A46a	30-Jul-10	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	A46	Reservoir	O26:H11
SAMN07430760	A65a	16-Aug-10	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	A65	Reservoir	O26:H11
SAMN07430761	A65b	16-Aug-10	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	A65	Reservoir	O26:H11
SAMN07430762	A87b	6-Sep-10	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	A87	Reservoir	O26:H11
SAMN07430763	ERL111686	6-May-11	New Zealand: Canterbury	human	43.75 S 171.16 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430764	VC1139e	11-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1139	Reservoir	O26:H11
SAMN07430765	VC1140e	11-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1140	Reservoir	O26:H11
SAMN07430766	VC1186e	15-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1186	Reservoir	O26:H11
SAMN07430767	VC1187e	15-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1187	Reservoir	O26:H11
SAMN07430768	VC1190e	15-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1190	Reservoir	O26:H11
SAMN07430769	VC1195e	15-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1195	Reservoir	O26:H11

SAMN07430770	VC1196e	15-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1196	Reservoir	O26:H11
SAMN07430771	VC1202e	15-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1202	Reservoir	O26:H11
SAMN07430772	VC1309e	17-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1309	Reservoir	O26:H11
SAMN07430773	VC1310e	17-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1310	Reservoir	O26:H11
SAMN07430774	VC1311e	17-Sep-14	New Zealand: Canterbury	bovine	43.75 S 171.16 E	O26:H11	Bos taurus taurus	Calf	VC1311	Reservoir	O26:H11
SAMN07430775	15ER0512	3-Feb-15	New Zealand: Combined Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430776	15ER2552	12-Aug-15	New Zealand: Combined Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430777	16ER0520	4-Feb-16	New Zealand: Combined Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430778	16ER1646	30-Apr-16	New Zealand: Combined Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430779	129ST2	5-Sep-08	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf	129ST	Reservoir	O26:H11
SAMN07430780	22ST2	19-Aug-08	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf	22ST	Reservoir	O26:H11
SAMN07430781	AGR373	12-Dec-02	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf		Reservoir	O26:H11
SAMN07430782	AGR72	21-Nov-02	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf		Reservoir	O26:H11
SAMN07430783	VC833e	27-Aug-14	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC833	Reservoir	O26:H11
SAMN07430784	VC833f	27-Aug-14	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC833	Reservoir	O26:H11
SAMN07430785	VC833g	27-Aug-14	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC833	Reservoir	O26:H11
SAMN07430786	VC833h	27-Aug-14	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC833	Reservoir	O26:H11
SAMN07430787	VC836e	27-Aug-14	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC836	Reservoir	O26:H11
SAMN07430788	VC837e	27-Aug-14	New Zealand: Manawatu-Wellington	bovine	39.73 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC837	Reservoir	O26:H11
SAMN07430789	14ER3837	13-Nov-14	New Zealand: Nelson Marlborough	human	41.57 S 173.42 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430790	15ER0146	7-Jan-15	New Zealand: Nelson Marlborough	human	41.57 S 173.42 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430791	15ER3567	13-Oct-15	New Zealand: Nelson Marlborough	human	41.57 S 173.42 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430792	16ER0248	13-Jan-16	New Zealand: Nelson Marlborough	human	41.57 S 173.42 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430793	16ER1973	26-May-16	New Zealand: Nelson Marlborough	human	41.57 S 173.42 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430794	ERL071565	10-May-07	New Zealand: Nelson Marlborough	human	41.57 S 173.42 E	O26:H11	Homo sapiens			STEC	O26:H11

SAMN07430795	16ER0243	14-Jan-16	New Zealand: Northland	human	35.58 S 173.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430796	16ER1209	23-Mar-16	New Zealand: Northland	human	35.58 S 173.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430797	ERL121992	20-Feb-12	New Zealand: Northland	human	35.58 S 173.76 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430798	H108a	21-Aug-09	New Zealand: Northland	bovine	35.58 S 173.76 E	O26:H11	Bos taurus taurus	Calf	H108	Reservoir	O26:H11
SAMN07430799	H113a	21-Aug-09	New Zealand: Northland	bovine	35.58 S 173.76 E	O26:H11	Bos taurus taurus	Calf	H113	Reservoir	O26:H11
SAMN07430800	16ER0946	4-Mar-16	New Zealand: Southland	human	45.85 S 167.68 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430801	T11i	1-Sep-09	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	T11	Reservoir	O26:H11
SAMN07430802	t126c	21-Sep-10	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Cow	t126	Reservoir	O26:H11
SAMN07430803	t128a	21-Dec-10	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Cow	t128	Reservoir	O26:H11
SAMN07430804	t173a	20-Apr-11	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Cow	t173	Reservoir	O26:H11
SAMN07430805	t25a	17-Dec-09	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Cow	t25	Reservoir	O26:H11
SAMN07430806	t25c	17-Dec-09	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Cow	t25	Reservoir	O26:H11
SAMN07430807	T27a	14-Sep-09	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	T27	Reservoir	O26:H11
SAMN07430808	T48b	17-Aug-10	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	T48	Reservoir	O26:H11
SAMN07430809	T66c	17-Aug-10	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	T66	Reservoir	O26:H11
SAMN07430810	VC1362e	22-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1362	Reservoir	O26:H11
SAMN07430811	VC1366e	22-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1366	Reservoir	O26:H11
SAMN07430812	VC1367e	22-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1367	Reservoir	O26:H11
SAMN07430813	VC1394e	23-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1394	Reservoir	O26:H11
SAMN07430814	VC1395e	23-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1395	Reservoir	O26:H11
SAMN07430815	VC1395f	23-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1395	Reservoir	O26:H11
SAMN07430816	VC1395g	23-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1395	Reservoir	O26:H11
SAMN07430817	VC1396e	23-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1396	Reservoir	O26:H11
SAMN07430818	VC1403e	23-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1403	Reservoir	O26:H11
SAMN07430819	VC1471e	24-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1471	Reservoir	O26:H11

SAMN07430820	VC1471f	24-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1471	Reservoir	O26:H11
SAMN07430821	VC1471g	24-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1471	Reservoir	O26:H11
SAMN07430822	VC1471h	24-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1471	Reservoir	O26:H11
SAMN07430823	VC1473e	24-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1473	Reservoir	O26:H11
SAMN07430824	VC1474e	24-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1474	Reservoir	O26:H11
SAMN07430825	VC1486e	24-Sep-14	New Zealand: Southland	bovine	45.85 S 167.68 E	O26:H11	Bos taurus taurus	Calf	VC1486	Reservoir	O26:H11
SAMN07430826	100ST2	1-Sep-08	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	100ST	Reservoir	O26:H11
SAMN07430827	11ST	19-Aug-08	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	11S	Reservoir	O26:H11
SAMN07430828	191ST2	18-Sep-08	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	191ST	Reservoir	O26:H11
SAMN07430829	66ST1	27-Aug-08	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	66ST	Reservoir	O26:H11
SAMN07430830	E13c	29-Jul-09	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	E13	Reservoir	O26:H11
SAMN07430831	e171b	22-Feb-11	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Cow	e171	Reservoir	O26:H11
SAMN07430832	e171c	22-Feb-11	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Cow	e171	Reservoir	O26:H11
SAMN07430833	e186a	22-Mar-11	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Cow	e186	Reservoir	O26:H11
SAMN07430834	E189b	13-Sep-10	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	E189	Reservoir	O26:H11
SAMN07430835	e195b	18-Apr-11	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Cow	e195	Reservoir	O26:H11
SAMN07430836	e26a	22-Dec-09	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Cow	e26	Reservoir	O26:H11
SAMN07430837	e37c	18-Jan-10	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Cow	e37	Reservoir	O26:H11
SAMN07430838	E46a	25-Aug-09	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	E46	Reservoir	O26:H11
SAMN07430839	E59b	7-Sep-09	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	E59	Reservoir	O26:H11
SAMN07430840	VC1113e	4-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC1113	Reservoir	O26:H11
SAMN07430841	VC1122f	4-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC1122	Reservoir	O26:H11
SAMN07430842	VC1125e	4-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC1125	Reservoir	O26:H11
SAMN07430843	VC880e	1-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC880	Reservoir	O26:H11
SAMN07430844	VC932f	1-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC932	Reservoir	O26:H11

SAMN07430845	VC936e	1-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC936	Reservoir	O26:H11
SAMN07430846	VC940e	1-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC940	Reservoir	O26:H11
SAMN07430847	VC943e	2-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC943	Reservoir	O26:H11
SAMN07430848	VC943f	2-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC943	Reservoir	O26:H11
SAMN07430849	VC943g	2-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC943	Reservoir	O26:H11
SAMN07430850	VC943h	2-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC943	Reservoir	O26:H11
SAMN07430851	VC946e	2-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC946	Reservoir	O26:H11
SAMN07430852	VC951e	2-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC951	Reservoir	O26:H11
SAMN07430853	VC955e	2-Sep-14	New Zealand: Taranaki	bovine	39.35 S 174.44 E	O26:H11	Bos taurus taurus	Calf	VC955	Reservoir	O26:H11
SAMN07430854	ER005420	22-Jun-05	New Zealand: Unknown	human	39.28 S 175.57 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430855	ER854674	7-Jun-05	New Zealand: Unknown	human	39.28 S 175.57 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430856	ER985544	20-Jun-05	New Zealand: Unknown	human	39.28 S 175.57 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430857	ERL023841	24-Jun-05	New Zealand: Unknown	human	39.28 S 175.57 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430858	14ER2056	12-Jul-14	New Zealand: Waikato	human	38.06 S 175.44 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430859	14ER2240	31-Jul-14	New Zealand: Waikato	human	38.06 S 175.44 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430860	16ER1892	20-May-16	New Zealand: Waikato	human	38.06 S 175.44 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430861	16ER2139	18-Jun-16	New Zealand: Waikato	human	38.06 S 175.44 E	O26:H11	Homo sapiens			STEC	O26:H11
SAMN07430862	H132a	21-Aug-09	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	H132	Reservoir	O26:H11
SAMN07430863	h148a	21-Feb-11	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Cow	h148	Reservoir	O26:H11
SAMN07430864	h148b	21-Feb-11	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Cow	h148	Reservoir	O26:H11
SAMN07430865	h199a	14-Jun-11	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Cow	h199	Reservoir	O26:H11
SAMN07430866	h199c	14-Jun-11	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Cow	h199	Reservoir	O26:H11
SAMN07430867	H31c	28-Jul-09	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	H31	Reservoir	O26:H11
SAMN07430868	Hide14d	10-Aug-10	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	Hide14	Reservoir	O26:H11
SAMN07430869	ND29	6-Nov-09	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	ND2	Reservoir	O26:H11

SAMN07430870	ND3	21-Oct-09	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	ND	Reservoir	O26:H11
SAMN07430871	ND35	12-Nov-09	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	ND3	Reservoir	O26:H11
SAMN07430872	ND62	24-Nov-09	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	ND6	Reservoir	O26:H11
SAMN07430873	Pre15a	26-Jul-10	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	Pre15	Reservoir	O26:H11
SAMN07430874	Pre4d	3-Aug-10	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	Pre4	Reservoir	O26:H11
SAMN07430875	VC396e	12-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC396	Reservoir	O26:H11
SAMN07430876	VC397e	12-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC397	Reservoir	O26:H11
SAMN07430877	VC401e	12-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC401	Reservoir	O26:H11
SAMN07430878	VC452e	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC452	Reservoir	O26:H11
SAMN07430879	VC456e	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC456	Reservoir	O26:H11
SAMN07430880	VC459e	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC459	Reservoir	O26:H11
SAMN07430881	VC473e	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC473	Reservoir	O26:H11
SAMN07430882	VC474e	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC474	Reservoir	O26:H11
SAMN07430883	VC474f	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC474	Reservoir	O26:H11
SAMN07430884	VC474g	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC474	Reservoir	O26:H11
SAMN07430885	VC474h	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC474	Reservoir	O26:H11
SAMN07430886	VC476e	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC476	Reservoir	O26:H11
SAMN07430887	VC477e	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC477	Reservoir	O26:H11
SAMN07430888	VC479e	13-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC479	Reservoir	O26:H11
SAMN07430889	VC545e	18-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC545	Reservoir	O26:H11
SAMN07430890	VC547e	18-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC547	Reservoir	O26:H11
SAMN07430891	VC550e	18-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC550	Reservoir	O26:H11
SAMN07430892	VC554e	18-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC554	Reservoir	O26:H11
SAMN07430893	VC555e	18-Aug-14	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VC555	Reservoir	O26:H11
SAMN07430894	VL0828h	20-Aug-15	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VL0828	Reservoir	O26:H11

SAMN07430895	VL0958f	1-Sep-15	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VL0958	Reservoir	O26:H11
SAMN07430896	VL1058g	7-Sep-15	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VL1058	Reservoir	O26:H11
SAMN07430897	VL1277e	12-Jul-16	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VL1277	Reservoir	O26:H11
SAMN07430898	VL2020e	9-Aug-16	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VL2020	Reservoir	O26:H11
SAMN07430899	VL2228e	29-Aug-16	New Zealand: Waikato	bovine	38.06 S 175.44 E	O26:H11	Bos taurus taurus	Calf	VL2228	Reservoir	O26:H11
SAMN07430900	15ER3804	1-Nov-15	New Zealand: Auckland	human	36.85 S 174.76 E	O26:H11	Homo sapiens			STEC	O26:H11

Appendix Table 6: Bioproject, BioSample, SRA run number, country, source, and unique identifier for all public sequence data used in this study (n=252)

BioProject	BioSample	SRA Run	country	source	current_study_ID
PRJNA230969	SAMN05607363	SRR5330941	Australia	human	MOD1EC1684_Australia_human
PRJDB5571	SAMD00075768	DRR103425	Belgium	bovine	357S89_Belgium_bovine
PRJDB5571	SAMD00075735	DRR103392	Belgium	human	EH031_Belgium_human
PRJDB5571	SAMD00075761	DRR103418	Belgium	bovine	631KH91_Belgium_bovine
PRJDB5571	SAMD00075766	DRR103423	Belgium	human	EH182_Belgium_human
PRJDB5571	SAMD00075767	DRR103424	Belgium	human	EH193_Belgium_human
PRJDB5571	SAMD00075736	DRR103393	Belgium	human	EH322_Belgium_human
PRJDB5571	SAMD00075764	DRR103421	Belgium	bovine	B44_Belgium_bovine
PRJDB5571	SAMD00075763	DRR103420	Belgium	bovine	B43_Belgium_bovine
PRJDB5571	SAMD00075771	DRR103428	Belgium	bovine	11KH263_Belgium_bovine
PRJDB5571	SAMD00075769	DRR103426	Belgium	bovine	11KH63_Belgium_bovine
PRJDB5571	SAMD00075770	DRR103427	Belgium	bovine	11KH245_Belgium_bovine
PRJDB5571	SAMD00075971	DRR103628	Belgium	human	EH2035_Belgium_human
PRJDB5571	SAMD00075976	DRR103633	Belgium	human	EH2083_Belgium_human
PRJDB5571	SAMD00075975	DRR103632	Belgium	human	EH2075_Belgium_human
PRJDB5571	SAMD00075974	DRR103631	Belgium	human	EH2068_Belgium_human
PRJDB5571	SAMD00075772	DRR103429	Belgium	bovine	12KH23_Belgium_bovine
PRJDB5571	SAMD00075981	DRR103638	Belgium	human	EH2208_Belgium_human
PRJDB5571	SAMD00075984	DRR103641	Belgium	human	EH2244_Belgium_human
PRJDB5571	SAMD00075983	DRR103640	Belgium	human	EH2219_Belgium_human
PRJDB5571	SAMD00075982	DRR103639	Belgium	human	EH2209_Belgium_human
PRJDB5571	SAMD00075988	DRR103645	Belgium	human	EH2258_Belgium_human
PRJDB5571	SAMD00075986	DRR103643	Belgium	human	EH2252_Belgium_human
PRJDB5571	SAMD00075987	DRR103644	Belgium	human	EH2257_Belgium_human
PRJDB5571	SAMD00075985	DRR103642	Belgium	human	EH2251_Belgium_human
PRJNA319494	SAMN05504941*	SRR6154941	Canada	bovine	OLC0637_Canada_bovine
PRJNA309770	SAMN04420181*	SRR6061322	Germany	human	126814_Germany_human
PRJNA301341	SAMN04254589	SRR3110022	Canada	human	EC120246_Canada_human
PRJEB10700	SAMEA3529294	ERR1010233	Denmark	human	AA044_Denmark_human
PRJDB5571	SAMD00075758	DRR103415	France	human	99109_France_human
PRJDB5571	SAMD00075757	DRR103414	France	human	02113_France_human
PRJDB5571	SAMD00075759	DRR103416	France	human	03139_France_human
PRJNA230969	SAMN05605330	SRR5330864	Germany	human	MOD1EC2814_Germany_human
PRJNA230969	SAMN05607379	SRR5330926	Germany	human	MOD1EC1664_Germany_human
PRJDB5571	SAMD00075765	DRR103422	Italy	bovine	ED80_Italy_bovine
PRJNA230969	SAMN06555271	SRR5336246	Mexico	food	MOD1EC5336_Mexico_food
PRJEB23743	SAMEA104413463	ERR2210764	Poland	human	10016_Poland_human
PRJDB5571	SAMD00075756	DRR103413	Switzerland	human	TC6167_Switzerland_human
PRJDB5571	SAMD00075762	DRR103419	UK	human	H19_UK_human
PRJNA419720	SAMN08095914	SRR6321366	UK	human	2M8BS8_UK_human
PRJNA419720	SAMN08095913	SRR6321365	UK	human	I20VK7_UK_human
PRJNA419720	SAMN08095930	SRR6321271	UK	human	KLAV92_UK_human
PRJNA419720	SAMN08095938	SRR6321331	UK	human	HKCVSH_UK_human
PRJEB4681	SAMEA2204500	ERR435109	UK	human	ECO0283_UK_humansepsis
PRJNA315192	SAMN06030740	SRR5031110	UK	bovine	211644_UK_bovine
PRJDB5571	SAMD00075752	DRR103409	USA	human	TC6165_USA_human
PRJDB5571	SAMD00075754	DRR103411	USA	bovine	TC6169_USA_bovine
PRJDB5571	SAMD00075747	DRR103404	USA	bovine	TC3486_USA_bovine
PRJDB5571	SAMD00075748	DRR103405	USA	bovine	TC3630_USA_bovine
PRJDB5571	SAMD00075749	DRR103406	USA	bovine	TC3656_USA_bovine
PRJDB5571	SAMD00075751	DRR103408	USA	bovine	TC4219_USA_bovine
PRJDB5571	SAMD00075753	DRR103410	USA	human	TC6168_USA_human
PRJNA218110	SAMN04498710	SRR3178054	USA	human	2009C3689_USA_human
PRJNA218110	SAMN04498712	SRR3178056	USA	human	2009C3996_USA_human
PRJNA218110	SAMN04633589	SRR3371771	USA	human	2009C4747_USA_human

PRJNA218110	SAMN04633622	SRR3371781	USA	human	2010C3051_USA_human
PRJNA218110	SAMN04625574	SRR3360206	USA	human	2010C3902_USA_human
PRJNA218110	SAMN04625585	SRR3360195	USA	human	2010C4430_USA_human
PRJNA218110	SAMN04625563	SRR3360216	USA	human	2011C3270_USA_human
PRJNA218110	SAMN04913811	SRR4113678	USA	human	PNUSAE002149_USA_human
PRJNA218110	SAMN04913824	SRR4300141	USA	human	PNUSAE002166_USA_human
PRJNA218110	SAMN04625466	SRR3360241	USA	human	2011C3506_USA_human
PRJNA218110	SAMN04578418	SRR3290033	USA	human	2012C3101_USA_human
PRJNA218110	SAMN04495854	SRR3171841	USA	human	2012C3912_USA_human
PRJNA218110	SAMN04578419	SRR3290038	USA	human	2012C3102_USA_human
PRJNA218110	SAMN04498549	SRR3178026	USA	human	2012C4606_USA_human
PRJNA218110	SAMN04192188	SRR3040537	USA	human	PNUSAE001578_USA_human
PRJNA218110	SAMN03838116	SRR2481234	USA	human	PNUSAE000885_USA_human
PRJNA218110	SAMN04227723	SRR3040532	USA	human	PNUSAE001573_USA_human
PRJNA218110	SAMN04075848	SRR2417066	USA	human	PNUSAE001154_USA_human
PRJNA218110	SAMN03272820	SRR1738019	USA	human	PNUSAE000133_USA_human
PRJNA218110	SAMN03151532	SRR1635531	USA	human	PNUSAE000002_USA_human
PRJNA218110	SAMN03840334	SRR2121025	USA	human	PNUSAE000539_USA_human
PRJNA218110	SAMN04075513	SRR2415794	USA	human	PNUSAE001379_USA_human
PRJNA218110	SAMN04075843	SRR2415808	USA	human	PNUSAE001373_USA_human
PRJNA218110	SAMN03775204	SRR2481344	USA	human	PNUSAE000779_USA_human
PRJNA218110	SAMN04500985	SRR3189440	USA	human	PNUSAE002228_USA_human
PRJNA218110	SAMN04588711	SRR3371981	USA	human	PNUSAE002615_USA_human
PRJNA218110	SAMN05209084	SRR3644551	USA	human	PNUSAE003275_USA_human
PRJNA218110	SAMN05203326	SRR3644569	USA	human	PNUSAE003211_USA_human
PRJNA218110	SAMN07373072	SRR5870554	USA	human	PNUSAE008468_USA_human
PRJNA230969	SAMN05605272	SRR5330849	USA	human	MOD1EC1750_USA_human
PRJNA230969	SAMN04902887	SRR3465501	USA	other animal	MOD1EC6201_USA_other animal
PRJNA230969	SAMN05605257	SRR5330857	USA	human	MOD1EC1919_USA_human
PRJNA230969	SAMN05591573	SRR5330824	USA	human	MOD1EC550_USA_human
PRJNA268206	SAMN06256289	SRR5202193	USA	bovine	FSIS1609416_USA_bovine
PRJNA268206	SAMN04908471	SRR3457631	USA	bovine	FSIS1606391_USA_bovine
PRJNA268206	SAMN06127049	SRR5091629	USA	bovine	FSIS1608854_USA_bovine
PRJNA268206	SAMN06127045	SRR5091628	USA	bovine	FSIS1608722_USA_bovine
PRJNA268206	SAMN07237071	SRR5683240	USA	bovine	FSIS1701668_USA_bovine
PRJNA268206	SAMN06700862	SRR5441623	USA	bovine	FSIS1710186_USA_bovine
PRJNA268206	SAMN07987839	SRR6265848	USA	bovine	FSIS11704781_USA_bovine
PRJNA268206	SAMN07774189	SRR6158105	USA	bovine	FSIS21720313_USA_bovine
PRJNA218110	SAMN05294505	SRR3883019	USA	human	PNUSAE003398_USA_human
PRJNA218110	SAMN02352904	SRR3213940	USA	human	643464_USA_human
PRJNA218110	SAMN02352964	SRR3371784	USA	human	2010C3472_USA_human
PRJNA218110	SAMN08129177	SRR6359280	USA	human	PNUSAE011184_USA_human
PRJNA230969	SAMN05439479	SRR3988028	USA	food	MOD1EC6029_USA_food
PRJNA230969	SAMN05605269	SRR5330852	USA	human	MOD1EC1753_USA_human
PRJNA230969	SAMN05605332	SRR5185399	Germany	human	MOD1EC2812_Germany_human
PRJNA230969	SAMN05605331	SRR5185402	Germany	human	MOD1EC2813_Germany_human
PRJNA230969	SAMN05605263	SRR5185394	Germany	bovine	MOD1EC1763_Germany_bovine
PRJNA230969	SAMN03743659	SRR2176280	Canada	human	CFSAN033951_Canada_human
PRJDB5571	SAMD00075828	DRR103485	Japan	human	M01_Japan_human
PRJDB5571	SAMD00075829	DRR103486	Japan	human	M02_Japan_human
PRJDB5571	SAMD00075830	DRR103487	Japan	human	M03_Japan_human
PRJDB5571	SAMD00075832	DRR103489	Japan	human	M05_Japan_human
PRJDB5571	SAMD00075905	DRR103562	Japan	human	O01_Japan_human
PRJDB5571	SAMD00075998	DRR103655	Japan	human	T02_Japan_human
PRJDB5571	SAMD00075999	DRR103656	Japan	human	T03_Japan_human
PRJDB5571	SAMD00075911	DRR103568	Japan	human	O07_Japan_human
PRJDB5571	SAMD00075785	DRR103442	Japan	human	F02_Japan_human
PRJDB5571	SAMD00076003	DRR103660	Japan	human	T08_Japan_human
PRJDB5571	SAMD00075908	DRR103565	Japan	human	O04_Japan_human

PRJDB5571	SAMD00075912	DRR103569	Japan	human	O08_Japan_human
PRJDB5571	SAMD00075786	DRR103443	Japan	human	F03_Japan_human
PRJDB5571	SAMD00076004	DRR103661	Japan	human	T09_Japan_human
PRJDB5571	SAMD00076010	DRR103667	Japan	human	T16_Japan_human
PRJDB5571	SAMD00075789	DRR103446	Japan	human	F06_Japan_human
PRJDB5571	SAMD00075848	DRR103505	Japan	human	M21_Japan_human
PRJDB5571	SAMD00075918	DRR103575	Japan	human	O16_Japan_human
PRJDB5571	SAMD00076017	DRR103674	Japan	human	T23_Japan_human
PRJDB5571	SAMD00075930	DRR103587	Japan	human	O29_Japan_human
PRJDB5571	SAMD00075927	DRR103584	Japan	human	O25_Japan_human
PRJDB5571	SAMD00075852	DRR103509	Japan	human	M25_Japan_human
PRJDB5571	SAMD00075793	DRR103450	Japan	human	F10_Japan_human
PRJDB5571	SAMD00075728	DRR103385	Japan	bovine	Aki01_Japan_bovine
PRJDB5571	SAMD00075856	DRR103513	Japan	human	M30_Japan_human
PRJDB5571	SAMD00076025	DRR103682	Japan	human	T32_Japan_human
PRJDB5571	SAMD00075936	DRR103593	Japan	human	O35_Japan_human
PRJDB5571	SAMD00076024	DRR103681	Japan	human	T31_Japan_human
PRJDB5571	SAMD00076026	DRR103683	Japan	human	T33_Japan_human
PRJDB5571	SAMD00076028	DRR103685	Japan	human	T35_Japan_human
PRJDB5571	SAMD00075799	DRR103456	Japan	human	F17_Japan_human
PRJDB5571	SAMD00076030	DRR103687	Japan	human	T39_Japan_human
PRJDB5571	SAMD00076032	DRR103689	Japan	human	T41_Japan_human
PRJDB5571	SAMD00075857	DRR103514	Japan	human	M31_Japan_human
PRJDB5571	SAMD00075947	DRR103604	Japan	human	O46_Japan_human
PRJDB5571	SAMD00075859	DRR103516	Japan	human	M33_Japan_human
PRJDB5571	SAMD00076034	DRR103691	Japan	human	T43_Japan_human
PRJDB5571	SAMD00076036	DRR103693	Japan	human	T45_Japan_human
PRJDB5571	SAMD00075869	DRR103526	Japan	human	M43_Japan_human
PRJDB5571	SAMD00075867	DRR103524	Japan	human	M41_Japan_human
PRJDB5571	SAMD00075875	DRR103532	Japan	human	M53_Japan_human
PRJDB5571	SAMD00075877	DRR103534	Japan	human	M56_Japan_human
PRJDB5571	SAMD00075874	DRR103531	Japan	human	M48_Japan_human
PRJDB5571	SAMD00075881	DRR103538	Japan	human	M64_Japan_human
PRJDB5571	SAMD00076063	DRR103720	Japan	bovine	YB02_Japan_bovine
PRJDB5571	SAMD00076062	DRR103719	Japan	bovine	YB01_Japan_bovine
PRJDB5571	SAMD00075882	DRR103539	Japan	human	M65_Japan_human
PRJDB5571	SAMD00076040	DRR103697	Japan	human	T49_Japan_human
PRJDB5571	SAMD00076044	DRR103701	Japan	human	T53_Japan_human
PRJDB5571	SAMD00075887	DRR103544	Japan	human	M73_Japan_human
PRJDB5571	SAMD00076047	DRR103704	Japan	human	T59_Japan_human
PRJDB5571	SAMD00076049	DRR103706	Japan	human	T61_Japan_human
PRJDB5571	SAMD00075891	DRR103548	Japan	human	M79_Japan_human
PRJDB5571	SAMD00075892	DRR103549	Japan	human	M80_Japan_human
PRJDB5136	SAMD00064344	DRR073024	Japan	human	NIID070765_Japan_human
PRJDB5571	SAMD00076064	DRR103721	Japan	bovine	YB03_Japan_bovine
PRJDB5571	SAMD00075955	DRR103612	Japan	human	O55_Japan_human
PRJDB5571	SAMD00075956	DRR103613	Japan	human	O56_Japan_human
PRJDB5136	SAMD00064355	DRR073035	Japan	human	NIID080884_Japan_human
PRJDB5571	SAMD00076067	DRR103724	Japan	bovine	YB13_Japan_bovine
PRJDB5571	SAMD00076065	DRR103722	Japan	bovine	YB05_Japan_bovine
PRJDB5571	SAMD00075782	DRR103439	Japan	bovine	BK13_Japan_bovine
PRJDB5571	SAMD00076066	DRR103723	Japan	bovine	YB06_Japan_bovine
PRJDB5571	SAMD00075893	DRR103550	Japan	human	M81_Japan_human
PRJDB5571	SAMD00076055	DRR103712	Japan	human	T67_Japan_human
PRJDB5571	SAMD00075958	DRR103615	Japan	human	O58_Japan_human
PRJDB5571	SAMD00075957	DRR103614	Japan	human	O57_Japan_human
PRJDB5571	SAMD00075898	DRR103555	Japan	human	M89_Japan_human
PRJDB5571	SAMD00075901	DRR103558	Japan	human	M92_Japan_human
PRJDB5571	SAMD00075899	DRR103556	Japan	human	M90_Japan_human

PRJDB5571	SAMD00075824	DRR103481	Japan	human	F44_Japan_human
PRJDB5571	SAMD00075780	DRR103437	Japan	bovine	BK10_Japan_bovine
PRJDB5571	SAMD00075776	DRR103433	Japan	bovine	BK05_Japan_bovine
PRJDB5571	SAMD00075781	DRR103438	Japan	bovine	BK11_Japan_bovine
PRJDB5571	SAMD00075773	DRR103430	Japan	bovine	BK01_Japan_bovine
PRJDB5571	SAMD00075962	DRR103619	Japan	human	O62_Japan_human
PRJDB5136	SAMD00064358	DRR073038	Japan	human	NIID111609_Japan_human
PRJDB5571	SAMD00075964	DRR103621	Japan	human	O64_Japan_human
PRJDB5571	SAMD00075902	DRR103559	Japan	human	M93_Japan_human
PRJDB5136	SAMD00064338	DRR073018	Japan	human	NIID122711_Japan_human
PRJDB5136	SAMD00064353	DRR073033	Japan	human	NIID121840_Japan_human
PRJDB5136	SAMD00064351	DRR073031	Japan	human	NIID122657_Japan_human
PRJDB5136	SAMD00064348	DRR073028	Japan	human	NIID122147_Japan_human
PRJDB5571	SAMD00075989	DRR103646	Japan	bovine	She01_Japan_bovine
PRJDB5571	SAMD00075733	DRR103390	Japan	bovine	Aki06_Japan_bovine
PRJDB5571	SAMD00075729	DRR103386	Japan	bovine	Aki02_Japan_bovine
PRJDB5571	SAMD00075991	DRR103648	Japan	bovine	She03_Japan_bovine
PRJDB5136	SAMD00064341	DRR073021	Japan	human	NIID132777_Japan_human
PRJDB5136	SAMD00064342	DRR073022	Japan	human	NIID141423_Japan_human
PRJDB5136	SAMD00064354	DRR073034	Japan	human	NIID132265_Japan_human
PRJDB5136	SAMD00064337	DRR073017	Japan	human	NIID130549_Japan_human
PRJDB5571	SAMD00075734	DRR103391	Japan	other animal	Aki07_Japan_other animal
PRJDB5136	SAMD00064350	DRR073030	Japan	human	NIID141424_Japan_human
PRJDB5136	SAMD00064343	DRR073023	Japan	human	NIID141425_Japan_human
PRJNA315192	SAMN04568167	SRR3241986	UK	human	93279_UK_human
PRJNA315192	SAMN04568166	SRR3241985	UK	human	93280_UK_human
PRJNA315192	SAMN04567831	SRR3240981	UK	human	93304_UK_human
PRJNA315192	SAMN04554948	SRR3226393	UK	human	224098_UK_human
PRJNA315192	SAMN05734323	SRR4181480	UK	human	146047_UK_human
PRJNA315192	SAMN05733983	SRR4179768	UK	human	143934_UK_human
PRJNA259827	SAMN03703971	SRR2035403	UK	human	46009_UK_human
PRJNA259827	SAMN03703997	SRR2035432	UK	human	18109_UK_human
PRJNA259827	SAMN03703954	SRR2120750	UK	human	25910_UK_human
PRJNA259827	SAMN03703951	SRR2120773	UK	human	60510_UK_human
PRJNA259827	SAMN03703966	SRR2120768	UK	human	46710_UK_human
PRJNA259827	SAMN03703953	SRR2120771	UK	human	51911_UK_human
PRJNA259827	SAMN03703952	SRR2035364	UK	human	48212_UK_human
PRJNA259827	SAMN03703969	SRR2035399	UK	human	62612_UK_human
PRJNA259827	SAMN03703970	SRR2035402	UK	human	68013_UK_human
PRJNA259827	SAMN03703965	SRR2035374	UK	human	63713_UK_human
PRJNA259827	SAMN03703967	SRR2035376	UK	human	67013_UK_human
PRJNA315192	SAMN04568149	SRR3241859	UK	human	93285_UK_human
PRJNA315192	SAMN05170656	SRR3578565	UK	human	129381_UK_human
PRJNA315192	SAMN05733904	SRR4176976	UK	human	143493_UK_human
PRJNA315192	SAMN05734341	SRR4181536	UK	human	154820_UK_human
PRJNA315192	SAMN05734380	SRR4181594	UK	human	152422_UK_human
PRJNA218110	SAMN02991228	SRR2014841	USA	human	2010C4800_USA_human
PRJNA218110	SAMN03838114	SRR2481236	USA	human	PNUSAE000887_USA_human
PRJNA218110	SAMN03785431	SRR2550380	USA	human	PNUSAE000800_USA_human
PRJNA268206	SAMN04909393	SRR3457938	USA	bovine	FSIS1503454_USA_bovine
PRJNA268206	SAMN03921926	SRR2125829	USA	bovine	FSIS1500874_USA_bovine
PRJNA268206	SAMN03922101	SRR2126001	USA	bovine	FSIS1500781_USA_bovine
PRJNA268206	SAMN03921925	SRR2125821	USA	bovine	FSIS1500873_USA_bovine
PRJNA268206	SAMN03921934	SRR2125835	USA	bovine	FSIS1500928_USA_bovine
PRJNA268206	SAMN03921935	SRR2125824	USA	bovine	FSIS1500948_USA_bovine
PRJNA268206	SAMN03940972	SRR2132070	USA	bovine	FSIS1503255_USA_bovine
PRJNA268206	SAMN03940974	SRR2132075	USA	bovine	FSIS1503257_USA_bovine
PRJNA268206	SAMN03940973	SRR2132074	USA	bovine	FSIS1503256_USA_bovine
PRJNA268206	SAMN03216753	SRR1693303	USA	bovine	FSIS1400371_USA_bovine

PRJNA268206	SAMN03921981	SRR2125877	USA	bovine	FSIS1500641_USA_bovine
PRJNA268206	SAMN03921972	SRR2125881	USA	bovine	FSIS1500632_USA_bovine
PRJNA268206	SAMN03922009	SRR2125932	USA	bovine	FSIS1500489_USA_bovine
PRJNA232925	SAMN02566899	SRR1272859	USA	bovine	ewgs1005_USA_bovine
PRJNA283914	SAMN05366684	SRR4011138	USA	other animal	WAPHLECOA00006_USA_caprine
PRJNA275276	SAMN03340674	SO174[#]	Norway	human	SO174_Norway_human
PRJNA275276	SAMN03340677	SO179[#]	Norway	human	SO179_Norway_human
PRJNA284656	SAMN03704971	36493[#]	France	human	36493_France_human
PRJNA284656	SAMN03704970	36293[#]	France	human	36293_France_human
PRJNA284656	SAMN03704969	36348[#]	France	human	36348_France_human
PRJNA284656	SAMN03704968	34870[#]	France	human	34870_France_human
PRJNA284656	SAMN03704966	36708[#]	France	human	36708_France_human
PRJNA284656	SAMN03704963	36084[#]	France	human	36084_France_human
PRJNA218110	SAMN04498563	SRR3178040	USA	human	2015C5206_USA_human
PRJNA218110	SAMN04498702	SRR3178046	USA	human	2016C3018_USA_human
PRJNA312475	SAMN05414569	SRR3931242	USA	food	CFSAN046724_USA_food
PRJNA312475	SAMN05414581	SRR3931256	USA	food	CFSAN046736_USA_food
PRJNA230969	SAMN05425824	SRR3938671	USA	food	FDA00010430_USA_food
PRJNA230969	SAMN05452918	SRR3987970	USA	avian	MOD1EC5703_USA_avian
PRJNA230969	SAMN05605333	SRR4340534	USA	human	MOD1EC2790_USA_human

*True origin of isolate found by contacting submitter and original source; not correctly listed on NCBI
#Raw sequence data provided by submitter of assembled genome directly to corresponding author of publication

Appendix Table 7: Serotype, multilocus sequence type (MLST), metadata (country and isolation source), and *stx* profile for serogroup O26 *E. coli* (n=404)

Isolate	O_type	H_type	MLST	Country	Source	*stx* profile
MOD1EC1684_Australia_human	O26	H11	ST-21	Australia	human	*stx1*
11KH245_Belgium_bovine	O26	H11	ST-21	Belgium	bovine	*stx1*
11KH263_Belgium_bovine	O26	H11	ST-21	Belgium	bovine	*stx1&stx2*
11KH63_Belgium_bovine	O26	H11	ST-21	Belgium	bovine	*stx1*
12KH23_Belgium_bovine	O26	H11	ST-21	Belgium	bovine	*stx1*
357S89_Belgium_bovine	O26	H11	ST-21	Belgium	bovine	*stx1*
631KH91_Belgium_bovine	O26	H11	ST-29	Belgium	bovine	*no_stx*
B43_Belgium_bovine	O26	H11	ST-1705	Belgium	bovine	*stx1*
B44_Belgium_bovine	O26	H11	ST-21	Belgium	bovine	*stx1*
EH031_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH182_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH193_Belgium_human	O26	H11	ST-21	Belgium	human	*stx2*
EH2035_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH2068_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH2075_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1&stx2*
EH2083_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH2208_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH2209_Belgium_human	O26	H11	ST-1705	Belgium	human	*stx1*
EH2219_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH2244_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH2251_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH2252_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH2257_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
EH2258_Belgium_human	O26	H11	ST-1705	Belgium	human	*stx1*
EH322_Belgium_human	O26	H11	ST-21	Belgium	human	*stx1*
CFSAN033951_Canada_human	O26	H11	ST-21	Canada	human	*stx1*
EC120246_Canada_human	O26	H11	ST-21	Canada	human	*stx1*
OLC0637_Canada_bovine	O26	H11	ST-21	Canada	bovine	*stx1*
AA044_Denmark_human	O26	H27	ST-10	Denmark	human	*no_stx*
02113_France_human	O26	H11	ST-21	France	human	*stx1*
03139_France_human	O26	H11	ST-1705	France	human	*stx1*
34870_France_human	O26	H11	ST-29	France	human	*stx2*
36084_France_human	O26	H11	ST-21	France	human	*stx2*
36293_France_human	O26	H11	ST-29	France	human	*stx2*
36348_France_human	O26	H11	ST-29	France	human	*stx2*

36493_France_human	O26	H11	ST-29	France	human	*stx2*
36708_France_human	O26	H11	ST-29	France	human	*stx2*
99109_France_human	O26	H11	ST-29	France	human	*stx2*
126814_Germany_human	O26	H11	ST-21	Germany	human	*stx2*
MOD1EC1664_Germany_human	O26	H11	ST-21	Germany	human	*stx1&stx2*
MOD1EC1763_Germany_bovine	O26	H11	ST-21	Germany	bovine	*stx1*
MOD1EC2812_Germany_human	O26	H11	ST-29	Germany	human	*stx2*
MOD1EC2813_Germany_human	O26	H11	ST-29	Germany	human	*stx2*
MOD1EC2814_Germany_human	O26	H11	ST-21	Germany	human	*stx1&stx2*
ED80_Italy_bovine	O26	H11	ST-29	Italy	bovine	*no_stx*
Aki01_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
Aki02_Japan_bovine	O26	H11	ST-21	Japan	bovine	*no_stx*
Aki06_Japan_bovine	O26	H11	ST-29	Japan	bovine	*no_stx*
Aki07_Japan_other~animal	O26	H11	ST-21	Japan	other animal	*stx1*
BK01_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
BK05_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
BK10_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
BK11_Japan_bovine	O26	H11	ST-21	Japan	bovine	*no_stx*
BK13_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
F02_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
F03_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
F06_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
F10_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
F17_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
F44_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M01_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M02_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M03_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M05_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M21_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M25_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M30_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M31_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M33_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
M41_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M43_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M48_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M53_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M56_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M64_Japan_human	O26	H11	ST-21	Japan	human	*stx1*

M65_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M73_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M79_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M80_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M81_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M89_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M90_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M92_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
M93_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
NIID070765_Japan_human	O26	H11	ST-21	Japan	human	*stx2*
NIID080884_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
NIID111609_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
NIID121840_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
NIID122147_Japan_human	O26	H11	ST-21	Japan	human	*stx2*
NIID122657_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
NIID122711_Japan_human	O26	H11	ST-29	Japan	human	*stx2*
NIID130549_Japan_human	O26	H11	ST-29	Japan	human	*stx2*
NIID132265_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
NIID132777_Japan_human	O26	H11	ST-29	Japan	human	*stx2*
NIID141423_Japan_human	O26	H11	ST-29	Japan	human	*stx2*
NIID141424_Japan_human	O26	H11	ST-21	Japan	human	*stx2*
NIID141425_Japan_human	O26	H11	ST-5172	Japan	human	*stx2*
O01_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O04_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O07_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O08_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O16_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O25_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O29_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O35_Japan_human	O26	H11	ST-21	Japan	human	*no_stx*
O46_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O55_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O56_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
O57_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O58_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O62_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
O64_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
She01_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
She03_Japan_bovine	O26	H11	ST-21	Japan	bovine	*no_stx*
T02_Japan_human	O26	H11	ST-21	Japan	human	*stx1*

T03_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T08_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T09_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T16_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T23_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T31_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T32_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
T33_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T35_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T39_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T41_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T43_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T45_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T49_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T53_Japan_human	O26	H11	ST-21	Japan	human	*stx1&stx2*
T59_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T61_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
T67_Japan_human	O26	H11	ST-21	Japan	human	*stx1*
YB01_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
YB02_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
YB03_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
YB05_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
YB06_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
YB13_Japan_bovine	O26	H11	ST-21	Japan	bovine	*stx1*
MOD1EC5336_Mexico_food	O26	H12	ST-10	Mexico	food	*no_stx*
SO174_Norway_human	O26	H11	ST-29	Norway	human	*stx2*
SO179_Norway_human	O26	H11	ST-29	Norway	human	*stx2*
100ST2_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
11ST_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
129ST2_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
14ER2056_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
14ER2240_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
14ER3837_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
15ER0146_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
15ER0512_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
15ER0857_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
15ER2552_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
15ER2837_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
15ER3334_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
15ER3567_NZ_human	O26	H11	ST-21	NZ	human	*stx1*

15ER3804_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
15ER4241_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER0243_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER0248_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER0520_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER0672_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER0946_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER1209_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER1432_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER1646_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER1892_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER1973_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
16ER2139_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
191ST2_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
22ST2_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
66ST1_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
A14a_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
A17a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
a185c_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
a234d_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
A46a_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
A65a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
A65b_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
A87b_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
AGR373_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
AGR72_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
e171b_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
e171c_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
e186a_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
E189b_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
e195b_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
e26a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
e37c_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
E46a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
E59b_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
ER005420_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
ER854674_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
ER985544_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
ERL023841_NZ_human	O26	H11	ST-21	NZ	human	*no_stx*
ERL071565_NZ_human	O26	H11	ST-21	NZ	human	*stx1*

ERL093655_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
ERL111686_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
ERL121992_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
H108a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
H113a_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
H132a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
H13ESR01843_NZ_human	O26	H11	ST-21	NZ	human	*stx1*
h148a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
h148b_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
h199a_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
h199c_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
H31c_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
Hide14d_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
ND29_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
ND3_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
ND35_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
ND62_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
Pre15a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
Pre4d_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
T11i_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
t126c_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
t128a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
t173a_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
t25a_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
t25c_NZ_bovine	O26	H11	ST-29	NZ	bovine	*no_stx*
T27a_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
T48b_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
T66c_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1113e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1122f_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1125e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1139e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1140e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1186e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1187e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1190e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1195e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1196e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1202e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1309e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*

VC1310e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1311e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1362e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1366e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1367e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1394e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1395e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1395f_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1395g_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1396e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1403e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC1471e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1471f_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1471g_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1471h_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1473e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1474e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC1486e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC396e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC397e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC401e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC452e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC456e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC459e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC473e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC474e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC474f_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC474g_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC474h_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC476e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC479e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC545e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC547e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC550e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC554e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC555e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC833e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC833f_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC833g_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC833h_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*

VC836e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC837e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC880e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC932f_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC936e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC940e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VC943e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC943f_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC943g_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC943h_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC946e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC951e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VC955e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VL0828h_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VL0958f_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VL1058g_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VL1277e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*no_stx*
VL2020e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
VL2228e_NZ_bovine	O26	H11	ST-21	NZ	bovine	*stx1*
10016_Poland_human	O26	H11	ST-29	Poland	human	*stx2*
TC6167_Switzerland_human	O26	H11	ST-29	Switzerland	human	*no_stx*
129381_UK_human	O26	H11	ST-21	UK	human	*stx1&stx2*
143493_UK_human	O26	H11	ST-21	UK	human	*stx1&stx2*
143934_UK_human	O26	H11	ST-21	UK	human	*stx1*
146047_UK_human	O26	H11	ST-21	UK	human	*stx1*
152422_UK_human	O26	H11	ST-21	UK	human	*stx1*
154820_UK_human	O26	H11	ST-21	UK	human	*stx1&stx2*
18109_UK_human	O26	H11	ST-21	UK	human	*stx1&stx2*
211644_UK_animal	O26	H11	ST-29	UK	other animal	*no_stx*
224098_UK_human	O26	H11	ST-21	UK	human	*stx1*
25910_UK_human	O26	H11	ST-21	UK	human	*stx2*
2M8BS8_UK_human	O26	H11	ST-21	UK	human	*stx1*
46009_UK_human	O26	H11	ST-21	UK	human	*stx1&stx2*
46710_UK_human	O26	H11	ST-21	UK	human	*stx2*
48212_UK_human	O26	H11	ST-21	UK	human	*stx2*
51911_UK_human	O26	H11	ST-21	UK	human	*stx1&stx2*
60510_UK_human	O26	H11	ST-21	UK	human	*stx1&stx2*
62612_UK_human	O26	H11	ST-21	UK	human	*stx2*
63713_UK_human	O26	H11	ST-21	UK	human	*stx2*
67013_UK_human	O26	H11	ST-29	UK	human	*no_stx*

68013_UK_human	O26	H11	ST-29	UK	human	*no_stx*
93279_UK_human	O26	H11	ST-21	UK	human	*stx1&stx2*
93280_UK_human	O26	H11	ST-21	UK	human	*stx2*
93285_UK_human	O26	H11	ST-21	UK	human	*stx2*
93304_UK_human	O26	H11	ST-21	UK	human	*stx1*
ECO0283_UK_humansepsis	O26	H27	ST-10	UK	human	*no_stx*
H19_UK_human	O26	H11	ST-21	UK	human	*stx1*
HKCVSH_UK_human	O26	H11	ST-21	UK	human	*stx1*
I20VK7_UK_human	O26	H11	ST-21	UK	human	*stx1*
KLAV92_UK_human	O26	H11	ST-21	UK	human	*stx1*
2009C3689_USA_human	O26	H11	ST-29	USA	human	*stx2*
2009C3996_USA_human	O26	H11	ST-21	USA	human	*stx1*
2009C4747_USA_human	O26	H11	ST-21	USA	human	*stx1*
2010C3051_USA_human	O26	H11	ST-21	USA	human	*stx1*
2010C3472_USA_human	O26	H11	ST-21	USA	human	*stx1*
2010C3902_USA_human	O26	H11	ST-21	USA	human	*stx1&stx2*
2010C4430_USA_human	O26	H11	ST-21	USA	human	*stx1*
2010C4800_USA_human	O26	H11	ST-29	USA	human	*no_stx*
2011C3270_USA_human	O26	H11	ST-21	USA	human	*stx1*
2011C3506_USA_human	O26	H11	ST-21	USA	human	*stx1&stx2*
2012C3101_USA_human	O26	H11	ST-21	USA	human	*stx1*
2012C3102_USA_human	O26	H11	ST-21	USA	human	*stx1*
2012C3912_USA_human	O26	H11	ST-21	USA	human	*stx1*
2012C4606_USA_human	O26	H11	ST-21	USA	human	*stx1*
2015C5206_USA_human	O26	H11	ST-21	USA	human	*stx1*
2016C3018_USA_human	O26	H11	ST-21	USA	human	*stx1&stx2*
643464_USA_human	O26	H11	ST-21	USA	human	*stx1*
CFSAN046724_USA_food	O26	H11	ST-21	USA	food	*stx1*
CFSAN046736_USA_food	O26	H11	ST-21	USA	food	*stx1*
ewgs1005_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FDA00010430_USA_food	O26	H11	ST-21	USA	food	*stx1*
FSIS11704781_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1400371_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1500489_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1500632_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1500641_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1500781_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1500873_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1500874_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1500928_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*

FSIS1500948_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1503255_USA_bovine	O26	H11	ST-29	USA	bovine	*stx2*
FSIS1503256_USA_bovine	O26	H11	ST-29	USA	bovine	*stx2*
FSIS1503257_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1503454_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1606391_USA_bovine	O26	H11	ST-29	USA	bovine	*stx2*
FSIS1608722_USA_bovine	O26	H11	ST-29	USA	bovine	*stx2*
FSIS1608854_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1609416_USA_bovine	O26	H11	ST-5172	USA	bovine	*stx2*
FSIS1701668_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS1710186_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
FSIS21720313_USA_bovine	O26	H11	ST-5172	USA	bovine	*stx2*
MOD1EC1750_USA_human	O26	H11	ST-21	USA	human	*stx1*
MOD1EC1753_USA_human	O26	H11	ST-21	USA	human	*stx1*
MOD1EC1919_USA_human	O26	H11	ST-21	USA	human	*stx1*
MOD1EC2790_USA_human	O26	H11	ST-21	USA	human	*stx1&stx2*
MOD1EC550_USA_human	O26	H11	ST-21	USA	human	*stx1*
MOD1EC5703_USA_avian	O26	H11	ST-29	USA	other animal	*no_stx*
MOD1EC6029_USA_food	O26	H11	ST-21	USA	food	*stx1*
MOD1EC6201_USA_other~animal	O26	H6	ST-6236	USA	other animal	*no_stx*
PNUSAE000002_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE000133_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE000539_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE000779_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE000800_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE000885_USA_human	O26	H11	ST-21	USA	human	*stx1&stx2*
PNUSAE000887_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE001154_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE001373_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE001379_USA_human	O26	H11	ST-29	USA	human	*stx2*
PNUSAE001573_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE001578_USA_human	O26	H11	ST-29	USA	human	*stx2*
PNUSAE002149_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE002166_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE002228_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE002615_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE003211_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE003275_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE003398_USA_human	O26	H11	ST-21	USA	human	*stx1*
PNUSAE008468_USA_human	O26	H11	ST-21	USA	human	*stx1*

PNUSAE011184_USA_human	O26	H11	ST-21	USA	human	*stx1*
TC3486_USA_bovine	O26	H11	ST-29	USA	bovine	*no_stx*
TC3630_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
TC3656_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
TC4219_USA_bovine	O26	H11	ST-21	USA	bovine	*no_stx*
TC6165_USA_human	O26	H11	ST-21	USA	human	*stx1*
TC6168_USA_human	O26	H11	ST-21	USA	human	*stx1*
TC6169_USA_bovine	O26	H11	ST-21	USA	bovine	*stx1*
WAPHLECOA00006_USA_caprine	O26	H11	ST-21	USA	other animal	*stx1*

Appendix Table 8: Antibiotic resistance gene class (n=8) by serogroup O26 *E. coli* bacterial isolate (n=404)

Isolate	Aminoglycoside	Beta-lactam	Macrolide	Phenicol	Quinolone	Sulphonamide	Tetracycline	Trimethoprim
02113_France_human								
03139_France_human	Y	Y		Y		Y	Y	
10016_Poland_human								
100ST2_NZ_bovine							Y	
11KH245_Belgium_bovine	Y	Y		Y		Y	Y	Y
11KH263_Belgium_bovine	Y					Y		
11KH63_Belgium_bovine	Y					Y		
11ST_NZ_bovine	Y							
126814_Germany_human								
129381_UK_human								
129ST2_NZ_bovine							Y	
12KH23_Belgium_bovine	Y					Y	Y	
143493_UK_human								
143934_UK_human								
146047_UK_human								
14ER2056_NZ_human	Y					Y	Y	
14ER2240_NZ_human								
14ER3837_NZ_human								
152422_UK_human								
154820_UK_human								
15ER0146_NZ_human								
15ER0512_NZ_human								
15ER0857_NZ_human								
15ER2552_NZ_human	Y					Y	Y	
15ER2837_NZ_human								
15ER3334_NZ_human								

15ER3567_NZ_human								
15ER3804_NZ_human								
15ER4241_NZ_human								
16ER0243_NZ_human								
16ER0248_NZ_human								
16ER0520_NZ_human								
16ER0672_NZ_human	Y					Y	Y	
16ER0946_NZ_human								
16ER1209_NZ_human								
16ER1432_NZ_human								
16ER1646_NZ_human	Y							
16ER1892_NZ_human								
16ER1973_NZ_human								
16ER2139_NZ_human	Y	Y				Y		Y
18109_UK_human								
191ST2_NZ_bovine								
2009C3689_USA_human								
2009C3996_USA_human								
2009C4747_USA_human								
2010C3051_USA_human								
2010C3472_USA_human								
2010C3902_USA_human								
2010C4430_USA_human								
2010C4800_USA_human							Y	
2011C3270_USA_human								
2011C3506_USA_human								
2012C3101_USA_human								
2012C3102_USA_human		Y						
2012C3912_USA_human								

2012C4606_USA_human								
2015C5206_USA_human								
2016C3018_USA_human								
211644_UK_animal								
224098_UK_human								
22ST2_NZ_bovine	Y					Y	Y	
25910_UK_human	Y	Y				Y	Y	
2M8BS8_UK_human								
34870_France_human								
357S89_Belgium_bovine								
36084_France_human								
36293_France_human								
36348_France_human								
36493_France_human	Y	Y				Y		
36708_France_human			Y					
46009_UK_human								
46710_UK_human								
48212_UK_human	Y	Y				Y	Y	
51911_UK_human								
60510_UK_human								
62612_UK_human	Y	Y				Y	Y	
631KH91_Belgium_bovine							Y	
63713_UK_human								
643464_USA_human								
66ST1_NZ_bovine	Y							
67013_UK_human								
68013_UK_human								
93279_UK_human	Y	Y				Y		
93280_UK_human								

93285_UK_human								
93304_UK_human	Y					Y	Y	
99109_France_human								
A14a_NZ_bovine								
A17a_NZ_bovine								
A46a_NZ_bovine								
A65a_NZ_bovine	Y							
A65b_NZ_bovine	Y							
A87b_NZ_bovine								
AA044_Denmark_human							Y	
AGR373_NZ_bovine	Y							
AGR72_NZ_bovine								
Aki01_Japan_bovine								
Aki02_Japan_bovine	Y					Y		
Aki06_Japan_bovine								
Aki07_Japan_other~animal								
B43_Belgium_bovine	Y	Y		Y		Y	Y	Y
B44_Belgium_bovine	Y		Y			Y		
BK01_Japan_bovine		Y						
BK05_Japan_bovine								
BK10_Japan_bovine								
BK11_Japan_bovine								
BK13_Japan_bovine								
CFSAN033951_Canada_human							Y	
CFSAN046724_USA_food								
CFSAN046736_USA_food								
E189b_NZ_bovine								
E46a_NZ_bovine								
E59b_NZ_bovine							Y	

EC120246_Canada_human							Y	
ECO0283_UK_humansepsis	Y	Y		Y		Y	Y	Y
ED80_Italy_bovine			Y					
EH031_Belgium_human								
EH182_Belgium_human	Y			Y		Y	Y	Y
EH193_Belgium_human	Y					Y		
EH2035_Belgium_human	Y	Y				Y	Y	
EH2068_Belgium_human	Y					Y		
EH2075_Belgium_human								
EH2083_Belgium_human	Y	Y				Y		
EH2208_Belgium_human							Y	
EH2209_Belgium_human								
EH2219_Belgium_human	Y	Y		Y		Y	Y	Y
EH2244_Belgium_human	Y					Y		
EH2251_Belgium_human	Y					Y		
EH2252_Belgium_human	Y					Y	Y	
EH2257_Belgium_human								
EH2258_Belgium_human								
EH322_Belgium_human	Y	Y				Y	Y	
ER005420_NZ_human								
ER854674_NZ_human								
ER985544_NZ_human								
ERL023841_NZ_human	Y					Y	Y	
ERL071565_NZ_human								
ERL093655_NZ_human	Y							
ERL111686_NZ_human								
ERL121992_NZ_human	Y					Y		
F02_Japan_human								
F03_Japan_human								

F06_Japan_human								
F10_Japan_human								
F17_Japan_human								
F44_Japan_human								
FDA00010430_USA_food								
FSIS11704781_USA_bovine								
FSIS1400371_USA_bovine	Y	Y					Y	
FSIS1500489_USA_bovine								
FSIS1500632_USA_bovine								
FSIS1500641_USA_bovine								
FSIS1500781_USA_bovine								
FSIS1500873_USA_bovine								
FSIS1500874_USA_bovine								
FSIS1500928_USA_bovine								
FSIS1500948_USA_bovine	Y	Y		Y		Y	Y	Y
FSIS1503255_USA_bovine								
FSIS1503256_USA_bovine								
FSIS1503257_USA_bovine								
FSIS1503454_USA_bovine								
FSIS1606391_USA_bovine	Y			Y		Y	Y	
FSIS1608722_USA_bovine	Y	Y		Y		Y	Y	
FSIS1608854_USA_bovine								
FSIS1609416_USA_bovine								
FSIS1701668_USA_bovine								
FSIS1710186_USA_bovine								
FSIS21720313_USA_bovine								
H108a_NZ_bovine	Y							
H113a_NZ_bovine	Y							
H132a_NZ_bovine								

H13ESR01843_NZ_human	Y					Y	Y	
H19_UK_human	Y					Y	Y	
H31c_NZ_bovine								
HKCVSH_UK_human								
Hide14d_NZ_bovine	Y							
I20VK7_UK_human								
KLAV92_UK_human								
M01_Japan_human								
M02_Japan_human	Y					Y	Y	
M03_Japan_human								
M05_Japan_human								
M21_Japan_human	Y					Y	Y	
M25_Japan_human								
M30_Japan_human	Y					Y	Y	
M31_Japan_human								
M33_Japan_human								
M41_Japan_human								
M43_Japan_human	Y	Y				Y	Y	
M48_Japan_human	Y	Y				Y		
M53_Japan_human								
M56_Japan_human	Y	Y				Y		
M64_Japan_human	Y	Y				Y		
M65_Japan_human	Y					Y	Y	
M73_Japan_human								
M79_Japan_human	Y					Y	Y	
M80_Japan_human	Y	Y				Y		
M81_Japan_human	Y	Y		Y		Y	Y	Y
M89_Japan_human								
M90_Japan_human	Y					Y	Y	

M92_Japan_human	Y					Y	Y	
M93_Japan_human	Y					Y		
MOD1EC1664_Germany_human	Y					Y		
MOD1EC1684_Australia_human	Y			Y		Y		
MOD1EC1750_USA_human								
MOD1EC1753_USA_human	Y	Y				Y	Y	Y
MOD1EC1763_Germany_bovine	Y					Y	Y	
MOD1EC1919_USA_human	Y	Y				Y	Y	
MOD1EC2790_USA_human	Y	Y				Y	Y	
MOD1EC2812_Germany_human								
MOD1EC2813_Germany_human								
MOD1EC2814_Germany_human	Y	Y				Y		
MOD1EC5336_Mexico_food	Y	Y			Y	Y	Y	Y
MOD1EC550_USA_human	Y					Y	Y	
MOD1EC5703_USA_avian								
MOD1EC6029_USA_food								
MOD1EC6201_USA_other~animal								
ND29_NZ_bovine	Y					Y	Y	
ND35_NZ_bovine	Y					Y	Y	
ND3_NZ_bovine	Y					Y	Y	
ND62_NZ_bovine	Y					Y	Y	
NIID070765_Japan_human								
NIID080884_Japan_human	Y					Y	Y	
NIID111609_Japan_human	Y	Y				Y	Y	
NIID121840_Japan_human								
NIID122147_Japan_human							Y	
NIID122657_Japan_human	Y					Y		
NIID122711_Japan_human								
NIID130549_Japan_human	Y			Y		Y	Y	

NIID132265_Japan_human								
NIID132777_Japan_human								
NIID141423_Japan_human	Y			Y		Y	Y	
NIID141424_Japan_human								
NIID141425_Japan_human								
O01_Japan_human								
O04_Japan_human								
O07_Japan_human	Y					Y	Y	
O08_Japan_human	Y					Y	Y	
O16_Japan_human								
O25_Japan_human								
O29_Japan_human								
O35_Japan_human								
O46_Japan_human								
O55_Japan_human								
O56_Japan_human								
O57_Japan_human								
O58_Japan_human								
O62_Japan_human								
O64_Japan_human	Y	Y		Y	Y	Y	Y	Y
OLC0637_Canada_bovine	Y		Y			Y	Y	
PNUSAE000002_USA_human								
PNUSAE000133_USA_human								
PNUSAE000539_USA_human								
PNUSAE000779_USA_human								
PNUSAE000800_USA_human								
PNUSAE000885_USA_human								
PNUSAE000887_USA_human								
PNUSAE001154_USA_human								

PNUSAE001373_USA_human								
PNUSAE001379_USA_human								
PNUSAE001573_USA_human								
PNUSAE001578_USA_human								
PNUSAE002149_USA_human								
PNUSAE002166_USA_human								
PNUSAE002228_USA_human								
PNUSAE002615_USA_human								
PNUSAE003211_USA_human								
PNUSAE003275_USA_human								
PNUSAE003398_USA_human								
PNUSAE008468_USA_human								
PNUSAE011184_USA_human		Y						
Pre15a_NZ_bovine	Y							
Pre4d_NZ_bovine	Y					Y	Y	
SO174_Norway_human	Y							
SO179_Norway_human								
She01_Japan_bovine								
She03_Japan_bovine	Y	Y				Y	Y	
T02_Japan_human								
T03_Japan_human	Y					Y	Y	
T08_Japan_human								
T09_Japan_human								
T11i_NZ_bovine	Y							
T16_Japan_human	Y	Y				Y	Y	
T23_Japan_human								
T27a_NZ_bovine	Y					Y		
T31_Japan_human								
T32_Japan_human								

T33_Japan_human	Y					Y	Y	
T35_Japan_human								
T39_Japan_human								
T41_Japan_human								
T43_Japan_human								
T45_Japan_human								
T48b_NZ_bovine	Y					Y		
T49_Japan_human								
T53_Japan_human								
T59_Japan_human								
T61_Japan_human								
T66c_NZ_bovine								
T67_Japan_human	Y					Y	Y	
TC3486_USA_bovine	Y	Y				Y	Y	
TC3630_USA_bovine								
TC3656_USA_bovine								
TC4219_USA_bovine	Y	Y				Y	Y	
TC6165_USA_human								
TC6167_Switzerland_human								
TC6168_USA_human								
TC6169_USA_bovine								
VC1113e_NZ_bovine								
VC1122f_NZ_bovine								
VC1125e_NZ_bovine								
VC1139e_NZ_bovine	Y							
VC1140e_NZ_bovine	Y							
VC1186e_NZ_bovine								
VC1187e_NZ_bovine								
VC1190e_NZ_bovine								

VC1195e_NZ_bovine								
VC1196e_NZ_bovine								
VC1202e_NZ_bovine								
VC1309e_NZ_bovine								
VC1310e_NZ_bovine								
VC1311e_NZ_bovine								
VC1362e_NZ_bovine								
VC1366e_NZ_bovine								
VC1367e_NZ_bovine								
VC1394e_NZ_bovine								
VC1395e_NZ_bovine								
VC1395f_NZ_bovine								
VC1395g_NZ_bovine								
VC1396e_NZ_bovine								
VC1403e_NZ_bovine								
VC1471e_NZ_bovine								
VC1471f_NZ_bovine								
VC1471g_NZ_bovine								
VC1471h_NZ_bovine								
VC1473e_NZ_bovine								
VC1474e_NZ_bovine								
VC1486e_NZ_bovine								
VC396e_NZ_bovine								
VC397e_NZ_bovine								
VC401e_NZ_bovine								
VC452e_NZ_bovine								
VC456e_NZ_bovine								
VC459e_NZ_bovine								
VC473e_NZ_bovine								

VC474e_NZ_bovine								
VC474f_NZ_bovine								
VC474g_NZ_bovine								
VC474h_NZ_bovine								
VC476e_NZ_bovine								
VC479e_NZ_bovine								
VC545e_NZ_bovine								
VC547e_NZ_bovine								
VC550e_NZ_bovine								
VC554e_NZ_bovine								
VC555e_NZ_bovine								
VC833e_NZ_bovine	Y							
VC833f_NZ_bovine	Y							
VC833g_NZ_bovine	Y							
VC833h_NZ_bovine	Y							
VC836e_NZ_bovine	Y							
VC837e_NZ_bovine	Y							
VC880e_NZ_bovine								
VC932f_NZ_bovine								
VC936e_NZ_bovine								
VC940e_NZ_bovine								
VC943e_NZ_bovine								
VC943f_NZ_bovine								
VC943g_NZ_bovine								
VC943h_NZ_bovine								
VC946e_NZ_bovine								
VC951e_NZ_bovine								
VC955e_NZ_bovine								
VL0828h_NZ_bovine	Y					Y	Y	

Sample								
VL0958f_NZ_bovine								
VL1058g_NZ_bovine	Y					Y	Y	
VL1277e_NZ_bovine	Y					Y	Y	
VL2020e_NZ_bovine								
VL2228e_NZ_bovine								
WAPHLECOA00006_USA_caprine								
YB01_Japan_bovine	Y					Y		
YB02_Japan_bovine								
YB03_Japan_bovine	Y	Y		Y		Y	Y	Y
YB05_Japan_bovine			Y					
YB06_Japan_bovine								
YB13_Japan_bovine				Y				
a185c_NZ_bovine			Y					
a234d_NZ_bovine							Y	
e171b_NZ_bovine								
e171c_NZ_bovine								
e186a_NZ_bovine								
e195b_NZ_bovine								
e26a_NZ_bovine								
e37c_NZ_bovine								
ewgs1005_USA_bovine								
h148a_NZ_bovine								
h148b_NZ_bovine								
h199a_NZ_bovine								
h199c_NZ_bovine								
t126c_NZ_bovine								
t128a_NZ_bovine	Y							
t173a_NZ_bovine								
t25a_NZ_bovine								

t25c_NZ_bovine								

Aminoglycoside genes detected: strA_1, strA_4, strB_1, aac(3)-IId_1, aac(3)-IVa_1, aadA1_1, aadA1_3, aadA2_2,   aph(3')-IIa_1, aph(3')-Ia_1, aph(3')-Ic_1, aph(4)-Ia_1. aph(6)-Ic_1

Beta-lactam genes detected: blaCMY-2_1, blaOXA-1_1, blaTEM-104_1, blaTEM-1A_4, blaTEM-1B_1, blaTEM-1C_5

Macrolide genes detected: erm(B)_18, lnu(F)_1, mph(A)_1, mph(A)_2, mph(B)_1

Phenicol genes detected: floR_2, catA1_1, catB2_1, catB8_1, cmlA1_1

Quinolone genes detected: oqxA_1   o, qxB_1, QnrS1_1

Sulphonamide genes detected: sul1_2, sul2_11, sul2_14, sul2_2, sul2_3, sul3_2

Tetracycline genes detected: tet(34)_1, tet(A)_4, tet(B)_3, tet(B)_4, tet(C)_5

Trimethoprim genes detected: dfrA12_1, dfrA14_1, dfrA1_1, dfrA8_1

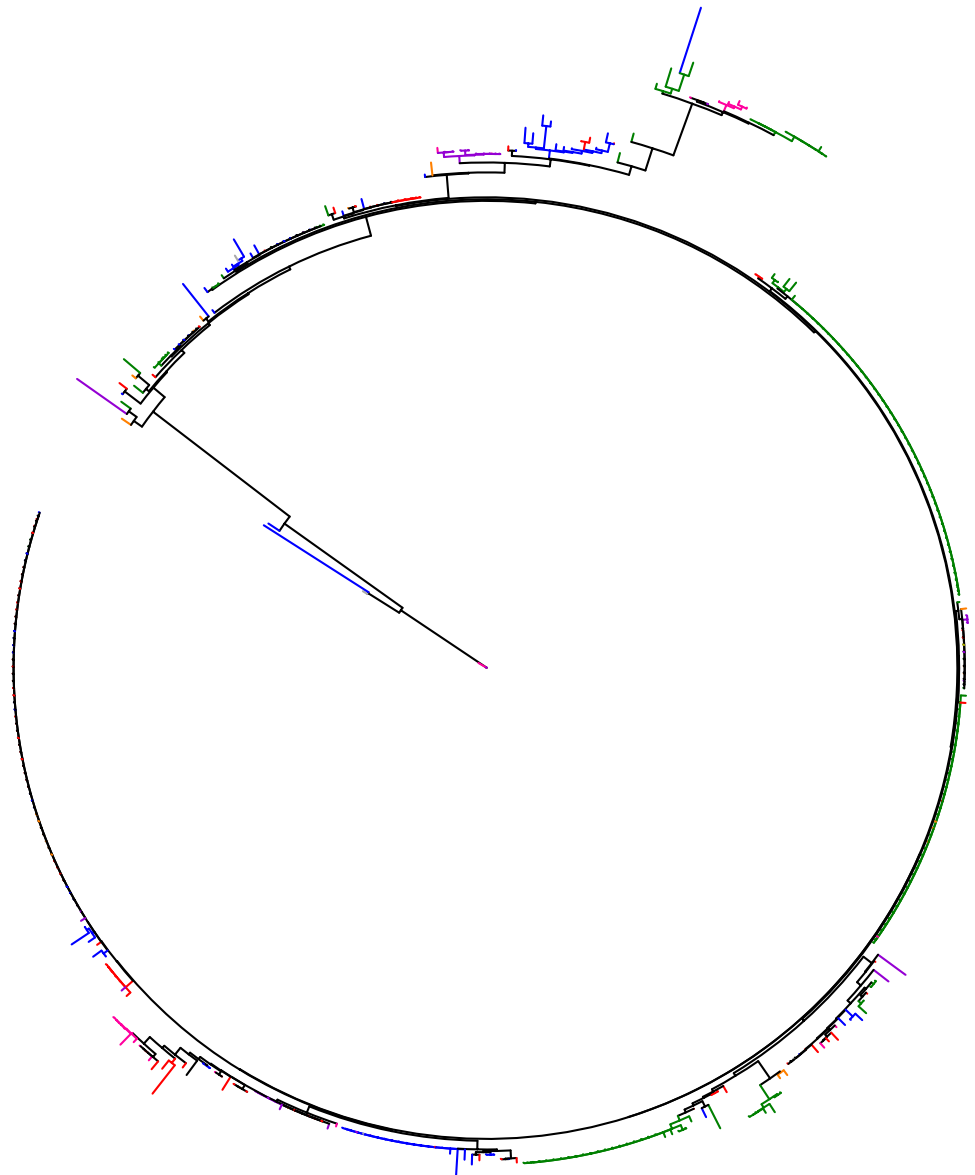Appendix Figure 1: Heatmap of SNP distances between *E. coli* serogroup O26 isolates (n=66); GREEN indicates larger SNP distances (dissimilar isolates), while RED indicates smaller SNP distances (very similar isolates)
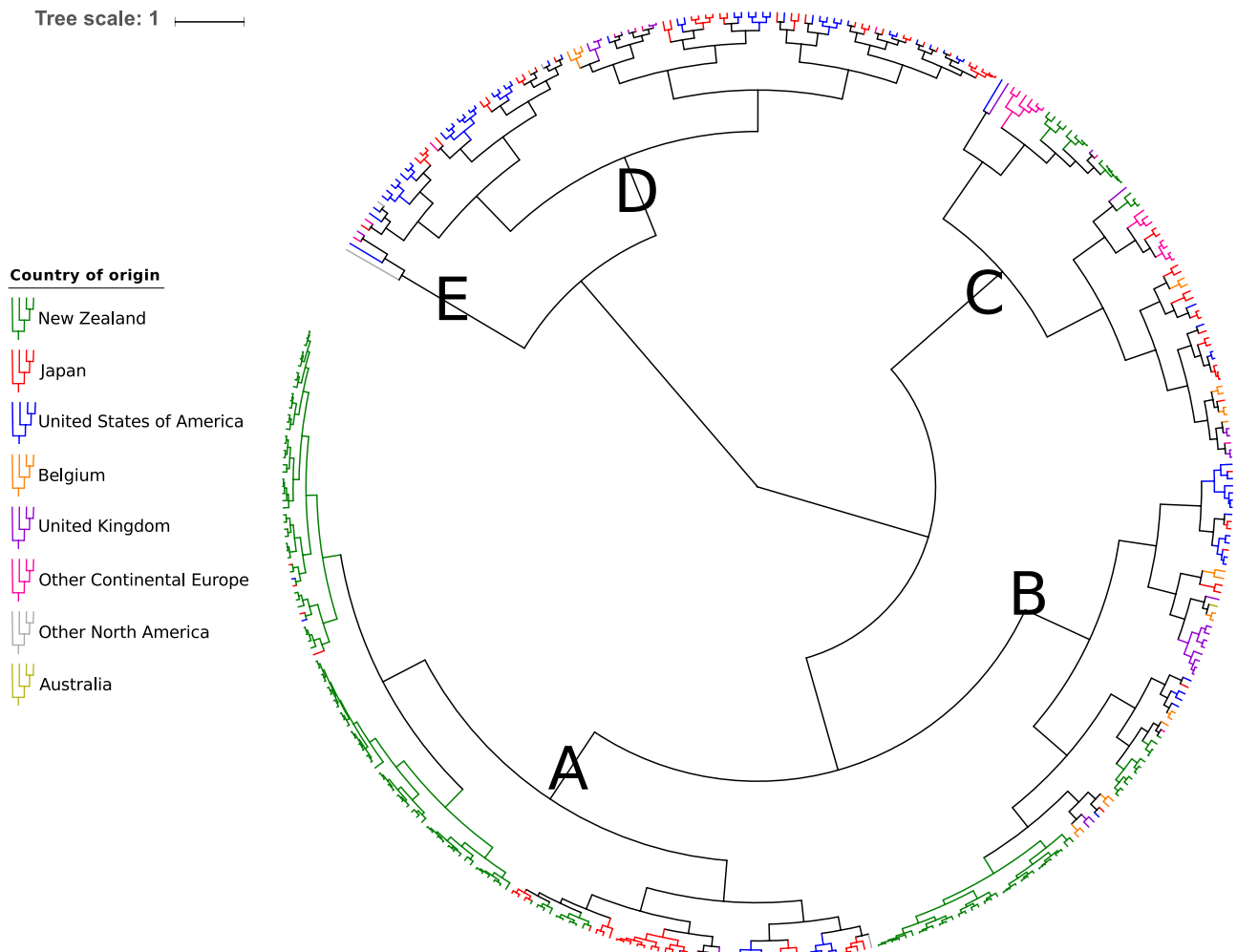
Appendix Figure 2. Maximum likelihood real branch length core and accessory genome trees of *E. coli* serogroup O26 isolates (n=404), annotated with country

Tree scale: 10

Country of origin
New Zealand
Japan
United States of America
Belgium
United Kingdom
Other Continental Europe
Other North America
Australia

Appendix Figure 3. Neighbor-joining tree of virulence genes (n=192) of serogroup O26 isolates (n=404) with real branch lengths annotated by country
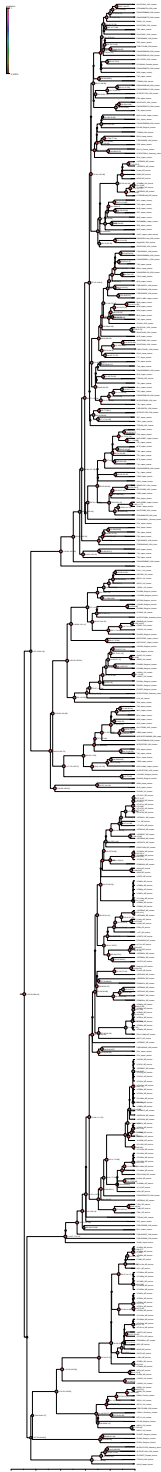
Appendix Figure 4: Maximum clade credibility tree of time of most recent common ancestor (TMRCA) analysis of *E. coli* serogroup O26 sequence type 21 (ST-21) isolates (n=344), annotated with posterior probability of tree branches. All convergence dates are annotated with 95% HPD intervals from the age of the newest isolate (2017.5 in decimal years)
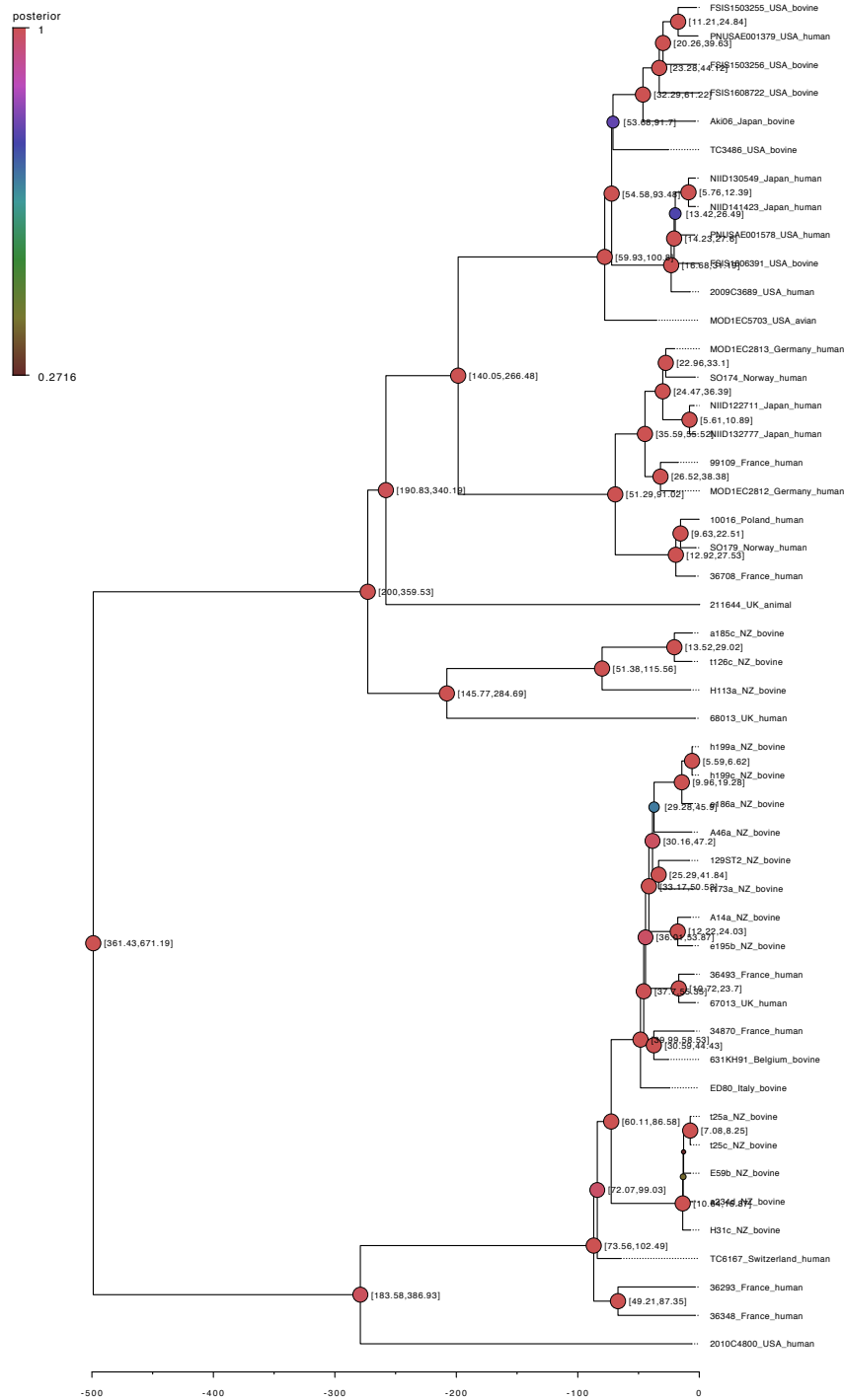
Appendix Figure 5: Maximum clade credibility tree of time of most recent common ancestor (TMRCA) analysis of *E. coli* serogroup O26 sequence type 29 (ST-29) isolates (n=48), annotated with posterior probability of tree branches. All convergence dates are annotated with 95% HPD intervals from the age of the newest isolate (2017.0411 in decimal years)