The Spatiotemporal Occurrence and Recovery of Shiga Toxin-producing *Escherichia coli* (STEC) in Well-sourced Drinking Water from Southern Alberta, Canada

by

Colin Reynolds

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

Master of Science

in

Environmental Health Sciences

School of Public Health

University of Alberta

Abstract

Shiga toxin-producing *Escherichia coli* (STEC) are important enteric pathogens that cause significant morbidity and mortality worldwide, as well as in Canada. Drinking water is a key exposure pathway for STEC infection and the occurrence of these pathogens in drinking water sources represents an important knowledge gap in the epidemiological understanding of STEC-associated illness, particularly with regards to non-O157 STEC serogroups as well as unregulated well-sourced drinking water systems. The purpose of this study was to investigate the frequency and spatiotemporal patterns of STEC occurrence in non-municipal well-sourced drinking water in the southern region of the province of Alberta, Canada. Using Shiga toxin gene quantitative polymerase chain reaction analysis (*stx* qPCR) as a proxy, STEC were found in 8.0% of *E. coli* positive drinking water samples and 0.2% of all drinking water samples from across southern Alberta submitted to the Alberta Provincial Laboratory for Public Health (ProvLab) between 2004-2016. A statistically significant seasonal pattern of STEC occurrence that peaked in the summer months was detected, which coincides with well-established seasonal increases of reported STEC cases in both Alberta and Canada. Significantly greater annual STEC occurrence corresponded with years where extreme precipitation events occurred across this region and a STEC contamination event in an area of southern Alberta was detected in the early summer of 2005 using Kulldorff scan statistics. Twenty-one distinct STEC serotypes, including 6 of the 7 most clinically

relevant serotypes in Alberta, were recovered from *stx* positive drinking water samples using a STEC-specific chromogenic growth media. Taken together, these findings demonstrate that STEC was consistently found in non-municipal well-sourced drinking water across southern Alberta and represents a public health risk to Canadians that rely on groundwater wells for their drinking water supplies. Furthermore, the seasonal alignment of increased STEC occurrence in non-municipal drinking water sources with patterns of increased STEC-associated illness highlights both the possibility of contaminated well water contributing to the seasonality of STEC infections in Canada, as well as the potential contribution of this exposure pathway to the overall burden of enteric disease.

Preface

Some of the research conducted for this thesis forms part of a research collaboration, led by Dr. Sylvia Checkley at the University of Calgary, with Dr. Norman Neumann being the lead collaborator at the University of Alberta. The research project, of which this thesis is a part, received research ethics approval from the University of Calgary Conjoint Faculties Research Ethics Board, Title: Assessing Water Quality, Microbial Risks and Waterborne Pathogens in Rural Alberta using a One Health Framework, ID: REB14-1474; and from the University of Alberta Research Ethics Board, Title: Assessing Water Quality, Microbial Risks and Waterborne Pathogens in Rural Alberta using a One Health Framework, ID: Pro00053656.

Acknowledgements

I would like to sincerely thank my supervisor Dr. Norman Neumann, my co-supervisor Dr. Linda Chui, and my committee members Dr. Sylvia Checkley and Dr. Byeonghwa Jeon for their invaluable wisdom and guidance throughout the entirety of this process.

I would also like to acknowledge the important contributions of Candis Scott and Dr. Graham Banting in my laboratory training and the countless other assistances they provided during the completion of the work for this thesis. I would like to thank Kelsey Meyer for her collaboration throughout this project, Shannon Braithwaite and the Environmental Waters team at ProvLab Edmonton, Jesse Invik for his help with GIS data transformation, Lorraine Ingham and Nancy Yuen from ProvLab Calgary, Dr. Roger Johnson and Kim Ziebell from the National Microbiology Laboratory at Guelph, Jacqueline Noga, as well as Sydney Rudko, Michelle Gordy and the rest of Dr. Patrick Hanington's Lab.

I am very grateful to the many friends and colleagues I had the privilege of getting to know during my time at the School of Public Health and all the help, laughter, and encouragement they provided. My graduate school experience was so much richer due to their involvement.

Truly, this would not have been possible without the patience, love, and support of my parents, Bill and Marion, my brother Vince and his wonderful family, and my girlfriend Jane. Thank you.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ABI – Applied Biosystems

ATCC – American Type Culture Collection

ATS – Alberta Township Survey System

BSA – Bovine Serum Albumin

CA-STEC – CHROMagar™ STEC

CI – Confidence Interval

CLRW – Clinical Laboratory Reagent Water

C_t – Cycle Threshold

DIAL –Data Integration for Alberta Laboratories Tool

DNA – Deoxyribonucleic Acid

E. coli – *Escherichia coli*

g – The Earth's gravitational force

HPLC – High Performance Liquid Chromatography

IAC – Internal Amplification Control

IRR – Incidence Rate Ratio

ISO – International Organization for Standardization

LoD – 95% Level of Detection

MAC – MacConkey

NTC – No Template Control

PCR – Polymerase Chain-reaction

pCR2.1 – Plasmid Vector pCR2.1

ProvLab – Alberta Provincial Laboratory for Public Health

QMRA – Quantitative Microbial Risk Assessment

qPCR – Quantitative Polymerase Chain-reaction

rpm – Revolutions per minute

stx – Shiga toxin gene aka. Verotoxin gene (*vtx*)

stx1 – Shiga toxin type 1 gene, aka. Verotoxin type 1 gene (*vtx1*)

stx2 – Shiga toxin type 2 gene, aka. Verotoxin type 2 gene (*vtx2*)

STEC – Shigatoxigenic or Shiga Toxin-Producing *E. coli*, aka Verotoxigenic or Verotoxin-Producing *E. coli* (VTEC)

TSA – Tryptic Soy Agar

TSB – Tryptic Soy Broth

UV – Ultraviolet

U of A – University of Alberta

CHAPTER 1: Introduction

1.1 Background

Escherichia coli is a rod-shaped, gram-negative bacterium from the *Enterobacteriaceae* family that is largely a commensal resident in the gastrointestinal tract of warm-blooded animals [1]. Through the loss and gain of various genes, commensal *E. coli* can become pathogenic. Human pathogenic *E. coli* have been assigned to six major diarrheagenic pathotypes: enteropathogenic *E. coli* (EPEC), *Shigella*/enteroinvasive *E. coli* (EIEC), enteroaggregative *E. coli* (EAEC), diffusely adherent *E. coli* (DAEC), enterotoxigenic *E. coli* (ETEC), and Shiga toxin-producing *E. coli* (STEC), which includes the enterohemorhaggic *E. coli* (EHEC) subset. Strains within each pathotype are genetically diverse and are often identified though the Kauffman classification scheme of serotyping somatic O polysaccharides (i.e., O157) and flagellar H antigens (i.e., H7) [2,3].

STEC are defined by the presence of Shiga toxin type 1 (*stx1*) genes, Shiga toxin type 2 (*stx2*) genes, or both, within an isolate[1]. Shiga toxins are a family of genetically and functionally related cytotoxins and are the definitive virulence factors for the STEC pathogroup [4]. Both the *stx1* and *stx2* genes originate from lamboid prophages that can be integrated into the *E. coli* chromosome [5,6]. While both *stx1* and *stx2* can cause disease, *stx2* carrying strains are generally associated

with more serious disease, particularly with the development of hemolytic uremic syndrome (HUS), the most serious sequelae associated with STEC infection that is characterized by renal failure, hemolytic anemia, and thrombocytopenia in infected patients [1,7–9]. Within each Shiga toxin type there are also multiple allelic subtypes (1a, 1c, 1d, 2a, 2b, 2c, 2d, 2e, 2f, 2g) that have varying pathogenic potentials [1,10–12]. Importantly, since *stx* genes are encoded on lysogenic bacteriophages, these genes can undergo horizontal gene transfer to *E. coli* strains (or other bacteria) not previously associated with *stx* genes and potentially give rise to new unexpected STEC strains [6,13,14]. The highly pathogenic *E. coli* O104:H4 strain responsible for a 2011 German STEC outbreak that caused nearly 4000 cases of illness, was able to acquire the virulence factors of both an EAEC (the ability to form an aggregative adherence pattern) as well as an STEC (Shiga toxin type 2) through horizontal gene transfer and serves as an example of the pathogenic potential associated with the uptake of *stx* genes by other *E. coli* [15–17]. There are a number of other important virulence factors that STEC serotypes often incorporate into their genome in addition to the Shiga toxins, of which two particularly notable factors include: i) the locus of enterocyte effacement (LEE) that encodes a variety of genes involved in attachment and delivery of toxic effects, and ii) EHEC-hemolysin (*hlxA*), a pore-forming toxin [1,18]. The virulence gene *eae*, which codes for the adhesion protein intimin and is located on the LEE pathogenicity island, and *hlxA* are often investigated alongside the *stx* genes when determining pathogenicity of an *E. coli* strain. The myriad of virulence factor combinations found in STEC strains play a defining role in determining the pathogenicity of each specific strain, the more

pathogenic of which can have very low infectious doses compared to other *E. coli* pathotypes (<10 – 100 cells), with an estimated ID_{50} (the dose at which 50% of the exposed population become ill) of 100-1000 organisms [19–22].

Pathogenic *E. coli* are a significant cause of disease and mortality worldwide and recent estimates attribute over 2.8 million annual acute illnesses to STEC infections, with close to 4000 annual cases of HUS [23]. From 2006-2015 there was approximately 730 annual reported STEC cases in Canada, however, for every reported Canadian case of STEC there is an estimated 10-47 unreported cases each year, suggesting a true incidence upwards of 7300 annual cases [24,25]. Alberta has been known to have one of the highest rates of *E. coli* O157 infection in Canada and from 2006 – 2016 there were 1526 reported cases of STEC in Alberta, an average of 139 reported cases per year [26,27]. STEC infections can cause uncomplicated diarrhea, bloody diarrhea, hemorrhagic colitis, HUS, and potentially death [11]. From a Canadian context, the most well-known STEC outbreak is probably the 2000 Walkerton, Ontario outbreak where over 2300 people became ill, 65 were hospitalized and 7 died due to the contamination of a municipal drinking water supply with *E. coli* O157:H7/*Campylobacter jejuni* [28]. In addition to the associated morbidity and mortality of this outbreak, there was an estimated tangible economic impact of over $64.5 million [29]. Due to the seriousness of the associated health outcomes, STEC are one of the more costly causes of waterborne gastroenteritis outbreaks [30].

The pathogenic *E. coli* O157 serotype has been the major focus of STEC research due to its association with multiple large-scale outbreaks and a higher risk of HUS development, however the importance of non-O157 STEC serotypes is

becoming increasingly accepted [19,31,32]. The previously mentioned 2011 outbreak of STEC O104:H4 in Germany, which caused 3816 cases of infection, 845 cases of HUS, and 54 deaths over a three month period, is one well known example that highlights the potential pathogenicity of non-O157 *E. coli* and the need to further investigate these 'under-rated' pathogens [15,32,33]. Non-O157 cases have likely been underreported historically, and recent increases in routine *stx* testing have resulted in substantial increases in the detection of these bacteria and their related illnesses [19,34–37]. Over 380 non-O157 serotypes have been isolated from humans, but the most commonly reported non-O157 STEC serotypes in North America belong to the 'big six' serogroups O26, O45, O103, O111, O121 and O145 [19,37–39]. The most common serotypes presenting clinically in Alberta are: O157:H7, O26:H11/O26:H-non-motile(HNM), O111:HNM/O111:H-nontypable(HNT) /O111:H8, and O121:H19 [40]. Different serogroups have been associated with varying degrees of virulence and it has also been suggested that the risk factors associated with STEC infection can differ across serotypes [1,41,42]. Across North America, and particularly within Alberta, studies have shown non-O157 *E. coli* are responsible for over 50% of the reported STEC cases [34,36,43–45].

Drinking water is an important, although perhaps overlooked, potential source of STEC infection and has been the root of a number of serious outbreaks, such as the aforementioned Walkerton outbreak [11,28,29,31,34,46–54]. A recently compiled review of North American *E. coli* O157 waterborne outbreaks has been provided by Saxena *et al.* [55], along with a corresponding table of relevant references. Waterborne disease in general, and more specifically enteric diseases such as STEC infections,

are largely under-reported for a variety of reasons such as self-treatment, lack of laboratory diagnosis during treatment, and short, self-limiting duration, all of which contribute to an estimated 300 enteric illnesses occurring in communities for every laboratory confirmed case [56–58]. The association between enteric disease and drinking water is further complicated due to issues such as constant exposure to drinking water from multiple sources, regional variations in diagnostic water testing and reportable infections, lack of proper sampling, low levels of target organisms, and misattributing or failing to fully determine the source of infection [52,58–61]. Despite this, from 1971-2006 there were 833 reported outbreaks associated with drinking water in the United States and from 1974-2001 there were 288 reported outbreaks linked to drinking water in Canada [58,62]. Importantly, these figures do not include any non-outbreak/sporadic waterborne disease and therefore likely underestimate the overall enteric disease burden associated with drinking water. Although sporadic cases of waterborne illness are often not well documented, they have accounted for 79% of reported STEC O157 cases in Alberta, and a recent study from British Columbia found that 100% of reported cases of drinking water associated enteric disease over a 10 year period were sporadic in nature [59,63,64]. Notably, due to the low infectious dose of STEC, only a single exposure to STEC-contaminated drinking water is required for a consumer to potentially become infected and experience related symptoms [65].

The importance of individual water systems to the overall burden of disease attributed to drinking water has been increasing in recent decades and the role of private wells and ground water has been specifically identified as an important

knowledge gap with regards to STEC-related illness [58,61,62,66]. Over 20% of waterborne illness outbreaks in Canada from 1974-2002 were attributed to private water systems and approximately 36% of drinking water outbreaks in the United Kingdom were also associated with private systems [67,68]. A recent Canadian study found that over 10-years, the risk of enteric illness was 5.2 times higher for residents that were served by private wells, and another Canadian study estimates that over 100,000 annual cases of acute gastrointestinal illness can be attributed to the consumption of water from private well and small water systems [59,69]. When estimating the relative role of different water sources in the transmission of waterborne enteric illness in Canada, Butler *et al.* [70] found that private well water was most frequently implicated for both bacterial and viral infections and Murphy *et al.* [69] estimated that close to 650 annual symptomatic cases of *E. coli* O157 infections are associated with consuming drinking water form private wells or small water systems.

In Canada over 4 million people (~13% of the population) rely on private water systems, with the majority relying on groundwater wells to supply their drinking water, and in Alberta approximately 450,000 - 600,000 people rely on private wells [67,71,72]. The Government of Canada recommends a multi-barrier approach to ensure drinking water quality in private sources, which includes source water protection, adequate treatment including disinfection, and a well-maintained distribution system [73]. Within Canada, as is common worldwide, total coliforms and *E. coli* are used as indicators of bacteriological water quality denoting vulnerability to contamination and fecal contamination respectively [73–75]. Despite similar

recommendations, in the United States during 2011 and 2012 untreated groundwater was found to be the second most common deficiency in drinking water outbreaks, and from 1971-2006 approximately 83% of private system outbreaks were due to contaminated untreated groundwater [61,62]. From 1971-2008 an estimated 67% of all reported groundwater outbreaks in the United States were attributed to improper design, maintenance and location of private wells and septic systems, and across Canada studies have found that approximately 25-35% of wells are contaminated with fecal indicator coliforms such as *E. coli* at levels that exceed government standards for safe drinking water [50,76–79]. Although the Government of Alberta suggests that well-owners submit well water for bacteriological analysis 2-4 times a year (depending on well-depth), the 2010 Alberta Well Water Survey found that less than 11% of respondents tested the quality of their well water on a yearly basis [67,71,80]. A survey of Alberta farms found that 59% of respondents had not tested their wells in the past 5 years or had never tested their wells, and only 40% of respondents used water treatment systems [67]. There is clearly a potential risk to public health in Alberta from STEC contamination of non-municipal drinking water supplies and a better understanding of the frequency and spatiotemporal patterns of occurrence of this pathogen can help inform public policy and prevention efforts to reduce the burden of enteric illness associated with this exposure pathway.

1.2 STEC in well water

For a groundwater well to become contaminated by STEC many contributing factors must converge in time and space [81,82]. First there must be a source of STEC in the external environment, and as naturally occurring enteric bacteria, this source is

likely through the fecal input of a primary or secondary STEC reservoir such as livestock or human sewage. The STEC must then be able to survive in this external environment for the duration of its transport into the freshwater environment and then further into groundwater, a process influenced by a myriad of pathogen-specific, environmental, and hydrogeological factors. These pathogens then need to survive in the groundwater source in high enough concentrations to be taken into the drinking water system, where they must persist through any existing water treatment, before exiting the tap where they can be detected or ingested.

Within North America cattle are generally recognized as the primary reservoir for STEC, although many secondary reservoirs or carriers such as sheep, goats, horses, pigs, alpaca, chickens, turkeys, deer, elk, coyote, and humans have been identified [47,65,83–85]. An extensive list of known STEC hosts can be found in the recent review of animal STEC reservoirs by Persad and Lejeune [86]. The exact reservoir structure of zoonotic STEC is still undetermined however and may be more complex than simply chains of infection within the cattle population [87]. Other animal populations and environmental niches that interact with cattle, such as local wildlife or water troughs, may play important roles in the maintenance of STEC within cattle and together may represent the true maintenance structure of the STEC reservoir in North America. Across the globe many animal populations have been identified as significant sources or even the primary reservoirs of STEC, and within North America cattle are likely only one of multiple important sources of STEC into the environment [86].

While the natural environment for most *E. coli* strains is within the gastrointestinal tract of warm-blooded hosts, these bacteria are able to survive, and in some cases grow, in the external environment [88,89]. STEC enter this environment through host defecation and numerous serotypes have been shown to survive within feces for up to 126 days at 15° C [90]. A field study carried out in alpine pastures of France found that these pathogens are able to survive in cow pats for the full duration that this feces exists on the landscape, despite large fluctuations in temperature and heavy rainfall [89]. STEC can also survive for long periods in soil. For example, an epidemiological investigation of an *E. coli* O157 outbreak in Scotland caused by sheep grazing in a field later used as a scout camp, recovered viable STEC from the soil up to 15 weeks after the exposure to sheep feces [91]. An *in-vitro* study using two types of soil sampled from Ireland, reported survival times of 14 non-O157 STEC serotypes at 10° C, for which survival ranged from 49-76 days depending on the soil type and, to a lesser degree, the serotype inoculated [92]. Numerous biotic and abiotic characteristics of soil can affect the survival of STEC in this environment, such as temperature, moisture, pH, soil composition, as well as competition and predation from the surrounding microflora [88,93,94].

Water is able to mobilize STEC present in soil or fecal matter during rainfall or snowmelt and transport these bacteria both vertically and/or horizontally into surface and ground water [88,95–98]. Once in a freshwater environment, STEC is still able to survive for extended periods of time. *In vitro* studies of O157 and non-O157 STEC in untreated well water found survival times from 15 days to greater than 56 days, depending on the water temperature and background predation [99,100]. In a

laboratory experiment investigating *E. coli* O157 survival in a variety of drinking and surface waters, viable STEC was still recovered from waters at lower temperatures (8° C) after 13 weeks [101]. Numerous results suggest that cool temperatures and low levels of background flora promote survival of STEC in freshwater [100–102]. Transport of STEC from the external environment into surface water and groundwater is largely determined by the local hydrology and hydrogeology of the area; characteristics such as aquifer type, subsoil type, strata thickness, preferential water flows, overland water flow, and in some cases even specific hydrogeological formations have been linked with groundwater outbreaks [50,66,103–105]. Within Alberta, it has been suggested that microbial contamination of well water may be more driven by sub-surface run-off processes influenced by shallow soil and topographic characteristics rather than deeper hydrogeological formations [106]. A well itself can also act as a transport pathway for STEC into groundwater due to improper design, construction or maintenance and this pathway has been specifically linked to numerous cases of waterborne illness and STEC infection [28,50,62,66]. There are many other localized situational risk factors that can influence well water contamination such as, well type (i.e., bored, drilled, etc.), well-depth, age of the well, proximity to livestock/animals and their density on the landscape, and the type of septic system and its location relative to a well on a property [50,60,77,107–109].

While the occurrence of STEC in well water is supported by a number of outbreaks worldwide, few studies have investigated the occurrence of these pathogens in well water supplies or even drinking water supplies in general

[52,55,102,110]. Table 1 outlines a list of studies deemed relevant for comparison to the current project and includes any study investigating STEC occurrence in drinking water that was not related to a specific outbreak scenario and any Canadian study investigating the occurrence of STEC in surface water. Canadian surface water studies were included due to both the absence of any STEC specific drinking water studies from Canada as well as the absence of any North American drinking water studies investigating the non-O157 STEC serogroups. Bacterial targets, location, type of water source, frequency of STEC occurrence, and methods of detection are outlined for ease of comparison to results in Chapter 3.

Three studies investigating STEC in drinking water were not included in the table - two due to the use of methods that prevented the calculation of the frequency of STEC occurrence per water sample tested, and a third because the study failed to confirm the presence of *stx* genes in the presumed STEC isolates. The first was a study from Brazil that recovered 12 STEC from both untreated and treated drinking water sources in both rural and urban areas across the Paraná State [111]. The second was a study from Uttar Pradesh, India, where 18 STEC were recovered from a number of separate sampling sites all served by the main treated and piped drinking water source for the city of Lucknow [112]. The third study investigated rural drinking water from various sources (ground water, surface water, piped supplies) in Uttar Pradesh, India and recovered STEC from 5/188 (2.7%) water samples tested [113]. These researchers used culture-based methods to isolate *E. coli* and subsequently identified STEC based on the serotype of the isolate. Any *E. coli* isolate belonging to O-groups commonly associated with STEC serotypes was considered to be an STEC.

However, serotype does not determine pathogenicity or resulting pathogroup classification and therefore, the results of this study should be interpreted with a degree of caution as these isolates may not be true STEC. For example, *E. coli* which belong to serogroups commonly associated with STEC are often found without the *stx* genes required for STEC classification [114–116].

Table 1 – Research papers investigating STEC in drinking water or Canadian surface water

E. coli Target	Water Source and Location	Occurrence (per water sample)	Reference
STEC	Rural drinking water (groundwater wells and springs), Upper Austria, Austria	1/2633 water samples (0.04%) 1/280 (0.4%) *E. coli* positive water samples	Halabi *et al.* (2008)[114]
O157:H7	Private drinking water (groundwater wells), Ohio, United States	7/180 (3.9%)	Won, Gill, and LeJeune (2013) [117]
O157:H7	Private drinking water (groundwater wells), Netherlands	4/147 (2.7%)	Schets *et al.* (2005) [108]
STEC	Surface Water, Southwestern B.C. (a) Serpentine b) Sumas c) Nicomekl River d) Lower Fraser watersheds)	a) 19/82 (23.2%) b) 21/97 (21.6%) c) 19/86 (19.5%) d) 6/65 (9.2%)	Nadya *et al.* (2016) [118]
a) O157:H7 b) STEC	Surface Water, Southwestern Ontario, Canada (Grand River watershed)	a) 9/893 (2.1%) b) 75/236 (31.8%)	Johnson *et al.* (2014) [51]
O157:H7	Surface Water, Southern Alberta, Canada (Oldman river watershed, 2000-2001)	27/1608 (1.7%)	Gannon *et al.* (2004) [119]
O157:H7	Surface Water, Southern Alberta, Canada (Oldman river watershed,1999-2000)	0.9% ± 0.2% (13/1483)	Johnson *et al.* (2003)[120]
O157:H7	Surface Water, Canada (Sumas, Oldman River, South Nation River, Bras d'Henri and Fourchette watersheds)	27/902 (3.0%)	Edge *et al.* (2012) [121]
O157:H7	Surface Water, Southern Alberta, Canada (Oldman river watershed, 2002-2004)	12/406 (3.0%)	Walters, Gannon, and Field (2007) [122]
O157:H7	Surface Water, Southern Alberta, Canada (Oldman river watershed, 2005-2007)	8/342 (2.3%)	Jokinen *et al.* (2011) [124]
O157:H7	Surface Water, Eastern Ontario, Canada (South Nation river watershed)	15/1186 (1.3%)	Wilkes *et al.* (2011) [125]
O157:H7	Surface Water, Eastern Ontario, Canada (South Nation river watershed)	5/823 (0.6%)	Wilkes *et al.* (2009) [126]
O157:H7	Surface Water, Southwestern B.C., Canada (Salmon River watershed)	5/186 (2.7%)	Jokinen *et al.* (2010) [127]

1.3 Seasonality

Seasonality in infectious disease occurrence refers to the regular recurring patterns of disease incidence that align with a particular season and that alternate with lower background levels of incidence [128,129]. These seasonal patterns represent relationships between intrinsic host attributes, pathogen physiology, environmental processes, and social behaviours, all of which are context dependent in both time and space. Seasonality is a global phenomenon and in Canada the major enteric pathogens such as *Cryptosporidium* spp., *Giardia* spp., *Campylobacter* spp., *Salmonella* spp., and *E. coli* all exhibit seasonal occurrence patterns [128,130–133]. Each pathogen has a respective pattern, but in Canada's temperate climate they generally peak during the warm months of the year and are lowest during the cold winter months. Important environmental effects on pathogen occurrence, transmission pathways, as well as social behaviours, are believed to be strong drivers of seasonality in enteric disease [128,131,134].

STEC cases in North America generally occur in the period from early spring to early fall, and usually peak during the summer months [34,37,39,135,136]. A study in Ontario found marked seasonal patterns in *E. coli* O157 infection that peaked in July of each year over a period of six years [130]. A more recent Canadian study found that *E. coli* O157 incidence in the province of New Brunswick followed a seasonal pattern with large summer peaks over a 9-year study period [131]. Five temporal clusters of *E. coli* O157 cases were also identified in this study, all of which occurred within the late spring or summer of the calendar year. An Alberta based study investigating the

clustering of *E. coli* O157 outbreak cases found that all of the temporal clusters occurred within the late spring to early autumn period [137]. Another study from Alberta found the largest number of cases of *E. coli* O157 reported over 9 years occurred in July, with 62% of total reported cases having occurred over the summer months and 82% having occurred between May and September [138]. Enteric outbreaks associated with drinking water in Canada tend to follow a similar seasonal pattern [58]. For instance Thomas *et al.* [81] found that while Canadian outbreaks of waterborne disease were reported in all 12 months of the year, the greatest proportion of these outbreaks (56%) occurred from March to July. Currently there are no known investigations into the seasonal occurrence of waterborne STEC infections attributed to drinking water or well water more specifically. A 2017 study from Alberta reported a statistically significant summer peak (peak date: July 24) in *E. coli* contamination in drinking water wells from 2004-2012, although they did not specifically determine the presence of any enteric pathogens [74].

1.3.1 Seasonality in STEC shedding

Seasonal patterns in the shedding of STEC bacteria from cattle have been found in numerous studies. Hancock *et al.* [139] detected seasonal patterns of *E. coli* O157 shedding in numerous cattle herds in the northwestern United States and found the highest prevalence of *E. coli* O157 excretion by herds was in June and the lowest was in December, with an overall pattern that suggested short periods of relatively high prevalence separated by longer periods of reduced shedding. A number of other studies, including a study from southern Alberta, have supported

this seasonal pattern, each finding seasonal patterns of *E. coli* O157 shedding in feces that peak in the summer (with some peaks extended into autumn) and that are lowest in the winter [140–143]. Ogden *et al.* [144] found the reverse of this pattern however, with *E. coli* O157 peaking in the winter in Scotland, although this study seems to be an anomaly and other studies have suggested that this alternate pattern can be attributed to cattle management practices of the study location (i.e., high density housing and the presence of super-shedders in the population) [93,144,145]. Studies investigating non-O157 STEC serotypes have found varying seasonal shedding patterns that are often serotype dependent, some of which do not necessarily peak in the summer, and some having no seasonal pattern at all [116,140,143,146]. Pearce *et al.* [146] and Dewsbury *et al.* [143] found that serotypes O26, O45, O111, O130, and O145 did follow similar seasonal patterns as O157 with prevalence highest during in the summer (and sometimes autumn) and lowest during the winter. Alternatively, Barkocy-Gallagher *et al.* [140] reported seasonal peaks in non-O157 recovery (serotypes were not identified) in the spring and fall. Recently a study by Stanford *et al.* [116] in southern Alberta found a wide variety of patterns within non-O157 STEC. Polymerase chain reaction (PCR)-based tests, supplemented with culture confirmation, indicated that prevalence of: i) serotypes O26 and O121 followed a pattern similar to O157; ii) serotypes O111 and O145 unexpectedly peaked in the winter and were lower in the summer; iii) O45 serotypes stayed constant from spring to fall and then dipped in the winter, and iv) O103 serotypes stayed constant the whole year. In general, most studies do show an increase in STEC shedding from cattle during the warmer months of the year however.

While the vast majority of research investigating seasonal patterns of STEC shedding has focused on cattle, there have been select studies that have investigated other hosts such as sheep, goats, swine and deer [147–154]. Low STEC positivity rates prevented the determination of seasonal patterns of STEC shedding in many of these studies, although a number of ovine studies have reported seasonal patterns of STEC shedding in warmer months of the year. Further investigation into the shedding of STEC from hosts other than cattle remains an important knowledge gap in understanding the sources of STEC into the environment.

1.3.2 Seasonality and weather

Meteorological conditions, primarily precipitation and temperature, are major determinants of overall seasonal change and are important factors in determining seasonal patterns of waterborne enteric illness [128,155]. Specifically for the climate of southern Alberta, above freezing temperatures, vernal melt, and vernal rainfall events can input concentrated pathogen-rich runoff into the freshwater environment during the spring, and periods of hot, dry conditions, separated by intense sporadic rainfall events in the summer, can also mobilize pathogens into the freshwater environment [58,82,156]. Numerous studies have found associations between rainfall, particularly extreme precipitation events, and cases or outbreaks of waterborne disease [155,157]. Extreme precipitation events generally refer to periods with notably higher than average rates of precipitation, whether that be a certain amount of rainfall within a specified time period, a precipitation rate above a specified threshold (i.e., the 90th percentile of average precipitation

across a study period) or similar measures, although a single working definition of this term has not yet been agreed upon in the literature [82,155,158].

Curriero *et al.* [158] retrospectively investigated reported waterborne disease outbreaks in the United States from 1948-1994 and found that 68% of outbreaks were preceded by precipitation events above the 80th percentile of average precipitation across this study period and 51% of outbreaks were preceded by precipitation events above the 90th percentile. Similarly, a Canadian study quantitatively analyzed associations between waterborne outbreaks and weather conditions and found that the probability of a waterborne outbreak increases by a factor of 2.28 for precipitation events above the 93rd percentile of average precipitation across the study period (1974-2001) [81]. Relationships between extreme precipitation events and STEC specifically in drinking water have not yet been investigated within a North American context, although associations between rainfall and STEC in surface water have been reported. Nadya *et al.* [118] reported significant correlation between temperature and average precipitation 3 days before recovering STEC positive surface water samples from the Lower Mainland of British Colombia, Canada. Likewise, a study from southern Alberta reported that the majority of *E. coli* O157 isolates were recovered from surface water after at least 8.9mm of rain fell during the 3 days previous to, and including, the day of sampling [123]. Within well water, Valeo *et al.* [106] reported a spatiotemporal cluster region in southern Alberta where drinking water wells were prone to *E. coli* contamination during extreme weather events, although *E. coli* pathotypes were not investigated specifically. A recent study from Ireland investigated weather events and

waterborne STEC illness and found that heavy rainfall events (a 24 hour period with greater than 33mm of rainfall), cumulative rainfall, and temperature, all had statistically significant associations with the occurrence of waterborne STEC outbreaks over a period of 8 years [82]. They determined that the relative odds of a waterborne STEC outbreak is 22.89 times greater if there is a heavy rainfall event in the preceding 7 days and that in terms of cumulative rainfall, the relative odds of a waterborne STEC outbreak increases by a factor of 1.05 for every additional 1mm of rainfall recorded.

Positive associations between temperature and the incidence of pathogenic *E. coli* infections are commonly reported by studies across the globe and while positive associations have also been reported between temperature and the incidence of waterborne disease, this relationship seems to be less straightforward [134,155]. Within Canada, Thomas *et al.* [81] found that with air temperatures above 0° C for a four week period before a waterborne outbreak, single degree (° C) increases in maximum ambient air temperature would result in the relative odds of an outbreak to increase by a factor of 1.007, meaning that a 5° C increase in maximum air temperature during this four week period would result in greater than a 4-fold increase in the relative probability of a waterborne outbreak. Fleury *et al.* [132] described a strong non-linear association between reported enteric *E. coli* infections and ambient temperature in Alberta for the six respective weeks prior to an infection, and found that the log relative risk of reported *E. coli* weekly case counts increased by 6% for every degree increase in weekly mean temperature [132]. Another study from Alberta found that mean temperature was a significant predictor of

monthly incidence rates of *E. coli* O157 in the province [159]. As for investigations specifically into relationships between temperature and waterborne STEC, the previously mentioned Irish study by O'Dwyer *et al*. [82], found that the relative odds of a waterborne STEC infection increases by a factor of 1.37 for every one degree (° C) increase in mean monthly temperature.

1.4 Research rationale and objectives

Even though waterborne enteric illnesses are a well-documented burden of disease across the globe, as well as in Canada, there remain important knowledge gaps in understanding the epidemiology of these diseases. One important gap is concerning endemic waterborne disease associated with groundwater and private water systems [59]. As mentioned previously, these water systems are important infection pathways for enteric disease in North America, and their contribution to the overall burden of enteric disease has been increasing. Currently there is a lack of information surrounding the prevalence of enteric pathogens within groundwater systems, particularly for enteric bacteria [66]. In order to improve exposure assessment and provide more accurate endemic risk estimates than currently afforded by the use of fecal indicator organisms as pathogen surrogates, direct pathogen detection within groundwater sources is required. Targeted ecological studies at the regional level that focus on area-specific temporal and spatial elements have been suggested in order to better inform local public-health policy

and decision-making [66,128,160]. The identification of disease-specific spatiotemporal occurrence patterns are necessary for the improvement of surveillance and prevention strategies, as well as increasing our understanding of the many interrelationships and interactions involved in the etiology of waterborne diseases. STEC in particular have been underrepresented within the literature thus far, and this shortage of information is exacerbated by the lack of studies that include both non-O157 and O157 serotypes in their investigations.

1.4.1 Objective 1: Determine the frequency of occurrence of STEC in non-municipal drinking water sources across southern Alberta

The presence of *stx* genes in submitted *E. coli* positive non-municipal drinking water samples was used as a proxy for the occurrence of STEC in these water sources. The hypothesis was that STEC would be consistently found within drinking water samples at a relatively low frequency of occurrence, similar to previously reported rates of STEC occurrence in private drinking water sources [108,114].

1.4.2 Objective 2: Evaluate STEC serotype occurrence in non-municipal drinking water samples across southern Alberta

It was hypothesized that a diverse range of STEC serotypes would be found within non-municipal drinking water sources consisting of both the most clinically

relevant serogroups in Alberta as well as serogroups not commonly recovered from clinical patients. The recovery of multiple genetically distinct STEC clones from the same *stx* positive drinking water sample was also expected in some, but not all, cases.

1.4.3 Objective 3: Investigate spatiotemporal patterns of occurrence of STEC in non-municipal drinking water samples across southern Alberta

It was hypothesized that statistically significant patterns in the occurrence of STEC in non-municipal drinking water samples in time, space, and space-time would be identified, including seasonal occurrence with a peak during the summer months, spatial clustering associated with stable geographic risk factors, and clustering in space-time with extreme precipitation events that occurred during the study period.

1.5 Thesis organization

The work presented in this thesis will address each of the aforementioned objectives in the following manner:

- Chapter 2 describes the methodologies used throughout the thesis;
- Chapter 3 presents and discusses both the findings of *stx* gene occurrence in submitted non-municipal drinking water samples as well as trends within these results;

- Chapter 4 includes a discussion of the methodologies used to recover STEC isolates from environmentally sourced drinking water samples, as well as presents and discusses the results from the STEC recovery protocol, the determination of clonality among recovered isolates, and the serotyping of genetically unique STEC isolates;
- Chapter 5 presents and discusses the results of statistical models testing patterns of STEC occurrence in time, space, and space-time; and
- Chapter 6 highlights key findings of the thesis, provides a general discussion for the results of the three previous chapters as well as some of the research limitations, and provides recommendations for future research.

CHAPTER 2: Research Methods

2.1 Non-municipal well-sourced drinking water samples

The groundwater well samples included in the analysis were all voluntarily submitted from non-municipal drinking water samples processed by the Provincial Laboratory for Public Health in Calgary, Alberta, Canada (ProvLab Calgary) for routine microbial contamination (i.e., *E. coli* and total coliforms) from March 2004 – July 2016. The catchment area for ProvLab Calgary included the southern third of the province of Alberta, the area south of the municipality of Red Deer. For this study, non-municipal drinking water samples included any well-sourced drinking water sample that were either submitted by a private landowner or from a water system that was not regulated by the Ministry of Environment and Parks in Alberta [161]. These samples were collected and submitted by the well-overseer who was also responsible for providing all the associated requisition information such as name, address, and location details. Each sample (250ml bottle) was transported on ice to the ProvLab within 24 hours of sampling in order for the sample to be accepted for routine processing. Requisition information was hand-written by the submitter at the time of collection and subsequently entered manually into the laboratory information system by the receiving ProvLab staff. The ProvLab is an ISO 17025 accredited laboratory for analysis of microbiological water quality. For the purposes

of this thesis, these passively acquired well-water samples were considered to be generalizable to the greater population of non-municipal drinking water sources across the catchment area of southern Alberta.

Due to the large variety of waterborne pathogens that may be present within drinking water sources, testing for each potential pathogen individually is not feasible for a public health laboratory and therefore indicator organisms are commonly used as surrogate markers for the risk of pathogenic contamination [162]. Since the majority of waterborne pathogens are associated with fecal sources of contamination, enteric coliform bacteria, such as *E. coli*, are often used as indicator organisms of fecal contamination. In Alberta, the ProvLab tests for both total coliforms and *E. coli* as indicators of microbial water quality. For this study, all submitted non-municipal drinking water samples were processed using the Colilert® presence/absence defined substrate method (IDEXX Laboratories Inc., Westbrook, ME, USA) to simultaneously detect the presence of total coliforms and *E. coli* according to the manufacturer's protocol. In diagnostic terms, when *E. coli* is present in a water sample, the media will change from a clear solution to a yellow colour (β-galactisidase activity) and will also fluoresces under UV light (β-glucaronidase activity), whereupon it is considered a positive result. Since this study focused strictly on STEC bacteria, Colilert® results were used to screen-out *E. coli* negative drinking water samples that would not likely contain STEC.

From March 2004-May 2015 any *E. coli* positive Colilert® enriched drinking water samples were archived for further analysis by adding 500 µl of the Colilert® sample/substrate solution to 500 µl of 2x Skim Milk and frozen at -70° C. These

samples were considered *retrospective* drinking water samples and operationally defined as those culture samples frozen and stored for a designated period of time. From June 2015 to July 2016, two 1ml aliquots of the sample/substrate solution of *E. coli* positive Colilert® enriched drinking water samples from ProvLab Calgary were sent to the University of Alberta (U of A, Edmonton, AB, Canada), one of which was immediately processed for *stx* occurrence. These *E. coli* positive Colilert® enriched drinking water samples received from June 2015 to July 2016 were considered *prospective* drinking water samples and operationally defined as culture samples **not** having being subjected to freezing or storage conditions before analysis.

2.2 Determining *stx* positivity of non-municipal well-sourced drinking water samples

Quantitative polymerase chain reaction (qPCR) analysis enables specific and sensitive detection of microorganisms in water matrices, and virulence factors are ideal qPCR gene targets since they are generally in low-copy numbers within a genome and often indicative of a specific species or pathotype [163–165]. qPCR analysis of the *stx1* and *stx2* genes was used as a proxy screen to determine the potential presence or absence of STEC in retrospective and prospective *E. coli* positive Colilert® drinking water samples processed from non-municipal well-sourced systems. The following section outlines the screening process.

2.2.1 Retrospective drinking water samples

Retrospective *E. coli* positive drinking water samples, which contained 500 µl *E. coli* positive Colilert® drinking water and 500 µl 2x skim milk, were thawed and a 1ml aliquot was transferred to 9ml of Tryptic Soy Broth (TSB). This culture was incubated at 35° C for 16-18 hours to aid in the resuscitation of the cryo-preserved bacteria [101]. After incubation the TSB enrichment was split two ways: 100 µl was transferred to a specified well in a 96-well Greiner plate (Greiner Bio One International, Monroe, NC, USA) for *stx1*/*stx2* qPCR analysis, and 500 µl was transferred to a screw-cap cryovial containing 500 µl of 2x skim milk for re-archival at -70° C.

2.2.2 Prospective drinking water samples

Upon receiving samples at the U of A, 1 mL aliquots of prospective *E. coli* positive Colilert® enriched drinking water were centrifuged at 13 000 x *g* for 5 minutes. Following centrifugation, the supernatant was removed and the bacterial pellet was stored at -70° C until processed for molecular testing. Five-hundred microliters (500µl) from the second 1 mL prospective aliquot was transferred to a screw-cap cryovial containing 500µl of 2x skim milk and kept at -70° C for further processing. Frozen sample pellets were subsequently re-suspended in 200 µl of clinical laboratory reagent water (CLRW) and 100µl of this solution was transferred to a specified well in a 96-well Greiner plate for *stx1*/ *stx2* qPCR analysis.

The prospective data set was included in the analysis in order to determine if cryopreservation and subsequent resuscitation had any effect on the rates of STEC recovery from Colilert® drinking water samples.

2.2.3 PCR control strains

Two *E. coli* strains were used as controls for the *stx1*/*stx2* qPCR analysis: i) the positive control strain ATCC 35150 that harbors both *stx1* and *stx2* genes; and ii) the negative control strain ATCC 35218 that does not harbor *stx1* or *stx2* genes. Control strains were grown in TSB broth at 35° C for 16 hours and aliquots of 1 ml were centrifuged at 13 000 x *g* for 5 minutes. Following centrifugation, the supernatant was removed and the bacterial pellet was stored at -70° C until needed. Pellets were re-suspended in 1mL of CLRW, and 100µl of this solution was transferred to a specified well in a 96-well Greiner plate for use in *stx1/stx2* qPCR analysis.

2.2.4 *stx1*/*stx2* qPCR analysis

Ninety-six-well Greiner plates containing processed TSB samples were stored at -20° C until thawed for *stx1*/*stx2* qPCR analysis. Thawed plates were boiled at 95° C for 10 minutes using an Eppendorf Mastercyler Thermal Cycler (Eppendorf Canada, Mississauga, ON, Canada) in order to lyse cells and release DNA. Five microliters (5µl) of each sample was transferred from the 96-well Greiner plate to a corresponding well on an ABI Fast 96-well Real-Time PCR plate (Thermo Fisher Scientific, Waltham, MA, USA) containing the corresponding *stx1* or *stx2* PCR

reagents (Table 2). Afterwards the Greiner plates containing processed TSB samples were resealed and stored at -20° C for archival and subsequent analysis (i.e., GTG(5) rep-PCR [see below]). Five microliters (5µl) of a pIDTsmart-internal amplification control (IAC) [20 copies/µl] (Integrated DNA Technologies, Inc., Coralville, IA, USA) was also added to each well for the *stx1*/IAC multiplex reaction. At minimum, three 'no template control' (NTC) wells containing only *stx* Master Mix solution and PCR water were included on each PCR plate, and acted as negative molecular controls. Each of the PCR control strains were also included in a respective well for each PCR plate. Each PCR plate included five 1:10 dilutions of plasmid DNA containing the *stx* gene under investigation (pCR2.1- *stx1*; pCR2.1- *stx2*) ranging from 5-50,000 gene copies per reaction in order to construct a standard quantification curve and determine the C_t value for the presence/absence cut off of each qPCR run.

After set-up, each PCR plate was centrifuged at 12000 rpm for 2 minutes before being loaded into an ABI 7500 Fast Real-time PCR System (Thermo Fisher Scientific, Waltham, MA, USA). Reaction conditions were as follows: a holding stage at 50° C for 2 minutes; a secondary holding stage at 95° C for 30 seconds; followed by a minimum of 40 cycles of 95° C for 3 seconds & 60° C for 30 seconds.

The *stx1*/stx2 primers and probes (Table 3) and TaqMan™ assay that were used had been previously validated (sensitivity: 100%; specificity: 100%) for the most common clinical *stx* subtypes (all known *stx* subtypes except *stx2f*) [166], and the limit of detection with 95% confidence (LoD_{95}) was previously determined to be 5 gene copies per reaction. The 5 gene copy LoD_{95} of this qPCR assay was used as the presence/absence cut off for the qPCR assay. This absolute cut off was calculated for

each qPCR run using the C_t value corresponding to 5 gene copies per reaction as determined by the standard quantification curve of each individual plate being analyzed.

Table 2 - *stx1*/*stx2* qPCR reagents

Solution	Reagents
stx1/IAC Master Mix (per reaction)	1 µl PCR water (Thermo Fisher Scientific, Waltham, MA, USA), 12.5 µl 2x TaqMan™ Fast Advanced Master Mix or TaqPath™ qPCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA), 0.5 µl 10mg/mL BSA (Merck KGaA, Darmstadt, Germany), 1 µl 25x *stx1* Primer/Probe mix
stx2 Master Mix (per reaction)	3.6 µl PCR water, 10 µl 2x TaqMan™ Fast Advanced Master Mix or TaqPath™ qPCR Master Mix, 0.4 µl 10mg/ml BSA, 1.0 µl 20x *stx2* Primer/Probe mix
stx1 Primer/Probe Mix	0.450 µM *stx1*-F primer (Thermo Fisher Scientific, Waltham, MA, USA), 0.450 µM *stx1*-R primer (Thermo Fisher Scientific, Waltham, MA, USA), 0.125 µM *stx1* Probe (Thermo Fisher Scientific, Waltham, MA, USA), 0.450 µM IAC-F primer (Thermo Fisher Scientific, Waltham, MA, USA), 0.450 µM IAC-R primer (Thermo Fisher Scientific, Waltham, MA, USA), 0.125 µM IAC Probe (Thermo Fisher Scientific, Waltham, MA, USA)
stx2 Primer/Probe Mix	0.450 µM *stx2*-F primer (Thermo Fisher Scientific, Waltham, MA, USA), 0.450 µM *stx2*-R primer (Thermo Fisher Scientific, Waltham, MA, USA), 0.125 µM *stx2* probe (Thermo Fisher Scientific, Waltham, MA, USA)

Table 3 – *stx* PCR primers and probes[a,b]

Primer/Probe	Sequence
stx1 Forward	5'-CATCGCGAGTTGCCAGAAT-3'
stx1 Reverse	5'-GCGTAATCCCACGGACTCTTC-3'
stx1 Probe	5'-[FAM]-CTGCCGGACACATAGAAGGAAACTCATCA-[TAMRA]-3'
IAC Forward	5' -CTAACCTTCGTGATGAGCAATCG-3'
IAC Reverse	5'-GATCAGCTACGTGAGGTCCTAC-3'
IAC Probe	5'-[VIC]-AGCTAGTCGATGCACTCCAGTCCTCCT-[MGBNFQ]-3'
stx2 Forward	5'-CCGGAATGCAAATCAGTC-3'
stx2 Reverse	5'-CAGTGACAAAACGCAGAACT-3'
stx2 Probe	5'-[FAM]-ACTGAACTCCATTAACGCCAGATATGA-[TAMRA]-3'

[a] Primer and probe sequences based on Chui *et al.* [166]
[b] All primers and probes were HPLC purified

2.3 Temporal analysis of *stx* positivity rates in non-municipal well-sourced drinking water supplies

Due to substantial differences in the number of voluntarily submitted drinking water samples over the study period, and particularly during certain periods of time compared to others, the frequency of occurrence of *stx* in non-municipal drinking water sources was investigated using positivity rates per sample. The term "positivity", when used in this thesis, is representative of frequency of occurrence rather than prevalence, since the denominator represents all of the wells tested by ProvLab Calgary and not all of the wells at risk across southern Alberta. Two distinct *stx* positivity rates were calculated: i) positivity rates per submitted non-municipal drinking water samples, and ii) positivity rates per *E. coli*-contaminated non-

municipal drinking water samples. The number of samples used in the denominator of each positivity rate was chosen for ease of understanding and comparison. Each rate was aggregated into annual and monthly time series for analysis. Graphs and tables for this thesis were created with Microsoft Excel Version 14.7.2 (Microsoft Corporation, 2010, Redmond, WA, USA). Flowcharts were created using the Drawing feature on Google Drive (Alphabet Inc., Mountain View, CA, USA).

2.3.1 Investigating the relationship of *stx* positivity rates and time

Poisson and negative binomial regression were used to model the *stx* positivity rates. Regression was performed using STATA 13.1 software (StataCorp. 2013, College Station, TX, USA). When using STATA, the 'exposure' model option was implemented in order to analyze rates of events and the 'irr' reporting option was implemented to report incidence rate ratios [167,168]. Time-aggregated totals of either submitted non-municipal drinking water samples or *E. coli* positive non-municipal drinking water samples were used as the exposure variable depending on the rate under investigation. The predictor variable of time was modeled as an indicator variable by year (2004 – 2016) or month (1 – 12). Poisson model fit was evaluated by deviance goodness-of-fit tests (P-value < 0.05). Poisson models with poor fit were subsequently modeled using negative binomial regression and over-dispersion was determined by likelihood-ratio chi-squared tests of the dispersion parameter alpha, testing if alpha was equal to zero (P-value < 0.05). Negative binomial model fit was assessed with deviance chi-squared tests (P-value < 0.05) and Anscombe residual plots were used to identify outliers and influential observation. Likelihood-ratio tests were used to evaluate the overall contribution of

the predictor variable (P-value < 0.05) and Wald tests were used to evaluate the contribution of individual indicator variables compared to the referent (P-value < 0.05).

2.4 Recovery of viable STEC from non-municipal well-sourced drinking water supplies

STEC lack a single, consistent, and determinable phenotypic trait that would allow simple differentiation and detection from background flora on growth media, and as a result differentiating between STEC and other *E. coli* can be difficult [118,169,170]. A number of laboratory and commercially developed selective agars are available to specifically aid in the recovery of STEC from mixed microbiological samples, however CHROMagar™ STEC agar [CA-STEC] (CHROMagar Microbiology Inc., Paris, France) has been singled out as a preferred option for STEC isolation [170–175]. For this study, CA-STEC was used as a selective growth media to aid in the recovery of viable STEC from within the population of background flora found in non-municipal well-sourced drinking water supplies.

The overgrowth of non-STEC background flora on CA-STEC during an initial STEC recovery protocol consistently prevented the successful isolation of presumptive STEC colonies [discussed extensively in Chapter 4]. A revised protocol for the recovery of STEC from drinking water samples using CA-STEC was developed to combat this overgrowth of non-STEC bacteria by including both a selective enrichment step and multiple incubation temperatures, into this

procedure. The revised protocol is described below and the initial protocol is described in the following section of this chapter.

2.4.1 Revised CHROMagar™ STEC isolation protocol

If a drinking water sample was determined to be *stx* positive by qPCR analysis, 100µl from the archived TSB/Skim Milk solution preserved at -70° C was added to 3ml tubes of MAC broth and incubated for 4 hours at 35° C to both resuscitate the bacteria and select for gram-negative, lactose-fermenting bacilli, such as *E. coli*.

After resuscitation, 100µl of the MAC enrichment was spread onto a CA-STEC plate and incubated in the dark at 44° C for 18-24 hours, with the remaining MAC enrichment kept at 4° C for the duration of the isolation protocol. Once incubation was complete, each plate was observed under a UV-light to assess the fluorescence of the colonies. As per the manufacturer's instructions, STEC should grow as mauve colonies on CA-STEC, with the majority of non-O157 STEC also fluorescing under UV light [176]. Non-pathogenic bacteria and other background flora should be inhibited or grow as blue or colourless colonies. Isolated mauve coloured colonies were picked at random from the incubated plate, added to 5ml tubes of TSB and incubated for 4 hours at 35° C while being shaken at 200rpm. This further enrichment step was done to increase the number of cells for successful cryopreservation and archival of each isolated colony.

If there was a shortage of isolated mauve colonies, non-isolated mauve colonies were picked and re-streaked to a CA-STEC plate and again incubated in the dark at 44° C for 18-24 hours. Plates were again assessed under UV-light and any

resulting isolated mauve colonies on the re-streaked plates were processed as above. If there were no, or very few mauve colonies present, a second 100µl aliquot was removed from the MAC enrichment being kept at 4° C, isolation spread onto a CA-STEC plate and incubated in the dark at 37° C. CA-STEC plates and mauve colonies were assessed and processed as above, with any re-streaked plates incubated at 37° C.

The TSB enriched culture sample from each mauve isolate was split two ways: 500µl of the enriched solution was transferred to a screw-cap cryovial containing 500µl of 2x skim milk and kept at -70° C for archival, and 100µl was transferred to a specified well in a 96-well Greiner plate for *stx1*/*stx2* qPCR analysis. The same *stx1*/*stx2* qPCR analysis protocol as described above was used to determine if mauve isolates were in fact STEC. A flowchart of this revised CHROMagar™ STEC isolation protocol is presented in Figure 1.

A select number of non-mauve coloured colonies were isolated and processed using the protocol outlined above in order to verify that environmental STEC strains behave similarly to clinical STEC strain and only form mauve coloured colonies on CA-STEC agar.

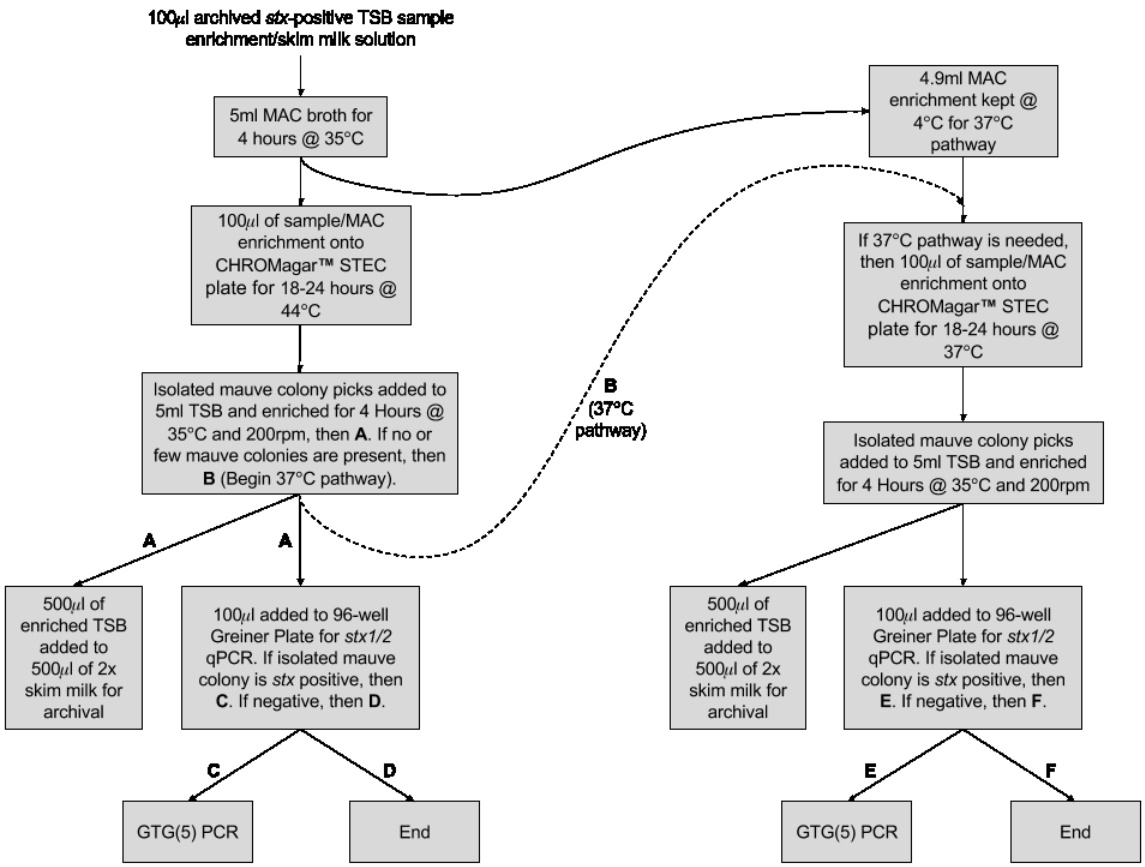

Figure 1 – Recovery of STEC isolates from non-municipal drinking water samples using CHROMagar™ STEC agar (revised protocol).

2.4.2 Initial CHROMagar™ STEC isolation protocol

Two portions of the initial CA-STEC isolation protocol differed from the revised version: i) the first enrichment step, and ii) the primary incubation temperature. The first enrichment step was as follows: If a drinking water sample was determined to be *stx* positive by qPCR analysis, 100µl from the archived TSB/Skim Milk solution preserved at -70° C was added to 5ml tubes of TSB and incubated for overnight at 35° C to resuscitate the bacteria. Subsequent incubation of CA-STEC plates was carried out at strictly 37° C for this initial protocol and

therefore re-plating only occurred when there was a lack of isolated mauve colonies and not a lack of mauve colonies generally. The remainder of the protocol was identical to the revised methods described previously.

2.5 Assessing genetic diversity of STEC in drinking water samples

Due to the possibility of multiple STEC strains being present in a single drinking water sample, the genetic similarity of each isolated strain from a respective water sample was investigated for clonality using a repetitive sequence-based PCR (rep-PCR) comparison analysis. rep-PCR analysis can be used as a genetic screen to determine relative relatedness of STEC strains and reduce the number of potential isolates that require more comprehensive typing methods [177]. A Canadian study comparing the efficacy of five commonly used rep-PCR primers to differentiate environmentally-derived aquatic *E. coli* strains found that the (GTG)5 primers had the highest discriminatory power and were the most suitable for identifying genetic differences in these bacteria [178]. A small subset of *stx* positive mauve isolates were tested using the two most discriminatory rep-PCR methods outlined in this study [(GTG)5 and BOX A1R (Table 4)] and similarly found that the (GTG)5 primer resulted in a greater number of bands per *E. coli* fingerprint. As a result, (GTG)5 rep-PCR was determined to be the most appropriate method for the purposes of this analysis. Genetically unique isolates, as determined by (GTG)5 rep-PCR were subsequently serotyped to enable comparison to previously identified STEC in the literature.

2.5.1 (GTG)5 rep-PCR

Archived 96-well Greiner plates containing processed CHROMagar™ isolates were thawed and 5µl from wells containing *stx* positive isolates were diluted 1:10 into RNA free water before being transferred to an ABI Fast 96-well Real-Time PCR plate containing 15µl of the (GTG)5 primer and Master Mix reagents (Table 5). A 2720 Thermal Cycler (Thermo Fisher Scientific, Waltham, MA, USA) was used for the PCR reaction under the following cycling conditions and as described by Mohapatra & Mazumder [178]: 95°C for 4 minutes, then 30 cycles of [94°C for 3 seconds; 92°C for 30 seconds; 40°C for 1 minute; 65°C for 8 minutes], with a final extension at 65°C for 8 minutes.

Table 4 - rep-PCR primers

(GTG)5[a]	5'-GTGGTGGTGGTGGTG-3'
BOX A1R[a]	5'-CTACGGCAAGGCGACGCTGACG-3'

[a] Based on Versalovic *et al.* [179]

Table 5 – (GTG)5 rep-PCR reagents

Solution	Reagents
(GTG)5 Master Mix (per reaction)	2.5 µl PCR Water, 10.0 µl 2x Maxima Hot Start Master Mix, 0.5 µl 10mg/ml BSA, 2.0 µl 100 µM (GTG)5 Primer

2.5.2 Comparison of (GTG)5 DNA-fragment fingerprints

A QIAxcel® Advanced (QIAGEN®, Montreal, QC, Canada) system and QIAxcel® DNA High-Resolution Kit (QIAGEN®, Montreal, QC, Canada) was used for high-resolution capillary electrophoresis DNA-fragment analysis to compare DNA-fragment fingerprints of the (GTG)5 PCR products from each isolate. One microliter (1 µl) of (GTG)5 rep-PCR product was diluted in 9 µl of QX DNA Dilution Buffer (QIAGEN®, Montreal, QC, Canada) and analyzed using the OM1200 preinstalled high-resolution method (QIAxcel ScreenGel v1.4.0, 2010-2014, QIAGEN®, Montreal, QC, Canada) with a 20ng/µl concentration of QX Size Marker 250bp – 8kb v2.0 (QIAGEN®, Montreal, QC, Canada).

2.6 Identification and typing of recovered STEC

2.6.1 Bacterial isolate species identification

In order to confirm that the *stx* positive mauve isolates determined to be genetically unique by (GTG)5 rep-PCR were *E. coli* and therefore true STEC, bacterial species identification was required. The archived TSB/CA-STEC isolate enrichments of genetically unique *stx* positive isolates were streaked to blood agar plates and incubated overnight (18-20 hours) at 37°C. Isolated colonies were picked and analyzed using a Vitek® Automated Bacterial Identification System (bioMerieux, Marcy-l'Étoile, France) per the manufacturer's instructions [180]. To confirm these bacteria still carried the *stx* genes, each identified strain was again analyzed via *stx1/stx2* qPCR.

2.6.2 Serotyping

All clonally unique, Vitek® confirmed, *E. coli* isolates were streaked onto TSA slants. Slants were incubated overnight (18-20 hours) at 37° C, sealed, and shipped to the *Escherichia coli* Reference Laboratory of the Public Health Agency of Canada's National Microbiology Laboratory at Guelph, Ontario to undergo serotyping. All isolates were serotyped by standard serum agglutination tests with antisera to all *E. coli* O antigens (O1-O188) and H antigens (H1-H56) [181].

2.7 Spatiotemporal Analysis

2.7.1 Seasonal Analysis of STEC occurrence

The Edwards's test is a commonly used statistical test of seasonality within epidemiological data that tests the null hypothesis of independent events uniformly distributed over twelve intervals against the alternate hypothesis of a simple harmonic trend with a peak and nadir six intervals apart [182,183]. Edwards' tests were performed using WINPEPI software (version 11.65, J.H. Abramson 2016) to identify any significant seasonal patterns in *stx* positivity during the study period, and the significance tests used Roger's modified procedure to help account for potentially small sample sizes [184–186]. Data was aggregated by month and both the number of monthly samples and the length of the month were adjusted for in the analysis. Edwards' tests were performed for the positivity rates of both submitted non-municipal drinking water samples and *E. coli*-contaminated non-municipal drinking water samples.

2.7.2 Cluster Analysis

Spatial scan statistics are used to detect statistically significant clusters of events that cannot be explained by the null hypothesis of spatial randomness [187]. The space-time scan statistic is a special case of the spatial scan statistic where a three-dimensional area defined by both space and time is scanned for the clustering of events. Clusters are detected by progressively scanning a window across space and/or time, noting the number of observed and expected events within the scanning window at each location, which are then compared to null hypothesis as defined by the underlying probability model being applied [188]. The clustering of *stx* positive drinking water samples in time, space and space-time was investigated using SaTScan™ software (version 9.4.2, M. Kulldorff and Information Management Services Inc., 2015, Boston, USA) to conduct sequential cluster detection tests [189]. To investigate purely temporal clustering, all *E. coli* positive drinking water samples with associated dates of collection (n=1910 *E. coli* positive drinking water samples of which 143 were *stx* positive from 2004/03/09 – 2016/07/31) were analyzed with Bernoulli model Kulldorff temporal scan tests [190]. This purely temporal statistic had a maximum cluster size of 50% of the study period, used 9999 model replications and had a P-value significance cut point of 0.05. All scan statistics used in this thesis which included temporal data, used the day as the base temporal unit and a 7-day temporal aggregation to account for known daily differences in ProvLab submission rates [191]. Purely spatial clustering was investigated with Bernoulli model Kulldorff spatial scan tests of all *E. coli* positive drinking water samples with accurate associated spatial data (n=1607 *E. coli* positive drinking water samples at

1120 locations of which 128 were *stx* positive) [190]. This statistic had a maximum cluster size of 50% of the population at risk, used 9999 model replications, used a circular scanning window and had a P-value significance cut point of 0.05. Spatial data associated with these drinking water samples was converted from the original Alberta Township Survey System (ATS) format outputted by the Data Integration for Alberta Laboratories (DIAL) tool to latitudinal and longitudinal coordinates of the centroid of the quarter section using Microsoft Excel templates created, and graciously provided by Jesse Invik [191,192]. ATS data allows for a location to be georeferenced to a resolution of 1 quarter section (~ 800m x 800m or 0.65km^2). Space-time clusters were investigated with space-time permutation model Kulldorff spatial scan tests of all *E. coli* positive drinking water samples with both associated spatial and temporal data (n=1580 water samples at 1109 locations of which 125 were *stx* positive) [193]. This space-time statistic had a maximum cluster size of 50% of the population at risk, used 999 model replications, used a circular scanning window and had a P-value significance cut point of 0.05. A maximum temporal window of 90-days was set for the space-time permutation model Kulldorff spatial scan tests to reduce computing time while still allowing for potential seasonal clusters to be detected. Space-time tests allowed for the spatial overlap of secondary clusters so long as they did not contain the centroid of a more likely cluster. Corresponding geospatial maps were created using ArcMap software (version 10.5, ESRI 2016, Redlands, CA, USA). Alberta and municipal base layer data was provided by the government of Alberta and used under an open government licence provided by Alberta Data Partnerships Ltd., through its partner AltaLIS Ltd [194].

CHAPTER 3: The Occurrence of *stx* Genes and Shiga Toxin-producing *Escherichia coli* (STEC) in Non-municipal Well-sourced Drinking Water Samples from Southern Alberta

3.1 Introduction

In Canada, some estimates place the number of STEC-related illnesses at nearly 22 000 cases per year, with only 1 in 10 cases being reported [24]. The province of Alberta has historically had *E. coli* O157:H7 disease rates more than double the national average and rates in southern Alberta are often among the highest in Canada [26,137]. The O157:H7 serotype makes up less than 50% of reported STEC infections in Alberta however, with non-O157 serotypes making the largest contribution to the burden of illness associated with these pathogens [43,45].

Although classically thought of as a foodborne pathogen, the importance of other routes of STEC infection are becoming better understood [195,196]. The outbreak of *E. coli* O157:H7/*Campylobacter jejuni* in Walkerton, Ontario that caused over 2300 illnesses and 7 deaths provides an unfortunate example of the importance of drinking water as a potential source of STEC infection in Canada [28]. Due to significant historical precedents such as the Walkerton outbreak, the O157:H7 serotype has dominated the focus of STEC research, however the importance of non-O157 STEC is increasingly appreciated within the field [19,32]. There have been over

380 non-O157 serotypes associated with human illness globally, but the most well-known belong to the 'big six' most clinically relevant serogroups in North America: O26, O45, O103, O111, O121, and O145 [19,37–39].

Despite the growing understanding of the importance of STEC related illness in Canada and across the globe, including the importance of drinking water as a potential pathway to disease, there has been very few studies investigating the occurrence and characteristics of these pathogenic bacteria in drinking water. The purpose of this portion of the project was to investigate the presence of *stx1* and *stx2* genes in submitted, Colilert® screened, *E. coli* positive, well water samples, as a proxy for STEC occurrence in non-municipal well-sourced drinking water systems across southern Alberta.

3.2 Results

3.2.1 *stx1* and *stx2* qPCR Analysis

During the study period, 95 675 non-municipal drinking water samples were submitted to ProvLab Calgary. Of these submitted samples, 2565 (2.7%) were determined to be *E. coli* positive via Colilert®, a positivity rate of 268.10 per 10 000 submitted non-municipal drinking water samples. Two thousand one hundred and ninety (2190) of these *E. coli* positive Colilert® enriched drinking water samples (85%) were archived and subsequently tested for *stx1* and *stx2* by qPCR analysis. Of these qPCR-tested samples, 1899 samples were included in the study (87% of archived samples and 74% of total *E. coli* positive Colilert® enriched drinking water

samples). Samples that were tested but not included in the study were removed as they were unable to be successfully linked to their ProvLab submission information. Seven percent (7%) of the *E. coli* positive Colilert® enriched drinking water samples (141/1899) were considered prospective samples having been submitted after May 21, 2015 and were processed accordingly. One-hundred and fifty-two (152) of the *E. coli* positive Colilert® enriched drinking water samples included in the study were qPCR positive for *stx1*, *stx2*, or both genes (Figure 2), resulting in an overall *stx* occurrence of 8% within included *E. coli* positive Colilert® enriched drinking water samples (152/1899) and an estimated *stx* positivity of 0.2% within voluntarily submitted non-municipal drinking water samples (152/95,675).

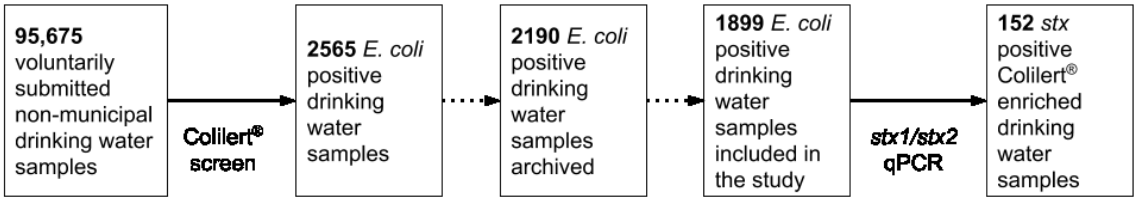

Figure 2 – Flow chart outlining the results obtained from drinking water samples submitted to ProvLab Calgary from March 2004 – July 2016 via the screening and identification of *stx* positive drinking water samples.

Of the 152 *stx* positive Colilert® enriched drinking water samples, 54 were *stx1* positive (35.5%), 53 were *stx2* positive (34.9%), and 45 were both *stx1* and *stx2* positive (29.6%) (Figure 3).

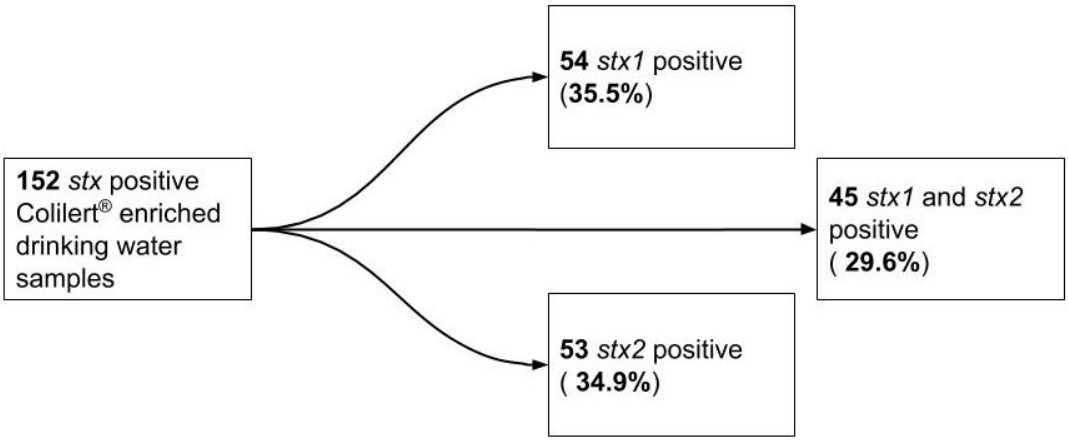

Figure 3 - *stx* positive Colilert® enriched drinking water samples separated by their respective *stx* positivity as determined by *stx1*/*stx2* qPCR.

3.2.2 Aggregated time-series analyses of *stx* positivity rates

Aggregated time-series of three distinct positivity rates were compared and analyzed across all years of the study period as well as within each individual month of the year in order to better understand temporal patterns of contamination of non-municipal drinking water with STEC.

3.2.2.1 *stx* positivity rates per submitted non-municipal drinking water samples

Shiga toxin gene (*stx*) positivity rates per 10 000 submitted non-municipal drinking water samples were calculated as a proxy for STEC occurrence rates within non-municipal drinking water sources in southern Alberta. Nine of the *stx* positive Colilert® enriched drinking water samples could not be linked to specific dates of collection, leaving 143 *stx* positive samples available for temporal analysis. Negative

binomial regression was used to test the relationship between time and annual *stx* positivity rates in submitted non-municipal drinking water samples (Table 6 and Table 7). As per convention, the first model uses the year 2010 as a referent as this is the year with the lowest number of positive samples [197]. Years without complete submission data (2004 and 2016) were not considered as referents.

Table 6 – Annual *stx* positivity rates per 10 000 submitted drinking water samples and negative binomial regression model of annual *stx* positivity rates in all non-municipal drinking water samples submitted to ProvLab Calgary compared to reference year 2010, March 2004 – July 2016

Year[c]	*stx* Positive Samples	Submitted Samples	*stx* Positivity Rate (per 10 000)	IRR	Wald Test p-value	95% CI
2004[a]	5	8548	5.85[a]	-	-	-
2005	70	14211	49.26[b]	8.39*	0.00	2.12 - 33.1
2006	10	9915	10.09	2.30	0.27	0.52 - 10.18
2007	8	9355	8.55	2.00	0.37	0.44 - 9.16
2008	9	7583	11.87	2.78	0.18	0.62 - 12.47
2009	4	7199	5.56	1.31	0.75	0.25 - 6.95
2010	3	6896	4.35	Referent	-	-
2011	9	6590	13.66	3.20	0.13	0.71 - 14.39
2012	4	6055	6.61	1.63	0.57	0.31 - 8.69
2013	10	6370	15.70	3.48	0.10	0.78 - 15.45
2014	5	5026	9.95	2.39	0.29	0.48 - 12.00
2015	4	5014	7.98	1.91	0.45	0.36 - 10.12
2016[a]	2	2913	6.87[a]	-	-	-

[a] Submission data incomplete for the full year

[b] Significant outlier (> upper fence of corresponding boxplot)

[c] Likelihood ratio Chi-squared (12) = 22.13 ; P-value = 0.036

* Indicates statistical significance, P-value < 0.05

The highest annual positivity rate per 10 000 submitted drinking water samples was in 2005 (49.3) and the lowest annual positivity rate for a full year of submitted

samples was 2010 (4.4). The statistical significance of the likelihood-ratio chi-squared test (75.9; P-value < 0.001) of the dispersion parameter alpha (0.81) suggests that the response variable is over-dispersed and is best described by a negative binomial distribution rather than a Poisson distribution. The likelihood ratio test for this model suggests that year had an effect on *stx* positivity in submitted drinking water samples and accounts for more variation in positivity than chance alone (P-value < 0.05). The year 2005 had a significantly higher *stx* positivity rate within submitted drinking water samples than the reference year of 2010, with the 2005 *stx* positivity rate in submitted drinking water samples 8.1 times the annual *stx* positivity rate in 2010 (P-value < 0.05). Model fit was poor when the June 2005 observation was included in the model (P-value < 0.05). Residual plots indicated that this observation was highly influential to the model. The model fit improved drastically when this observation was removed, however due to the biological significance of this observation the negative binomial model with June 2005 was reported.

In addition, when examining the annual *stx* positivity rates per 10 000 submitted drinking water samples using a box plot (Appendix A, Figure A 1), 2005 fell well-outside the upper fence and was determined to be a significant outlier. The 2005 *stx* positivity rate was over 5 times greater than the average rate of all other annual rates and over 2 times greater than the overall *stx* positivity rate per 10 000 submitted drinking water samples across the aggregate dataset.

The second annual *stx* positivity rate model used the year 2005 as a referent to investigate any statistically significant effects in the annual *stx* positivity rates in

submitted water samples when compared to the year with the highest number of *stx* positive samples (2005).

Table 7 – Annual *stx* positivity rates per 10 000 submitted drinking water samples and negative binomial regression model of annual *stx* positivity rates in all non-municipal drinking water samples submitted to ProvLab Calgary compared to reference year 2005, March 2004 – July 2016

Year[c]	*stx* Positive Samples	Submitted Samples	*stx* Positivity Rate (per 10 000)	IRR	Wald Test p-value	95% CI
2004[a]	5	8548	5.85[a]	-	-	-
2005	70	14211	49.26[b]	Referent	-	-
2006	10	9915	10.09	0.27*	0.01	0.10 - 0.74
2007	8	9355	8.55	0.24*	0.01	0.08 - 0.68
2008	9	7583	11.87	0.33*	0.03	0.12 - 0.91
2009	4	7199	5.56	0.16*	0.00	0.04 - 0.54
2010	3	6896	4.35	0.12*	0.00	0.03 - 0.47
2011	9	6590	13.66	0.38	0.06	0.14 - 1.05
2012	4	6055	6.61	0.19*	0.01	0.06 - 0.68
2013	10	6370	15.70	0.41	0.08	0.15 - 1.12
2014	5	5026	9.95	0.29*	0.03	0.09 - 0.92
2015	4	5014	7.98	0.23*	0.02	0.07 - 0.79
2016[a]	2	2913	6.87[a]	-	-	-

[a] Submission data incomplete for the full year

[b] Significant outlier (> upper fence of corresponding boxplot)

[c] Likelihood ratio Chi-squared (12) = 22.13 ; P-value = 0.036

* Indicates statistical significance, P-value < 0.05

Only 2011 and 2013 were found to not have significantly different (P-value < 0.05) annual *stx* positivity rates in submitted drinking water samples when

compared to 2005. The frequencies of *stx* occurrence for all other years were between 88% (2010) to 67% (2008) lower than that of 2005 (P-value < 0.05).

Negative binomial regression was also used to test the relationship between time and monthly *stx* positivity rates in submitted non-municipal drinking water samples (Table 8 and Table 9). Similar to the regression models testing annual *stx* positivity rates in submitted non-municipal drinking water samples, the time periods with both the lowest (November) and highest (June) monthly occurrence of *stx* positive samples were used as referents in respective negative binomial models.

Table 8 – Monthly *stx* positivity rates per 10 000 submitted drinking water samples and negative binomial regression model of monthly *stx* positivity rates in all non-municipal drinking water samples submitted to ProvLab Calgary compared to reference month of November, March 2004 – July 2016

Month[c]	*stx* Positive Samples	Submitted Samples	*stx* Positivity Rate (per 10 000)	IRR	Wald Test p-value	95% CI
Jan	2	5319	3.76	2.96	0.50	0.19 - 28.03
Feb	0	4862	0.00	< 0.01	0.99	0.00 - ∞
Mar	7	6439	10.87	7.75	0.08	0.78 - 61.92
Apr	5	6833	7.32	5.20	0.19	0.49 - 42.68
May	6	9503	6.31	4.47	0.22	0.43 - 36.18
Jun	64	11685	54.77	27.07*	0.00	3.20 - 203.81
July	29	12767	22.71	12.81*	0.02	1.47 - 97.46
Aug	13	10881	11.95	8.10	0.07	0.87 - 63.05
Sept	9	8869	10.15	6.31	0.12	0.65 - 49.89
Oct	4	7446	5.37	3.91	0.30	0.34 - 32.92
Nov	1	6385	1.57	Referent	-	-
Dec	3	4686	6.40	4.93	0.24	0.39 - 43.20

[c] Likelihood ratio Chi-squared (11) = 45.68 ; P-value < 0.000

* Indicates statistical significance, P-value < 0.05

The lowest monthly positivity rate per 10 000 submitted drinking water samples was February (0.0), without a single *stx* positive sample from the over 4000 submitted drinking water samples from that month. Negative binomial and Poisson regression models struggle to calculate accurate IRRs for values equal to zero, therefore statistical significance could not be ascribed to the February IRR for monthly analyses and this month was disregarded for comparisons. The statistical significance of the likelihood-ratio chi-squared test (68.9 ; P-value < 0.001) of the dispersion parameter alpha (0.61) suggests that the response variable is over-dispersed and is best described by a negative binomial distribution rather than a

Poisson distribution. The likelihood ratio test for this model suggests that month had an effect on *stx* positivity for submitted drinking water samples and accounts for more variation in positivity than chance alone (P-value < 0.05). The months of June and July had significantly higher *stx* positivity rates within submitted drinking water samples than the reference month of November (P-value < 0.05). The June *stx* positivity rate in submitted drinking water samples was 27.1 times the monthly *stx* positivity rate in November and the July *stx* positivity rate in submitted drinking water samples was 12.8 times the monthly *stx* positivity rate in November. An apparent peak in the monthly frequency of *stx* occurrence was observed in the early summer (June/July). Again, model fit was poor when the June 2005 observation was included in the model (P-value < 0.05), however due to the biological significance of this observation the negative binomial model with June 2005 was reported.

Table 9 – Monthly *stx* positivity rates per 10 000 submitted drinking water samples and negative binomial regression model of monthly *stx* positivity rates in all non-municipal drinking water samples submitted to ProvLab Calgary compared to reference month of June, March 2004 – July 2016

Month[c]	*stx* Positive Samples	Submitted Samples	*stx* Positivity Rate (per 10 000)	IRR	Wald Test p-value	95% CI
Jan	2	5319	3.76	0.09*	0.00	0.20 - 0.43
Feb	0	4862	0.00	< 0.01	0.98	0.00 - ∞
Mar	7	6439	10.87	0.27*	0.01	0.10 - 0.73
Apr	5	6833	7.32	0.18*	0.00	0.06 - 0.53
May	6	9503	6.31	0.16*	0.00	0.05 - 0.44
Jun	64	11685	54.77	Referent	-	-
July	29	12767	22.71	0.47	0.05	0.22 - 1.01
Aug	13	10881	11.95	0.29*	0.01	0.12 - 0.70
Sept	9	8869	10.15	0.22*	0.00	0.09 - 0.58
Oct	4	7446	5.37	0.13*	0.00	0.04 - 0.43
Nov	1	6385	1.57	0.04*	0.00	0.00 - 0.31
Dec	3	4686	6.40	0.16*	0.01	0.42 - 0.60

[c] Likelihood ratio Chi-squared (11) = 45.68; P-value < 0.000

* Indicates statistical significance, P-value < 0.05

When the month of June was used as a referent only July was found to not have significantly different (P-value < 0.05) monthly *stx* positivity rates in submitted drinking water samples. The frequencies of *stx* occurrence for all other months were found to be between 96% (November) to 71% (August) lower than that of June (P-value < 0.05).

Since the 2005 *stx* positivity rate per 10 000 submitted drinking water samples could be considered statistical outlier (Appendix A, Figure A 1), there was the possibility that this outlying year could bias results of the monthly time-series. Therefore, a time-series of monthly *stx* positivity rates per 10 000 voluntarily

submitted non-municipal drinking water samples and Poisson regression models testing monthly *stx* positivity rates in submitted drinking water samples, both with the 2005 sample data omitted, were calculated in order to investigate any potential changes to the overall pattern of monthly *stx* occurrence (Table 10 and Table 11). The time periods with both the lowest (November) and highest (June) monthly occurrence of *stx* positive samples were used as referents in respective Poisson models.

Table 10 – Monthly *stx* positivity rates per 10 000 submitted drinking water samples and Poisson regression model of monthly *stx* positivity rates in all non-municipal drinking water samples submitted to ProvLab Calgary compared to reference month of November, March 2004 – July 2016 (2005 data omitted)

Month[c]	*stx* Positive Samples	Submitted Samples	*stx* Positivity Rate (per 10 000)	IRR	Wald Test p-value	95% CI
Jan	1	4823	2.07	1.11	0.94	0.07 - 17.72
Feb	0	4298	0.00	0.00	0.99	0.00 - ∞
Mar	7	5821	12.03	6.43	0.08	0.79 - 52.24
Apr	5	6203	8.06	4.31	0.18	0.50 - 36.88
May	5	8469	5.90	3.16	0.29	0.37 - 27.01
Jun	20	9666	20.69	11.06*	0.02	1.48 - 82.41
July	15	10122	14.82	7.92*	0.05	1.05 - 59.97
Aug	9	8814	10.21	5.46	0.11	0.69 - 43.08
Sept	3	7468	4.02	2.15	0.51	0.22 - 20.64
Oct	4	6371	6.28	3.36	0.28	0.38 - 30.02
Nov	1	5345	1.87	Referent	-	-
Dec	3	4064	7.38	3.95	0.24	0.41- 37.93

[c] Likelihood ratio Chi-squared (11) = 34.92 ; P-value < 0.000

* Indicates statistical significance, P-value < 0.05

In general, the pattern of *stx* positivity rates when the 2005 results were omitted and when the 2005 results were included are similar with both series showing a summer peak, although this peak was notably reduced when the 2005 results were omitted (Figure 4). The lowest positivity rate per 10 000 submitted drinking water samples was still February (0.00) and the highest rate was again June (20.7). June, July and September did experience notable rate changes upon the omission of 2005 data, with the June *stx* positivity rate dropping from 54.8 to 20.7, the July rate dropping from 22.7 to 14.8 and the September rate dropping from 10.2 to 4.0. A deviance goodness-of-fit test for this model (102.1; Chi-squared (125) P-value = 0.77) suggests that a Poisson model is appropriate for this regression. The likelihood ratio test for this Poisson model suggests (P-value < 0.05) that month had an effect on *stx* positivity for submitted drinking water samples with 2005 data omitted. The months of June and July again had significantly higher *stx* positivity rates within submitted drinking water samples than the reference month of November, with the June *stx* positivity rate in submitted drinking water samples 11.1 times the monthly *stx* positivity rate in November and the July *stx* positivity rate in submitted drinking water samples 7.9 times the monthly *stx* positivity rate in November (P-value < 0.05). An apparent peak in the monthly frequency of *stx* occurrence was still observed in the early summer (June/July) with the 2005 data omitted.

Table 11 – Monthly *stx* positivity rates per 10 000 submitted drinking water samples and Poisson regression model of monthly *stx* positivity rates in all non-municipal drinking water samples submitted to ProvLab Calgary compared to reference month of June, March 2004 – July 2016 (2005 data omitted)

Month[c]	*stx* Positive Samples	Submitted Samples	*stx* Positivity Rate (per 10 000)	IRR	Wald Test p-value	95% CI
Jan	1	4823	2.07	0.10*	0.03	0.01 - 0.75
Feb	0	4298	0.00	0.00	0.99	0.00 - ∞
Mar	7	5821	12.03	0.58	0.22	0.25 - 1.37
Apr	5	6203	8.06	0.39	0.06	0.15 - 1.04
May	5	8469	5.90	0.29*	0.01	0.11 - 0.76
Jun	20	9666	20.69	Referent	-	-
July	15	10122	14.82	0.72	0.33	0.37 - 1.40
Aug	9	8814	10.21	0.49	0.08	0.22 - 1.08
Sept	3	7468	4.02	0.19*	0.01	0.06 - 0.65
Oct	4	6371	6.28	0.30*	0.03	0.10 - 0.89
Nov	1	5345	1.87	0.09*	0.02	0.01 - 0.67
Dec	3	4064	7.38	0.36	0.10	0.11 - 1.20

[c] Likelihood ratio Chi-squared (11) = 34.92 ; P-value < 0.000

* Indicates statistical significance, P-value < 0.05

The months of January, May, September, October, and November were all found to have significantly different (P-value < 0.05) monthly *stx* positivity rates in submitted drinking water samples when June was the referent and the 2005 data was omitted. The frequencies of *stx* occurrence for these five months were found to be between 91% (November) to 70% (October) lower than that of June (P-value < 0.05).

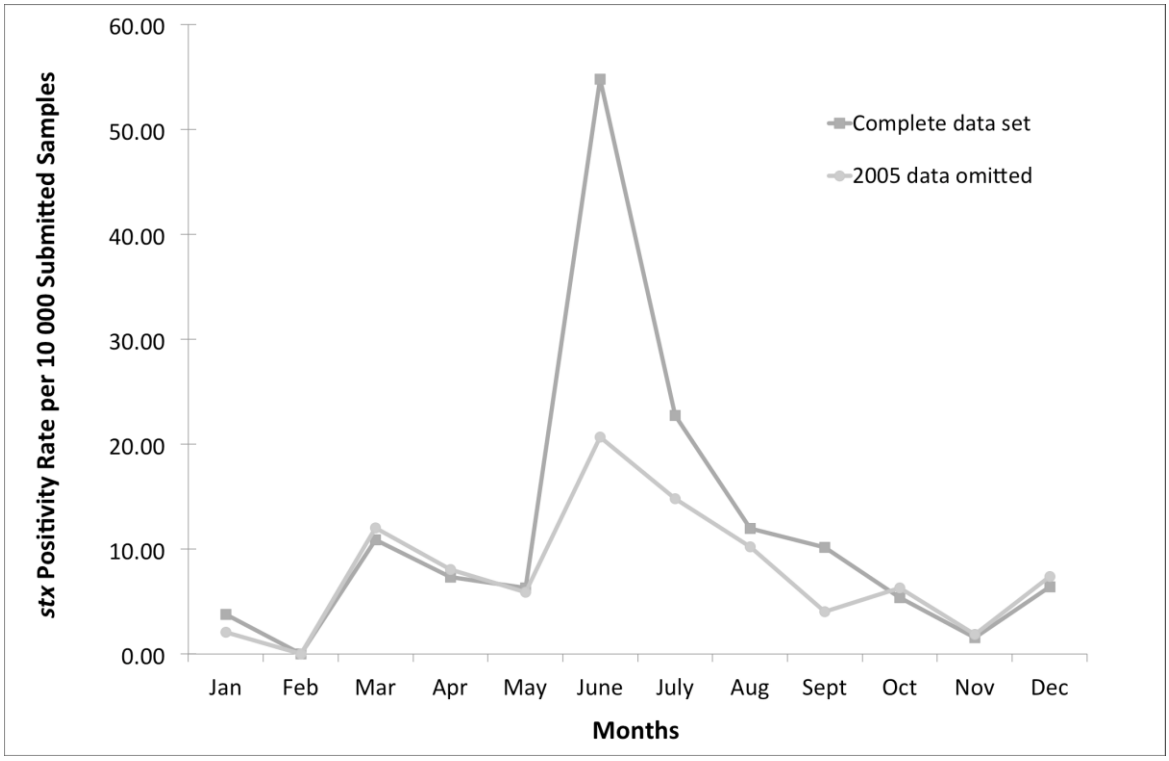

Figure 4 - Comparison of monthly *stx* positivity rates per 10 000 submitted drinking water samples aggregated from the complete data set, and the data set with the data from 2005 omitted.

3.2.2.2 *stx* positivity rates per *E. coli*-contaminated non-municipal drinking water samples

Shiga toxin gene (*stx*) positivity rates were also analyzed according to the number of *E. coli*-contaminated wells within the dataset (*stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples). In contrast to the time-series' described above for all non-municipal drinking water samples, this analysis was done to provide information on the risk of STEC within samples known to be

fecally contaminated and therefore already deemed to be at risk of causing enteric illness. The dominant factors affecting patterns of STEC occurrence in a well may be partially independent from the factors driving general fecal contamination of the well. For example, it could be argued that some water systems may be subject to routine, repetitive contamination with feces, but STEC occurrence within these fecal sources may be temporally-driven by other factors (i.e., increased seasonal periods of STEC shedding from animal hosts). For purposes of this thesis, it was hypothesized that similar patterns should be observed in respective *stx* positivity rates when grouped by month and year due to the generally accepted surrogate relationship between fecal indicators and enteric pathogens such as STEC. In addition, the *stx* positivity rates of *E. coli* positive drinking water samples provide a direct representation of the results of the *stx1*/*stx2* qPCR analysis of this project.

In this analysis, eighty-nine (89) of the *stx* negative *E. coli* positive Colilert® enriched drinking water samples could not be linked to specific dates of collection, leaving 1658 of these samples for temporal analysis. Importantly, due to the relatively rare occurrence of *stx* positive samples, as well as the relatively small sample sizes of *E. coli* positive Colilert® enriched drinking water samples, these results should be treated with a degree of caution.

Aggregated time-series of annual *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples as well as Poisson regression models are outlined in Table 12 and Table 13. The time periods with both the lowest (2010) and highest (2005) annual occurrence of *stx* positive samples were used as referents in respective Poisson models.

Table 12 – Annual *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples and Poisson model of annual *stx* positivity rates in all *E. coli* positive drinking water samples submitted to ProvLab Calgary compared to reference year 2010, March 2004 – July 2016

Year[c]	*stx* Positive Samples	*E. coli* Positive Samples	*stx* Positivity Rate (per 1000)	IRR	Wald Test p-value	95% CI
2004[a]	5	166	30.12[a]	-	-	-
2005	70	564	124.11	2.61	0.10	0.82 - 8.29
2006	10	125	80.00	1.74	0.40	0.4 - 6.32
2007	8	130	61.54	1.34	0.66	0.36 - 5.06
2008	9	100	90.00	1.94	0.32	0.52 - 7.15
2009	4	69	57.97	1.25	0.77	0.28 - 5.57
2010	3	68	44.12	Referent	-	-
2011	9	76	118.42	2.48	0.17	0.67 - 9.15
2012	4	61	65.57	1.43	0.64	0.32 - 6.41
2013	10	149	67.11	1.51	0.53	0.41 - 5.48
2014	5	100	50.00	1.09	0.91	0.26 - 4.54
2015	4	81	49.38	1.05	0.95	0.24 - 4.70
2016[a]	2	59	33.90[a]	-	-	-

[a] Submission data incomplete for the full year

[c] Likelihood ratio Chi-squared (12) = 23.98 ; P-value = 0.021

* Indicates statistical significance, P-value < 0.05

The highest annual positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples was in 2005 (124.1) and the lowest annual positivity rate for a full year of submitted samples was 2010 (44.2). The results of this time-series suggest that STEC occurrence in contaminating fecal sources may be variable from year-to-year. Similar to the annual *stx* positivity rates per 10 000 submitted drinking water samples, 2005 had the highest annual *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples. There were some notable

and unexpected differences from the annual *stx* positivity rates per 10 000 submitted drinking water samples, such as the relatively high 2011 *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples and the relatively low 2013 *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples. A deviance goodness-of-fit test for this model (116.7; Chi-squared (129) P-value = 0.77) suggests that a Poisson model is appropriate for this regression. The likelihood ratio test for this Poisson model suggests that year has an effect on *stx* positivity for *E. coli* positive Colilert® enriched drinking water samples and accounts for more variation in positivity than chance alone (P-value < 0.05). However, there were no years where annual *stx* positivity rates were determined to be statistically different (P-value < 0.05) when compared to 2010.

Table 13 – Annual *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples and Poisson model of annual *stx* positivity rates in all *E. coli* positive drinking water samples submitted to ProvLab Calgary compared to reference year 2005, March 2004 – July 2016

Year[c]	*stx* Positive Samples	*E. coli* Positive Samples	*stx* Positivity Rate (per 1000)	IRR	Wald Test p-value	95% CI
2004[a]	5	166	30.12[a]	-	-	-
2005	70	564	124.11	Referent	-	-
2006	10	125	80.00	0.67	0.23	0.34 - 1.29
2007	8	130	61.54	0.51	0.08	0.25 - 1.07
2008	9	100	90.00	0.74	0.40	0.37 - 1.49
2009	4	69	57.97	0.48	0.15	0.17 - 1.31
2010	3	68	44.12	0.38	0.10	0.12 - 1.22
2011	9	76	118.42	0.95	0.88	0.47 - 1.90
2012	4	61	65.57	0.55	0.25	0.20 - 1.51
2013	10	149	67.11	0.58	0.11	0.30 - 1.12
2014	5	100	50.00	0.42	0.06	0.17 - 1.03
2015	4	81	49.38	0.40	0.08	0.15 - 1.10
2016[a]	2	59	33.90[a]	-	-	-

[a] Submission data incomplete for the full year

[c] Likelihood ratio Chi-squared (12) = 23.98 ; P-value = 0.021

* Indicates statistical significance, P-value < 0.05

When 2005 was used a referent there were still no years where annual *stx* positivity rates were determined to be statistically different (P-value < 0.05).

One important consideration in this time-series analysis was also the comparison of prospective samples (not cryopreserved) versus retrospective samples (cryopreserved for up to 14 years) to identify any effect cryopreservation had on rates of STEC recovery in cultured Colilert® enriched drinking water samples. A comparative analysis revealed no appreciable difference between retrospective and prospective samples in terms of *stx* positivity rates per 1000 *E.*

coli positive Colilert® enriched drinking water samples. For prospective samples (the 13 months spanning June 2015 and July 2016) a *stx* positivity rate of 42.6 per 1000 *E. coli* positive Colilert® enriched drinking water samples was observed, which fell within the range of annual retrospective sample rates (42.3 [2010] – 110.4 [2005]) and was not significantly different from the mean annual rate of 67.9 for retrospective samples. Furthermore, there was no discernable pattern among annual *stx* positivity rates to suggest that time spent frozen had any influence on the recovery of STEC and detection of *stx* genes via qPCR analysis.

Poisson regression was also used to test monthly *stx* positivity rates in *E. coli* positive Colilert® enriched drinking water samples (Table 14 and Table 15). The time periods with both the lowest (November) and highest (June) monthly occurrence of *stx* positive samples were used as referents in respective Poisson models.

Table 14 – Monthly *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples and Poisson regression model of monthly *stx* positivity rates in all *E. coli* positive drinking water samples submitted to ProvLab Calgary compared to reference month of November, March 2004 – July 2016

Month[c]	*stx* Positive Samples	*E. coli* Positive Samples	*stx* Positivity Rate (per 1000)	IRR	Wald Test p-value	95% CI
Jan	2	21	95.24	4.00	0.26	0.36 - 44.12
Feb	0	22	0.00	0.00	0.99	0.00 - ∞
Mar	7	51	137.25	5.33	0.12	0.66 - 43.35
Apr	5	54	92.59	4.07	0.20	0.48 - 34.82
May	6	107	56.07	2.46	0.40	0.30 - 20.45
Jun	64	428	149.53	5.79	0.08	0.80 - 41.71
July	29	428	67.76	2.94	0.29	0.40 - 21.56
Aug	13	272	47.79	2.50	0.38	0.33 - 19.08
Sept	9	181	49.72	2.35	0.42	0.30 - 18.53
Oct	4	104	38.46	1.86	0.58	0.21 - 16.68
Nov	1	46	21.74	Referent	-	-
Dec	3	34	88.24	4.11	0.22	0.43 - 39.56

[c] Likelihood ratio Chi-squared (11) = 27.75 ; P-value = 0.004

* Indicates statistical significance, P-value < 0.05

The lowest monthly *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples was February (0.0) and the highest positivity rates were June (149.5) and March (137.3). The results of this time-series suggest that STEC occurrence in contaminating fecal sources may be variable from month-to-month. A deviance goodness-of-fit test for this model (112.9; Chi-squared (130) P-value = 0.86) suggests that a Poisson model is appropriate for this regression. The likelihood ratio test for this Poisson model suggests (P-value < 0.05) that month has an effect on *stx* positivity for *E. coli* positive Colilert® enriched drinking water samples. However, there were no months where monthly *stx* positivity rates were

determined to be statistically different (P-value < 0.05) when compared to November.

Table 15 – Monthly *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples and Poisson regression model of monthly *stx* positivity rates in all *E. coli* positive drinking water samples submitted to ProvLab Calgary compared to reference month of June, March 2004 – July 2016

Month[c]	*stx* Positive Samples	*E. coli* Positive Samples	*stx* Positivity Rate (per 1000)	IRR	Wald Test p-value	95% CI
Jan	2	21	95.24	0.69	0.61	0.17 - 2.82
Feb	0	22	0.00	0.00	0.99	0.00 - ∞
Mar	7	51	137.25	0.92	0.84	0.42 - 2.01
Apr	5	54	92.59	0.70	0.45	0.28 - 1.75
May	6	107	56.07	0.43*	0.05	0.18 - 0.98
Jun	64	428	149.53	Referent	-	-
July	29	428	67.76	0.51*	0.002	0.33 - 0.79
Aug	13	272	47.79	0.43*	0.01	0.24 - 0.78
Sept	9	181	49.72	0.41*	0.01	0.20 - 0.82
Oct	4	104	38.46	0.32*	0.03	0.12 - 0.88
Nov	1	46	21.74	0.17	0.08	0.02 - 1.25
Dec	3	34	88.24	0.71	0.56	0.22 - 2.26

[c] Likelihood ratio Chi-squared (11) = 27.75 ; P-value = 0.004

* Indicates statistical significance, P-value < 0.05

The months of May, July, August, September, and October were found to have significantly lower *stx* positivity rates for *E. coli* positive Colilert® enriched drinking water samples than the reference month of June, with the frequencies of *stx* occurrence for these five months between 68% (October) to 49% (July) lower than that of June (P-value < 0.05).

Similar to the monthly *stx* positivity rates per 10 000 submitted drinking water samples, there was an apparent early summer peak in positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples. However, the general pattern of monthly the *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples unexpectedly differed from the pattern of monthly *stx* positivity rates per 10 000 submitted drinking water samples, with various months having higher than average positivity rates such as January, and a potential spring peak in March and April. This finding suggests that the early spring (March/April) and early summer (June) are potentially important seasonal periods of increased risk of waterborne STEC infection due to the fecal sources impacting water quality in non-municipal drinking water systems having increased STEC positivity.

Again, the degree to which the 2005 *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples deviated from the other yearly rates, as well as the corresponding boxplot, which had the 2005 positivity rate positioned at the edge of the upper fence (Appendix A, Figure A 2), suggested that this year could bias results of the monthly time-series. Accordingly, monthly *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples as well as Poisson regression models testing monthly *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples were calculated with the 2005 sample data omitted (Table 16 and Table 17). Again, both November and June were used as referents in respective Poisson models.

Table 16 – Monthly *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples and Poisson regression model of monthly *stx* positivity rates in all *E. coli* positive drinking water samples submitted to ProvLab Calgary compared to reference month of November, March 2004 – July 2016 (2005 data omitted)

Month[c]	*stx* Positive Samples	*E. coli* Positive Samples	*stx* Positivity Rate (per 1000)	IRR	Wald Test p-value	95% CI
Jan	1	20	50.00	1.68	0.71	0.11 - 26.89
Feb	0	17	0.00	0.00	0.99	0.00 - ∞
Mar	7	49	142.86	4.25	0.18	0.52 - 34.51
Apr	5	53	94.34	3.19	0.29	0.37 - 27.30
May	5	95	52.63	1.78	0.60	0.21 - 15.23
Jun	20	197	101.52	2.89	0.30	0.39 - 21.54
July	15	270	55.56	1.84	0.55	0.24 - 13.96
Aug	9	226	39.82	1.67	0.63	0.21 - 13.14
Sept	3	112	26.79	1.02	0.99	0.11 - 9.79
Oct	4	80	50.00	1.87	0.57	0.21 - 16.76
Nov	1	35	28.57	Referent	-	-
Dec	3	30	100.00	3.58	0.27	0.37 - 34.43

[c] Likelihood ratio Chi-squared (11) = 11.82 ; P-value = 0.377

* Indicates statistical significance, P-value < 0.05

There were some notable changes in the monthly *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples when 2005 data was omitted from the time-series (Figure 5). Without the 2005 data, the apparent summer peak in positivity rates was reduced and the apparent spring peak became more pronounced. As a result of this change, the month of March became the time period with the greatest *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples (142.9). January also experienced a notable rate change upon the omission of 2005 data, falling from a *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples of 95.3 to 50.0. As mentioned

previously, due to limited sample sizes, the calculated rates per 1000 *E. coli* positive Colilert® enriched drinking water samples should be treated with a degree of caution, and in particular, the *stx* positivity rates of the winter months (November - February). These months all had less than 50 *E. coli* positive samples over the study period and rates could be easily inflated by single *stx* positive samples. For example, the difference between January rates when 2005 data was omitted represents the omission of a single *E. coli* positive sample that was also *stx* positive.

A deviance goodness-of-fit test for this model (93.8; Chi-squared (118) P-value = 0.95) suggests that a Poisson model is appropriate for this regression. The likelihood ratio test for this Poisson model (P-value = 0.38) does not reject the null hypothesis (P-value < 0.05) that month has no effect on *stx* positivity in *E. coli* positive drinking water samples with 2005 data omitted. When 2005 data was omitted, there were no significant differences (P-value < 0.05) between monthly *stx* positivity rates in *E. coli* positive Colilert® enriched drinking water samples when November was used as a referent.

Table 17 – Monthly *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples and Poisson regression model of monthly *stx* positivity rates in all *E. coli* positive drinking water samples submitted to ProvLab Calgary compared to reference month of June, March 2004 – July 2016 (2005 data omitted)

Month[c]	*stx* Positive Samples	*E. coli* Positive Samples	*stx* Positivity Rate (per 1000)	IRR	Wald Test p-value	95% CI
Jan	1	20	50.00	0.58	0.60	0.08 - 4.34
Feb	0	17	0.00	0.00	0.99	0.00 - ∞
Mar	7	49	142.86	1.47	0.38	0.62 - 3.47
Apr	5	53	94.34	1.10	0.84	0.41 - 2.94
May	5	95	52.63	0.62	0.33	0.23 - 1.64
Jun	20	197	101.52	Referent	-	-
July	15	270	55.56	0.64	0.19	0.33 - 1.25
Aug	9	226	39.82	0.58	0.17	0.26 - 1.26
Sept	3	112	26.79	0.35	0.09	0.10 - 1.19
Oct	4	80	50.00	0.65	0.43	0.22 - 1.90
Nov	1	35	28.57	0.35	0.30	0.05 - 2.58
Dec	3	30	100.00	1.24	0.73	0.37 - 4.17

[c] Likelihood ratio Chi-squared (11) = 11.82 ; P-value = 0.377

* Indicates statistical significance, P-value < 0.05

With June as the referent and 2005 data omitted, there were again no significantly different *stx* positivity rates between months (P-value < 0.05).

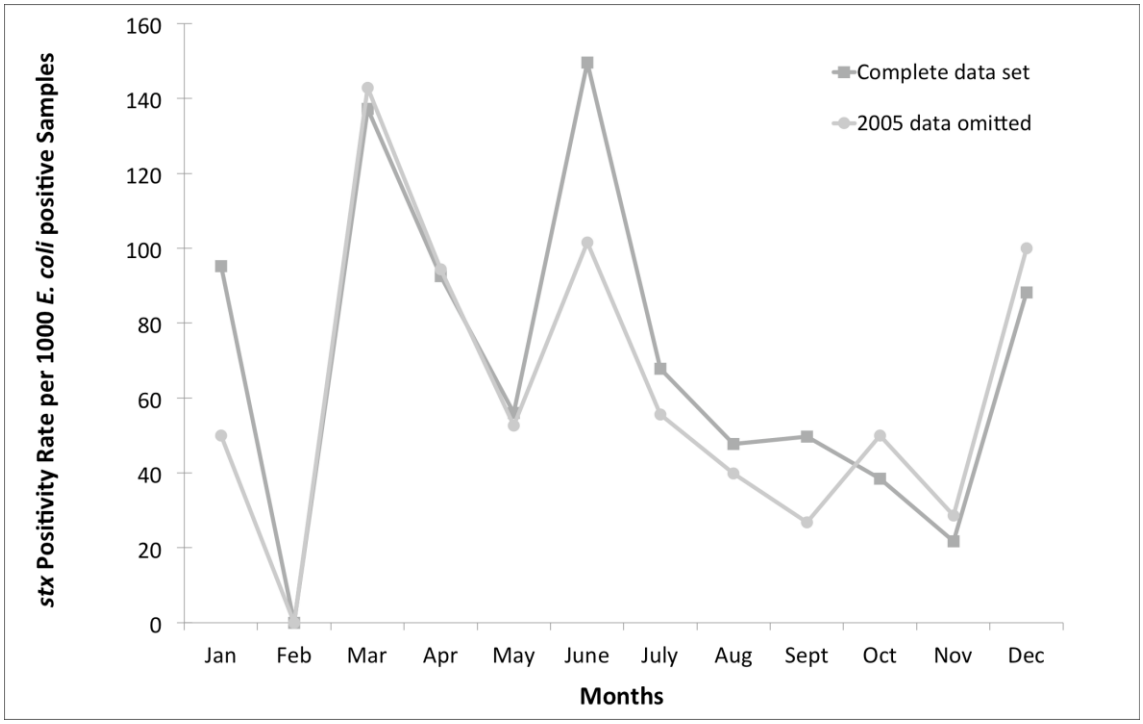

Figure 5 - Comparison of monthly *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples aggregated from a) the complete data set and b) the data set with 2005 data omitted

3.2.2.3 Comparing *stx* positivity rates with overall *E. coli* positivity rates in groundwater wells

While the focus of this project was not assessing *E. coli* contamination of non-municipal well-sourced drinking water (a subject that has been previously investigated by Invik [191]), the *E. coli* positivity rate per 1000 submitted drinking water samples was calculated to better understand the factors influencing the *stx* positivity rates of submitted drinking water samples during the study period. The calculated annual and monthly *stx* positivity rates of submitted drinking water

samples were equally dependent on the *E. coli* positivity rate of submitted drinking water samples and the *stx* positivity rate of *E. coli* positive Colilert® enriched drinking water samples, as only *E. coli* positive drinking water samples were tested for *stx* genes via qPCR and therefore *stx* positive samples were necessarily *E. coli* positive as well. To better understand the unexpected differences in the patterns of annual and monthly *stx* positivity rates of submitted drinking water samples and *E. coli* positive Colilert® enriched drinking water samples, the *E. coli* positivity rate per 1000 submitted drinking water samples was included in the comparisons due to it's direct relationship to both of these calculated *stx* positivity rates (Table 18 and Table 19). For *E. coli* positivity rates per 1000 submitted drinking water samples, deviations from the mean suggests months/years of lower or higher amounts of general fecal contamination in well-water sources, and for *stx* positivity rates per 1000 of *E. coli* positive Colilert® enriched drinking water samples, deviation suggests months/years of lower or higher amounts of STEC occurrence within *E. coli*-contaminated well-water sources.

Table 18 – Mean and annual *stx* and *E. coli* positivity rates for drinking water samples, 2005 – 2015

Year[a]	*stx* Positivity Rate per 1000 Submitted Samples[b]	*E. coli* Positivity Rate per 1000 Submitted Samples	*stx* Positivity Rate per 1000 *E. coli* Positive Samples
2005	4.93	44.61	124.11
2006	1.01	13.62	80.00
2007	0.86	14.75	61.54
2008	1.19	14.37	90.00
2009	0.56	10.14	57.97
2010	0.44	10.30	44.12
2011	1.37	12.90	118.42
2012	0.66	10.73	65.57
2013	1.57	24.96	67.11
2014	0.99	20.89	50.00
2015	0.80	16.95	49.38
Mean	**1.31**	**17.66**	**73.48**

[a] Years without complete data sets (2004, 2016) were omitted

[b] Positivity rate was converted to 1000 submitted samples for comparison

Annual *E. coli* positivity rates per 1000 submitted drinking water samples were relatively consistent during the study period, with three years above average (2005, 2013, 2014) and 2005 markedly so. For some years both the *E. coli* positivity rates per 1000 submitted drinking water samples and *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples mirrored one another, such as 2005, where both general *E. coli* contamination and STEC occurrence within *E. coli*-contaminated wells were above average, or 2010, where both general *E. coli* contamination and STEC occurrence within *E. coli*-contaminated wells were below average. Some years that experienced increased annual *stx* positivity per 1000 submitted drinking water samples were driven more by a single contributing positivity rate, such as 2013 with above average general *E. coli* contamination or

2011 with above average STEC occurrence within *E. coli*-contaminated wells. Conversely, some years that experienced decreased annual *stx* positivity per 1000 submitted drinking water samples were driven more by below average general *E. coli* contamination (2006) or by below average STEC occurrence within *E. coli*-contaminated wells (2015).

Table 19 - Mean and monthly *stx* and *E. coli* positivity rates for drinking water samples, 2004 – 2016

Month	*stx* Positivity Rate per 1000 Submitted Samples[a]	*E. coli* Positivity Rate per 1000 Submitted Samples	*stx* Positivity Rate per 1000 *E. coli* Positive Samples
Jan	0.38	4.32	95.24
Feb	0.00	4.52	0.00
Mar	1.09	9.01	137.25
Apr	0.73	8.63	92.59
May	0.63	11.89	56.07
June	5.48	42.11	149.53
July	2.27	35.80	67.76
Aug	1.20	26.19	47.79
Sept	1.02	21.42	49.72
Oct	0.54	14.50	38.46
Nov	0.16	7.36	21.74
Dec	0.64	7.90	88.24
Mean	**1.18**	**16.13**	**70.37**

[a] Positivity rate was converted to 1000 submitted samples for comparison

As expected, and similar to the statistically significant seasonality reported by Invik *et al.* [74], *E. coli* positivity rates per 1000 submitted drinking water samples demonstrated a smooth seasonal pattern with rates that peaked in the summer months. This comparison suggests that the previously highlighted rate differences

between the *stx* positivity rates of submitted drinking water samples and *stx* positivity rates of *E. coli* positive Colilert® enriched drinking water samples seen in January were offset by a below average *E. coli* positivity rates per 1000 submitted drinking water samples. In other words, the higher than average STEC occurrence within *E. coli*-contaminated well-water sources was offset by below average general *E. coli* contamination in well-water sources. For March, the greater than average STEC occurrence within *E. coli*-contaminated well-water sources were offset by overall below average general *E. coli* contamination in well-water sources during this month. This helps to explain that although the *stx* positivity rates per 1000 *E. coli* positive drinking water samples of both March and June are similarly high, the corresponding *stx* positivity rates of submitted drinking water samples for March and June are dissimilar. Notably, the summer months of June and July had high rates of both general *E. coli* contamination in well-water sources as well as STEC occurrence within *E. coli*-contaminated well-water sources, which corresponds to the increased *stx* positivity rates per 1000 submitted drinking water samples during the summer. The late summer and early fall (August and September) had above average general *E. coli* contamination in well-water sources that were offset by below average STEC occurrence within *E. coli*-contaminated well-water sources.

3.3 Discussion

This portion of the project set out to investigate the occurrence of STEC in non-municipal well-sourced drinking water systems in southern Alberta by determining the presence of the *stx1* and *stx2* STEC virulence genes in submitted

drinking water samples using a Colilert® screen to determine *E. coli* positivity and qPCR analysis to determine *stx* positivity.

It is useful to examine how the *stx* positivity results in well water from southern Alberta relate to the larger body of knowledge surrounding waterborne STEC. The results presented by Halabi *et al*. [114] are particularly notable due to similarities between this study and the current project. Similarities included: i) investigating non-municipal drinking water from mostly groundwater sources in areas with substantial agricultural activity; ii) investigating the occurrence of diverse STEC strains by using *stx1/stx2* PCR; and iii) determining distinct frequencies of occurrence for drinking water samples overall, as well as for *E. coli* positive drinking water samples. Shiga toxin gene (*stx*) positivity in well water was shown to be greater in southern Alberta than was found by Halabi *et al.* [114] in Upper Austria. The present study found *stx* positivity to be 0.2% within voluntarily submitted non-municipal drinking water samples as compared to 0.04% within all collected drinking water samples as determined by Halabi *et al*. [114]. In addition, the *stx* positivity within *E. coli* positive Colilert® enriched drinking water samples from the present study were determined to be 8.0% as compared to the 0.4% within collected *E. coli* positive drinking water samples from Upper Austria. Three other previous studies have investigated STEC in drinking water supplies. Ramteke and Tewari [113] examined STEC in rural drinking water in India, although the culture-based methods of detection used in this study prevent any accurate comparisons to the results of the present project. Both Won *et al.* [117] and Schets *et al.* [108] investigated STEC in private drinking water, however they only focused on *E. coli* O157:H7. In

rural Ohio, U.S.A., Won *et al.* [117] found *E. coli* O157:H7 in 3.9% of drinking water samples and Schets *et al.* [108] found *E. coli* O157:H7 in 2.7% of drinking water samples in the Netherlands. The STEC positivity of both studies were higher than the *stx* positivity of 0.2% from the present study and notably these occurrence rates did not include any non-O157 STEC present in the drinking water samples, which would likely have increased the rates even further. In summation, the positivity of *stx* in non-municipal drinking water sources in southern Alberta fell within the range of STEC occurrence defined by previous studies. Due to historical precedents of STEC-related disease and the availability of known risk factors (i.e., high agricultural density) in southern Alberta, relatively high rates of STEC occurrence were expected, and although *stx* positivity in this region was higher than the study by Halabi *et al*. [114], it was notably lower than the other studies specifically investigating drinking water. Methodological differences, such as the volume of water being investigated or the STEC being targeted, may explain some of the variability between these studies. In addition, the geographical location of each study area likely plays an important role in determining the STEC occurrence due to the importance of local risk factors in groundwater contamination.

Other relevant studies have investigated surface water rather than drinking water, but importantly, were either Canadian studies or had sampling locations from southern Alberta (Table 1). The frequency of occurrence of STEC in surface waters can still relate to the drinking water samples tested in the current study since contaminated runoff and surface water are potentially significant sources of contamination for groundwater wells [96,98,123,159]. STEC positivity in surface water is

likely to be greater than in groundwater, but the degree to which they differ will depend largely on the local hydrogeology of the area and specific traits of the pathogens present. For example, the type of soil in an area can have a substantial impact on the degree in which surface water, and any pathogens therein, can reach the groundwater table [95,98]. For pathogens, traits that aid the bacteria in surviving difficult environmental conditions and predation will play a considerable role in their ability to persist in surface and groundwater [88]. These Canadian studies found STEC positivity in surface waters to be quite high (9.2-31.8%) in comparison to the *stx* positivity results of the current study in groundwater (0.2%). As expected, studies investigating strictly *E. coli* O157:H7 had lower positivity (0.6-3.0%) than those investigating all STEC, yet the *E. coli* O157:H7 positivity in surface water was still notably higher than the *stx* positivity results in groundwater of the present study.

The extended duration of the study period reflected in this thesis enabled the aggregation of annual time-series to investigate yearly variation within *stx* occurrence, as well as monthly time-series to investigate seasonal variations within *stx* occurrence. Positivity rates were used in order to take into account any differences in the number of samples submitted during certain months or certain years, and were considered a proxy for STEC occurrence rates. Shiga toxin gene (*stx*) positivity rates of submitted drinking water samples provide evidence for assessing the risk of STEC contamination in non-municipal well-sourced drinking water more generally, while the *stx* positivity rates of *E. coli* positive Colilert® enriched drinking

water samples provide evidence for assessing the risk of STEC contamination within well-water considered to be fecally contaminated.

The two highest *stx* positivity rates per 10 000 submitted drinking water samples were in 2005 and 2013, which is notable due to the extreme precipitation events that occurred in southern Alberta during these two years, and that extreme weather is an associated risk factor for waterborne illness [81,158]. In June 2005, three major storms slowly made their way across southern Alberta causing substantial flooding in the region [198]. Areas in the southwestern portion of the province recorded close to 400mm of monthly precipitation, approximately four times the normal June average [106]. In June 2013, a shorter duration event occurred over 48 hours with intense precipitation rates across southwestern Alberta and upwards of 200-300mm of rain falling in the mountains and foothills, which along with warm temperatures, helped to melt an above-average snowpack, and together caused massive flash flooding across southern Alberta [199]. Two previous studies demonstrated that there were statistically significant increases in *E. coli* contamination rates of non-municipal groundwater systems across southern Alberta in 2005 and noted that this increase appeared to be the result of the extreme precipitation events that occurred in June of that year [74,106]. In line with these studies, it appears that increases in *stx* positivity rates within non-municipal groundwater systems across southern Alberta may also be associated with extreme precipitation events in this region, both in 2005 and 2013. Not only does it appear that these extreme precipitation events contributed to increased *stx* positivity rates during each respective year, but the large difference in rates between these two

years may be due to differences between the extreme precipitation events themselves. For instance, the event in 2005 was preceded by a dry spring across the south of the province and a snow pack that was the smallest in 40 years [198]. This precipitation event caused a steady influx of water, over an extended period of time, across a large geographical area, and into a very dry landscape, which engorged rivers and caused reservoirs to overflow. Previous research has suggested that protracted dry periods followed by heavy precipitation, similar to the environmental conditions experienced in 2005, may lead to increased overland movement of pathogens, substantial groundwater recharge, and higher pathogen loads in water sources [200–202]. Alternatively, the 2013 event was preceded by a wet spring and warm weather that had begun to melt an exceptionally large snowpack, both of which had contributed to higher than average groundwater levels [199]. This precipitation event caused a large, localized influx of water, over a short period of time, and into an already saturated landscape, which caused violent and geographically limited flash flooding in affected flood plains. Previous research has also suggested that heavy rainfall onto saturated landscapes, similar to the environmental conditions experienced in 2013, may increase the movement of pathogens into water sources [200,203,204]. The differences in how STEC were mobilized and transported from the landscape into groundwater during these two events may have played a role in determining *stx* contamination rates and suggests that the environmental conditions surrounding extreme precipitation events could influence the degree to which these events can impact the risk of *stx* contamination within groundwater sources.

The highest annual *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples was in 2005, which was not unexpected. In line with the generally held assumption of a positive association between fecal contamination and waterborne pathogen occurrence, the previously mentioned increase reported in *E. coli* contamination of groundwater systems in southern Alberta corresponding to the June 2005 extreme precipitation event indicated that there would likely be an increase in STEC occurrence. However, the unexpected differences between relative increases of *stx* positivity rates between *E. coli* positive Colilert® enriched drinking water samples and submitted drinking water samples in 2013 and 2011 challenge this assumption. It was hypothesized that the increased *stx* positivity rate per 10 000 submitted drinking water samples seen in 2013 would be mirrored by an increased *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples in the same year, yet this was not the case. Likewise, with the increased rate of *stx* positivity per 1000 *E. coli* positive Colilert® enriched drinking water samples in 2011, an increased rate of *stx* positivity per 10 000 submitted drinking water samples was expected, which again was not observed. This suggests that STEC occurrence in feces may be temporally variable from year to year and may exhibit patterns of seasonality (spring/summer peaks). This temporal variability in STEC occurrence is further complexed by the general increase in fecal-contamination of wells over the summer [74].

Due to nature of the Colilert® screen used for this study, only *E. coli* positive samples were tested for *stx* genes and therefore *stx* positive samples were necessarily *E. coli* positive samples, which ensured an equal contribution of both *E.*

coli positivity rates of submitted drinking water samples and *stx* positivity rates of *E. coli* positive Colilert® enriched drinking water samples when calculating *stx* positivity rates of submitted drinking water samples. Including the *E. coli* positivity rates of submitted drinking water samples in time-series comparisons allowed for greater understanding of which contributing rate (*E. coli* positivity of drinking water samples or *stx* positivity of *E. coli* positive drinking water samples) was driving any increases or decreases in the overall *stx* positivity rate of submitted drinking water samples. For example, in 2013 the increased *stx* positivity rate of submitted drinking water samples was driven by an above average number of samples contaminated with *E. coli* but with an average rate of STEC occurrence within these *E. coli* positive samples. Alternatively, in 2011 the above average rate of STEC occurrence within *E. coli* positive samples did not visibly influence the *stx* positivity rate of submitted drinking water samples because 2011 had a below average rate of *E. coli* contamination within groundwater sources that offset the increased STEC numbers. In 2005, it was both the above average rates of *E. coli* contamination and STEC occurrence within the *E. coli*-contaminated groundwater sources that contributed to the drastically higher *stx* positivity rate of submitted drinking water samples.

Seasonality is a well-established characteristic of human zoonotic enteric diseases and waterborne enteric illnesses in particular are influenced by ecological and climactic factors that will change throughout a calendar year, especially in temperate climates [128,160,205]. Monthly aggregated time-series of *stx* positivity rates helped to elucidate any notable patterns and variations between seasons or months of the year. The strong summer peak apparent in the monthly *stx* positivity rates per

10 000 submitted drinking water samples, as demonstrated by the significantly greater *stx* positivity rates of June and July outlined in negative binomial regression models, suggests a seasonal pattern of *stx* occurrence in non-municipal well-sourced drinking water sources and was in agreement with previous studies investigating STEC O157:H7 occurrence in surface waters of southern Alberta [119]. A similar summer peak in *stx* contamination within of *E. coli* positive Colilert® enriched drinking water samples also suggests a seasonal pattern of *stx* occurrence within fecally contaminated well water supplies, although the seasonal pattern within *E. coli*-contaminated water samples also contained a pulse of occurrence during the early spring (March-April). These multiple peaks imply a seasonal pattern of STEC contamination within fecally contaminated well water that is distinct from the seasonal pattern of general STEC contamination across non-municipal well-sourced drinking water.

Similar to the annual *stx* positivity rates comparison, monthly *E. coli* positivity rates of submitted drinking water samples were calculated to further explore the differences between monthly *stx* positivity of submitted drinking water samples and monthly *stx* positivity of *E. coli* positive Colilert® enriched drinking water samples. The distinct spring peak of STEC occurrence within the fecally contaminated groundwater sources is only marginally represented in STEC positivity within submitted drinking water samples, as this increased severity of STEC contamination in *E. coli* positive drinking water samples is offset by below average rates of *E. coli* contamination within groundwater sources. The statistically significant increase in the June and July rates of *stx* occurrence in submitted

drinking water samples was aided by the alignment of both above average STEC positivity within *E. coli*-contaminated groundwater sources and above average *E. coli* contamination of groundwater sources, highlighting a period where non-municipal drinking water sources would be at particular risk of STEC contamination.

Seasonal patterns of STEC occurrence remained visible upon omission of the outlying 2005 data despite the substantial influence this data had on *stx* positivity rates, suggesting that these seasonal patterns are endemic. For instance, the June *stx* positivity rate per submitted drinking water samples remained statistically greater than the majority of other monthly positivity rates even when 2005 data was omitted. Within *stx* positivity rates per *E. coli* positive Colilert® enriched drinking water samples the omission of the 2005 data had an even more notable influence on the seasonal patterns of STEC occurrence as March became the month with the highest *stx* positivity rate per 1000 *E. coli* positive Colilert® enriched drinking water samples due to the decrease in June *stx* positivity. The corresponding seasonal peak in the early spring became the more prominent of the two seasonal pulses of *stx* occurrence, further highlighting the distinct seasonal patterns of STEC within fecally contaminated drinking water samples and STEC within all submitted drinking water samples.

The results of this chapter provide some very important and noteworthy insights. Firstly, the peak summer pattern of *stx* occurrence in non-municipal drinking water supplies matches the seasonal pattern of reported STEC-related illness in Canada [58,128,130]. This finding brings into question whether contaminated

non-municipal drinking water sources may play an important role in STEC-related illness in Canada, an association not previously described. Additionally, while both the risk factors that influence the input of STEC into environmental water sources as well as the risk factors that influence the susceptibility of well-water systems to *E. coli* contamination are important to overall STEC occurrence within groundwater sources, these results suggest that the degree to which each of these risk factors contribute to the overall risk of STEC contamination at any one time not only varies, but follow seasonal patterns distinct from one another. This apparent seasonality in *stx* positivity outlined by the time-series analysis in this chapter necessitated more detailed analysis and further discussion - the topic of which is described in-depth in Chapter 5 of this thesis.

One of the main difficulties of studying pathogens in water is the often rare and sporadic pattern of their occurrence and consequently, having a large enough sample size to capture a meaningful amount of cases for analysis [59,206]. Centralized water testing at the Alberta ProvLab provided the capacity to access a very large number of water samples from across a vast geographic area in order to address this challenge. During the almost 14 years of sample collection, 95 675 drinking water samples were submitted and screened for the presence of *E. coli* by the ProvLab in southern Alberta, and from this screen 1899 *E. coli* positive samples were analyzed for STEC occurrence through the detection of *stx* virulence genes. Sample sizes of this magnitude are rare in the relevant literature (Table 1). For instance, the largest sample set from the relevant studies outlined previously was 2633 total water samples with 280 being *E. coli* positive [114]. The length of the study

period was another unique aspect of this sample set and was considerably longer than any study periods found in the related literature. The extended duration of the study period allowed for more robust examination of both seasonal patterns of *stx* occurrence, as well as annual frequencies of *stx* occurrence. However, even with such a large sample size, the ability to statistically evaluate the occurrence of STEC in groundwater over time was still limited. For the Poisson-based regression models used in this analysis, the smaller the number of positive samples, the larger the sample size of aggregate submitted or *E. coli* positive samples were needed to achieve adequate statistical power [168,207]. The regression models investigating *stx* positivity rates in *E. coli* positive drinking water samples likely did not have large enough sample sizes for the low rates of positivity, thereby preventing meaningful statistical comparisons between indicators. For example, the Poisson regression model of monthly *stx* positivity rates in all *E. coli* positive drinking water samples submitted to ProvLab Calgary with the reference month of June found that the positivity rate of October (68% lower than June, 4 *stx* positive samples and 104 *E. coli* positive samples) was statistically different, while the month of November, which had a lower positivity rate than October, (83% lower than June, 1 *stx* positive sample and 46 *E. coli* positive samples) was not determined to be statistically different (P-value <0.05).

The *stx* qPCR methods employed were highly sensitive and specific for the detection of both the *stx1* and *stx2* genes. The sensitivity of this method as a screen for STEC in drinking water allowed for the detection of low levels of STEC in large and diverse background populations of environmental microbes, within which STEC

may be missed using strictly culture-based detection methods. The specificity of this method lends a significant amount of confidence that a positive result is a true positive and that only the *stx* gene targets are being amplified. The *stx* qPCR assay distinguished between *stx1* and *stx2* genes, which thereby provided additional information about the characteristics and potential pathogenicity of the presumptive STEC within the drinking water samples. Clinical studies have shown that *stx2* is more commonly associated with human illness and more severe health outcomes, such as the development of HUS, than *stx1* [7–9]. The results of the current study show that a significant portion of the *stx* contaminated drinking water samples contain virulence gene profiles associated with increased pathogenicity, with 64.5% harboring the *stx2* gene. Shiga toxin 1 (*stx1*) genes still pose a serious risk of causing enteric illness and HUS however, and STEC harboring strictly *stx1* should not be disregarded [1,7]. Although these toxins play a key role in the development of STEC related illness, it is important to note that there are many other virulence factors that can significantly affect the pathogenicity of an STEC.

Another advantage of using *stx* qPCR for screening water samples for the presence of STEC is that this method is able to capture all STEC serotypes, not only *E. coli* O157:H7 and the other the 'big six' non-O157 STEC serogroups. The majority of waterborne STEC studies, and notably the studies done in Canada, have only investigated *E. coli* O157:H7 (Table 1). This restricted focus is a serious shortcoming in the field, particularly due to the increased understanding of the clinical importance of non-O157 STEC in both Alberta and Canada [118,208]. Quantitative PCR (qPCR) analysis is also able to detect STEC within drinking water samples that have

entered a viable but non-culturable (VBNC) state – a physiological state in which the organism may not be detected using culture-based methods, yet is still able to cause infection in a host [209]. Overnight TSB enrichments, such as the resuscitation step of the qPCR protocol, have been shown to revert waterborne VBNC STEC back into a culturable state, however if any pathogens had stayed in a VBNC state the low LOD_{95} of this assay would help ensure that these cells would still be detected despite any lack of growth during the protocol [101].

One important consideration with regards to the non-municipal drinking water samples used in this study was that they were not randomly selected and instead, were submitted voluntarily by those responsible for the quality of water being tested. Because of this passive sampling approach there is a possibility of volunteer bias and in this case, the volunteer bias could skew the results in multiple ways. For example, it could skew results towards the over-representation of problem drinking water systems as well overseers may be more likely to send in samples if there is a history of poor water quality or if there is a current water quality issue they are trying to address. Conversely, the volunteer bias could skew results towards the over-representation of well-maintained and low-risk drinking water systems, where overseers who routinely test water quality as part of a comprehensive maintenance program are submitting water samples more frequently. Voluntary sampling may also lead to certain geographic areas being more frequently sampled from then others. For instance, if drinking water samples from higher or lower-risk areas are submitted more consistently than other regions across southern Alberta, results could be skewed in the direction of the risk profile

of these areas. Inherent in these issues presented by the voluntary submission of water samples, is that samples from the same water source can be submitted multiple times over the study period. There was no definitive way to ascertain how many of the submitted samples were from the same water system using the information provided with the sample however. Multiple submissions may exacerbate the potential biases mentioned previously. While passive surveillance does have fundamental limitations, it is an effective and economical way to increase the power of a study, particularly when investigating baseline occurrence rates of rare pathogens or disease, or investigating temporal trends over extended study periods [74,206].

Due to the extremely large number of total samples submitted during the study period, Colilert® was used as a screen to reduce the total amount of water samples that needed to be analyzed via *stx* qPCR. As such, only *E. coli* positive Colilert® samples were analyzed by *stx* qPCR, having assumed that STEC would only be found in *E. coli* positive samples. Although most *E. coli* and STEC strains possess the β-galactosidase and β-glucuronidase activity needed to induce an *E. coli* positive Colilert® result, some STEC strains, such as *E. coli* O157, do not typically possess β-glucuronidase. For instance, a pure culture of *E. coli* O157 would cause Colilert® media to turn yellow but the media would not fluoresce, leading to a false negative result for the presence of *E. coli*. However, this Type II error is likely rare as the probability that a single strain of *E. coli* would be responsible for the color change and fluorescence in Colilert® when testing microbial water quality is low, and it is equally unlikely that only β-glucuronidase negative *E. coli* strains would be present

in a fecally contaminated water sample [210]. This *E. coli* screening process is widely accepted and is part of the certified water testing protocol developed and implemented by the ProvLab to ensure safe drinking water across Alberta [211].

Several qualifications exist when interpreting *stx* screening as a proxy for STEC occurrence in drinking water. First, since qPCR analysis can detect DNA from viable and non-viable microbes, *stx* positivity may indicate the presence of a cell unable to infect a host. However, due to enrichment steps during the *E. coli* and *stx* screening protocols the detection of non-viable cells was minimized, and even if non-viable cells were being detected, these cells are still representative of a public health risk since these non-viable cells were likely viable at some point within the water source. Alternatively, it is possible that the *stx* genes being amplified in the Colilert® sample are from free *stx*-encoding bacteriophages or from background host cells infected during *stx* phage replication, rather than actual STEC itself. If this were the case, there still exists a risk of infection, as these phages can interact with other *E. coli* in the water samples or in the gut, and through phage-mediated transduction of the *stx* genes, potentially give rise to viable STEC. The deadly 2011 outbreak STEC O104:H4 strain that caused almost 4000 illnesses in Germany provides an example of the potential consequences of other *E. coli* pathotypes integrating environmentally persistent *stx*-encoding bacteriophages [15]. Finally, the *stx1*/*stx2* qPCR assay used for this project, like many *stx*-specific PCR assays, was unable to detect the *stx2f* subtype [212–215]. Although historically *stx2f* has been rarely associated with human cases of STEC-related illness, STEC encoding this subtype are believed to be an emerging pathogen [214–217]. Thus far, the primary reservoir of *stx2f*-

encoding STEC are pigeons, and therefore the impact of including this subtype when analyzing rural groundwater in Alberta may be limited, however using a qPCR assay able to detect this subtype in the future would allow for a more complete estimation of STEC contamination in non-municipal drinking water sources [212,218].

CHAPTER 4: The Isolation and Recovery of Shiga Toxin-producing *Escherichia coli* (STEC) from Non-municipal Well-sourced Drinking Water Samples Using CHROMagar™ STEC Agar

4.1 Introduction

Although *E. coli* O157 has dominated much of the STEC-related research activities over the last several decades, the growing understanding of the importance of non-O157 STEC has led to the development of methods specifically targeted for the identification these bacteria [169,208,219,220]. While genetic tests have a number of advantageous characteristics, culture-based methods remain the gold standard for microbial identification due the capacity to isolate bacteria for further characterization, and as such, there are a number of selective agars for STEC [172,208]. However, there is currently no best-method for the detection and isolation of all STEC serotypes due to the well-recognized difficulties of differentiating between STEC and other *E. coli* [170]. These difficulties stem from the lack of a single, consistent, and determinable phenotypic trait that is specific to this *E. coli* pathotype [118,169]. Little is known about the performance of clinically validated agars for the detection and isolation of waterborne STEC, a group of pathogens that is likely to include a diverse repertoire of STEC strains from animals as well as humans. Furthermore, the background microflora in environmental samples is far different than clinical

samples (i.e., feces) and little is known about how these microbes may grow on selective agars developed for clinical samples.

STEC antigen typing (i.e., serotyping) is done to determine the genetic relatedness between known pathogens, which helps to increase our understanding of STEC epidemiology, allows for more effective public health interventions, and is an essential aspect of STEC surveillance [208]. As mentioned previously, certain STEC serotypes have been more commonly linked with serious human illness, such as O157:H7 and the 'big six' non-O157 serogroups. In Canada, the six most common clinical non-O157 serogroups are O26, O103, O111, O117, O121, and O145, slightly different from the 'big six' of the United States, with O117 included rather than O45 [221]. These six serogroups accounted for over 50% of the non-O157 STEC clinical isolates reported to the Public Health Agency of Canada from 1998-2012. Within Alberta, the most common clinically reported serogroups are O157, O26, O111, and O121 [40].

The purpose of this portion of the study was to isolate STEC from *E. coli* positive private well water that had screened positive for *stx1* and/or *stx2* genes by qPCR (Chapter 3). This was done in order to provide the most accurate representation possible of STEC occurrence in non-municipal drinking water sources across southern Alberta. As a result of this analysis, the performance of a STEC-specific clinically validated media (CHROMagar™ STEC agar) to selectively isolate waterborne STEC was also evaluated.

4.2 Results

4.2.1 STEC Isolation

Each of the 152 *stx* positive Colilert® enriched drinking water samples described in Chapter 3 were plated onto CHROMagar™ STEC agar (CA-STEC) and mauve-coloured isolates were collected as presumptive STEC colonies. Mauve colony colour is the morphological characteristic specified by the manufacturer to indicate a colony of STEC and the majority of non-O157 STEC are expected to fluoresce when exposed to UV light. Additionally, a selection of non-typical isolates (i.e., non-mauve) were also collected in order to verify that environmentally-derived strains of STEC would only appear as mauve colonies on CA-STEC plates.

In an original screening protocol, CA-STEC plates were incubated at 37°C after enrichment in TSB broth. However, after a preliminary evaluation of this protocol (29 samples), two notable challenges of this method were revealed: a) there was often significant bacterial background growth that prevented the isolation of colonies; and b) despite qPCR screening having identified these samples as *stx* positive, in many cases isolated mauve colonies were determined to be *stx* negative or were not present at all. Upon further investigation, an increase in the incubation temperature to 44°C reduced background growth and led to greater isolation of distinct mauve colonies on CA-STEC plates for some samples. This increase in incubation temperature also repeatedly improved the likelihood of mauve colonies forming on the plate (Figure 6). However, in some cases the opposite effect occurred, where a 44°C incubation resulted in few or no mauve colonies but re-

plating samples at 37°C resulted in greater growth of mauve colonies (Figure 7). It was also observed on occasion that culturing samples at 37°C resulted in the growth of only *stx* negative mauve colonies, but when incubated at 44°C, *stx* positive mauve colonies were isolated from these same samples. These results suggested that different environmental isolates of STEC might have distinct optimal growth temperatures.

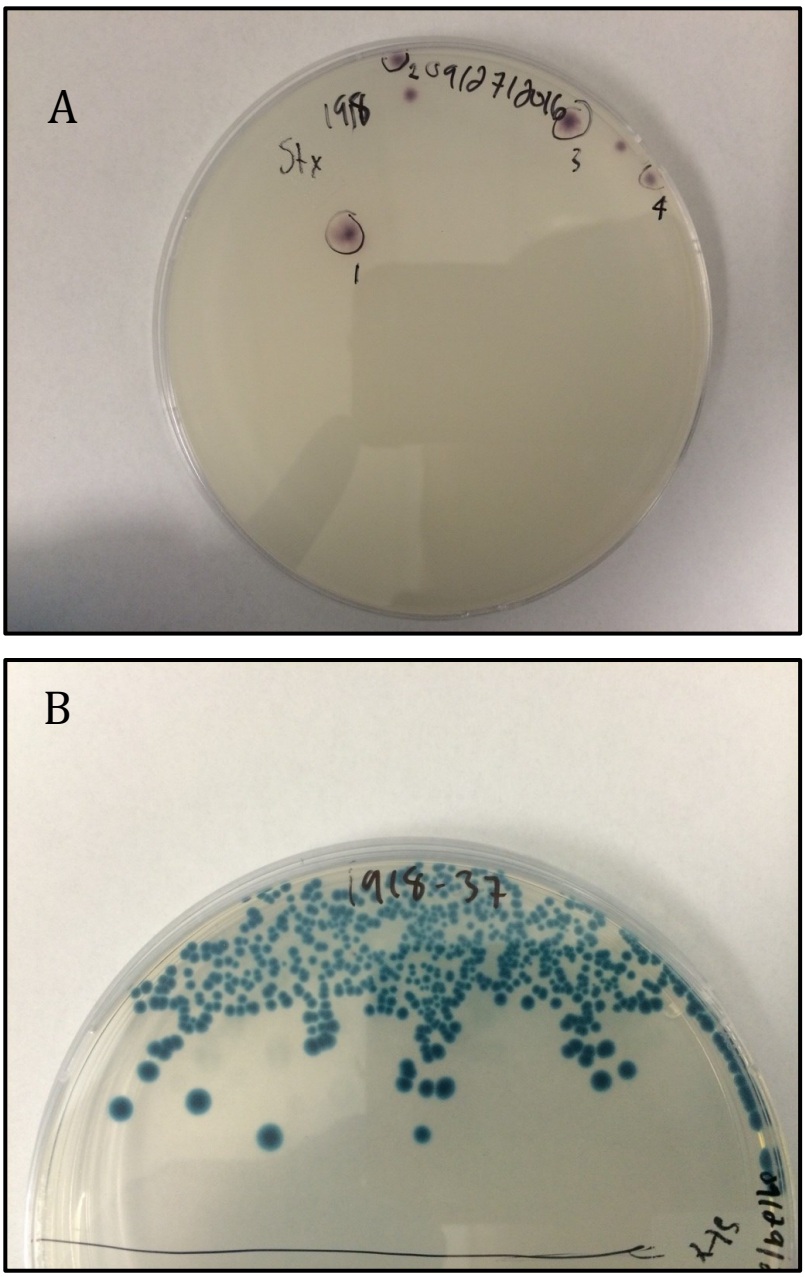

Figure 6 – An example of the effect that increased incubation temperature had on improving the isolation of waterborne STEC on CHROMagar™ STEC agar. Pictures depict overnight growth of 100µl of MAC enriched *stx* positive Colilert® drinking water samples on CHROMagar™ STEC agar (sample #1918), incubated at 44° C (A) or 37° C (B). In panel A, four mauve colonies (circled) were selected and confirmed as *stx* positive colonies via *stx1*/*stx2* qPCR.

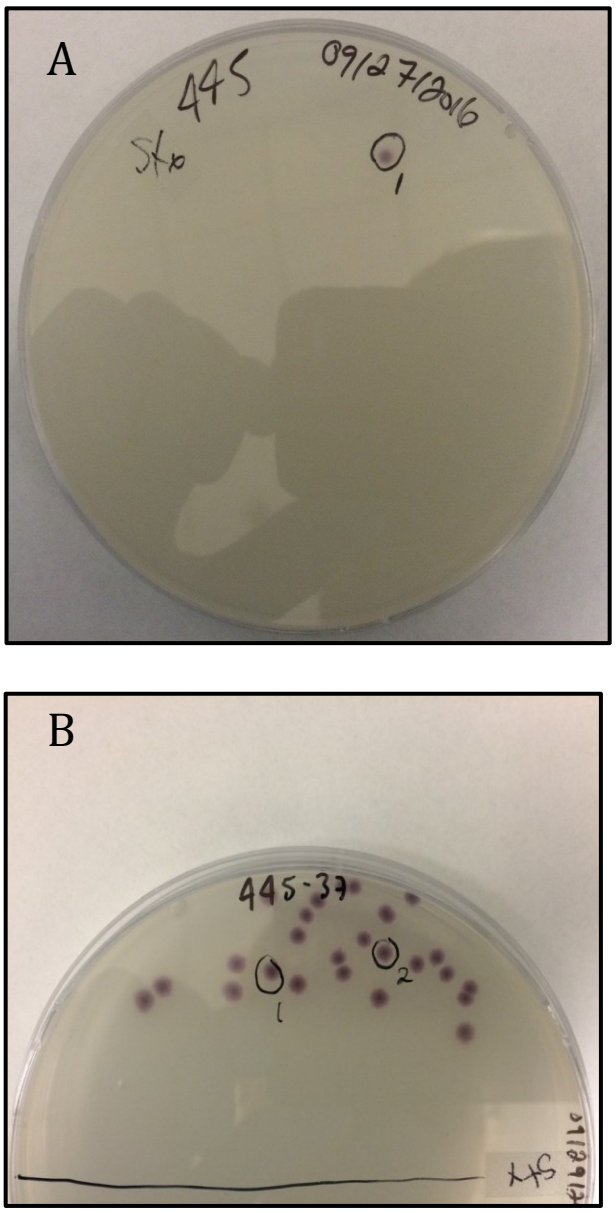

Figure 7 - An example of the effect that decreased incubation temperature had on improving isolation of waterborne STEC on CHROMagar™ STEC agar. Pictures depict overnight growth of 100µl of MAC enriched *stx* positive Colilert® drinking water samples on CHROMagar™ STEC agar (sample #445), incubated at 44° C (A) or 37° C (B). A single mauve coloured colony (circled) from the plate represented in Panel A was isolated and subsequently confirmed as *stx* negative via *stx1*/*stx2* qPCR. Two mauve coloured colonies (circled) from the plate depicted in Panel B were isolated and subsequently confirmed as *stx* positive colonies via *stx1*/*stx2* qPCR.

Samples with high amounts of background flora often inhibited the detection of viable STEC. For example, in sample #1072 (Figure 8), no mauve colonies were visible amongst the background flora due to overgrowth and the STEC present would have been undetected if it weren't for spots of fluorescence under UV light. Picks were taken from the areas of fluorescence and re-plated, resulting in isolated mauve colonies that were confirmed *stx* positive. In many instances however, the entire 'lawn' of background flora would fluoresce, preventing identification in this fashion (Figure 9 B). For some samples, the background growth produced 'lawns' of only mauve colonies, yet none of the colonies picked from these plates were confirmed as *stx* positive (Figure 9 A). It is possible that viable STEC may have existed in these samples where background flora was phenotypically similar to STEC on CA-STEC agar (mauve and/or fluorescent colonies), but the STEC were unable to be detected due to overgrowth of the non-STEC bacteria.

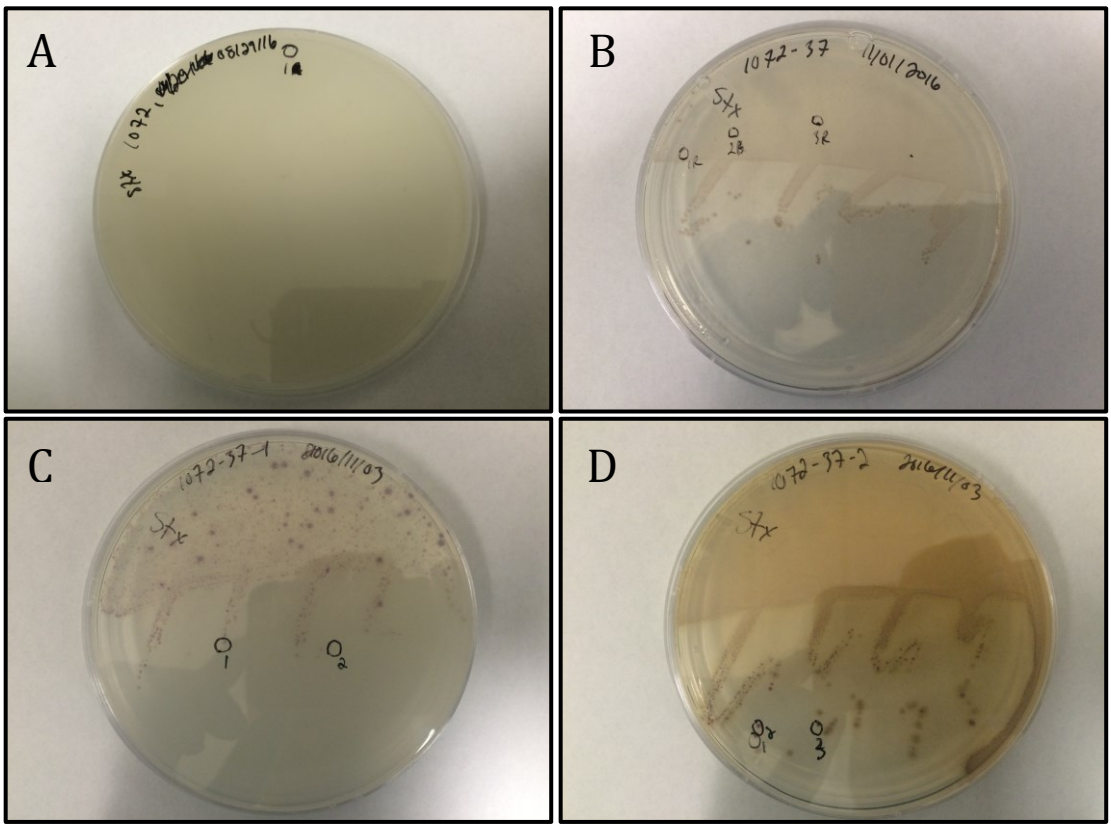

Figure 8 – An example of background growth inhibiting the detection of STEC on CHROMagar™ STEC agar. Pictures depict overnight growth of 100µl of MAC enriched *stx* positive Colilert® drinking water samples on CHROMagar™ STEC agar (sample #1072), incubated at 44° C (A) or 37° C (B,C,D). A single mauve coloured colony (circled) from the plate represented in Panel A was isolated and subsequently confirmed as *stx* negative via *stx1*/*stx2* qPCR, therefore the isolation procedure was attempted again at 37° C (Panel B). Only grey colonies were present upon visual inspection of the plate represented in Panel B, however there were three spots (circled) of UV fluorescence which were labeled, picked and re-plated for isolation at 37° C. The plate represented in Panel C is the re-plated pick 1R from Panel B from which two mauve colonies (circled) were isolated and subsequently confirmed as *stx* positive via *stx1*/*stx2* qPCR. The plate represented in Panel D is the re-plated pick 2B from Panel B from which three mauve colonies (circled) were isolated and subsequently confirmed as *stx* positive via *stx1*/*stx2* qPCR. The re-plate of pick 3R from Panel B did not result in the growth of any mauve colonies and is not pictured.

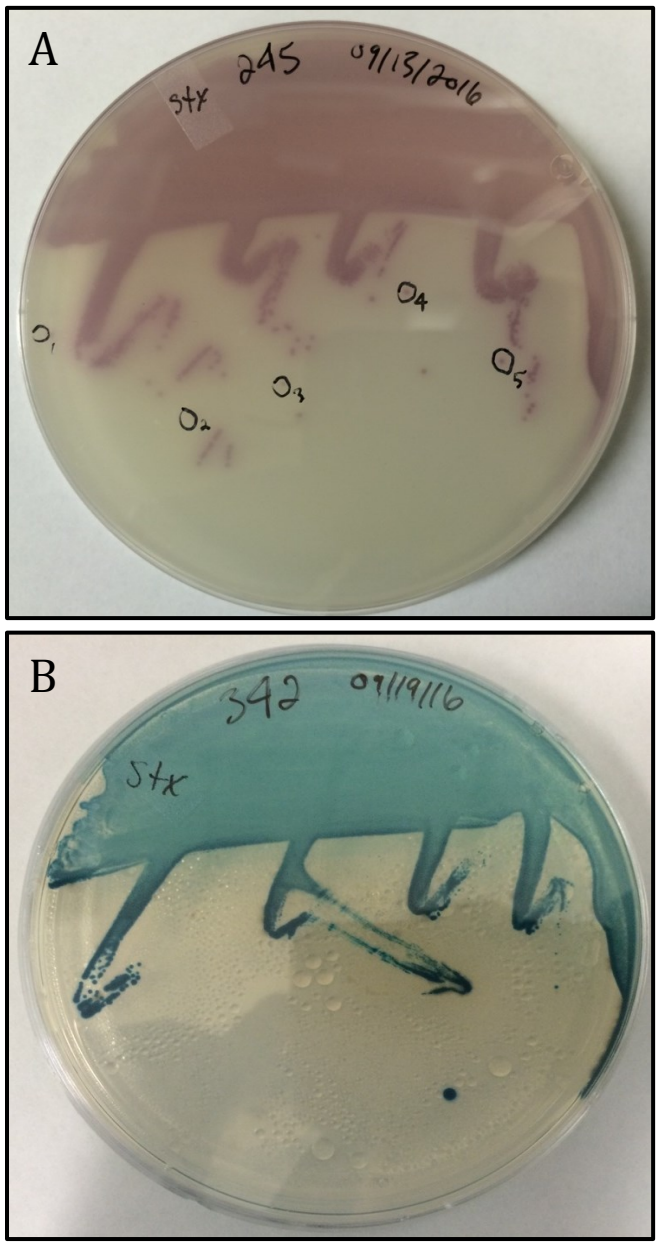

Figure 9 – Examples of background growth inhibiting the detection of STEC on CHROMagar™ STEC agar. Pictures depict overnight growth of 100µl of MAC enriched *stx* positive Colilert® drinking water samples on CHROMagar™ STEC agar (sample #245 and #342), incubated at 44° C. Five mauve coloured colonies (circled) from the plate represented in Panel A were isolated and subsequently confirmed as *stx* negative via *stx1*/*stx2* qPCR. No mauve coloured colonies were present upon visual inspection of the plate represented in Panel B and all visible colonies were UV fluorescent.

Bacterial competition also had potential effects on STEC growth during the enrichment phase of the protocol, which was observed during the process of troubleshooting the original CA-STEC isolation protocol. Aliquots of TSB culture broth removed at 2 hours and 18 hours of enrichment were analyzed by qPCR to monitor changes in *stx* concentrations over time and ensure that *stx* genes were still present in enrichment solutions prior to plating. As an example, sample #346, considered to have low level growth of presumptive STEC (i.e., 130 *stx* gene copies/5 ul of enriched TSB sample from the qPCR screening protocol), was seeded into TSB and after 2 hours of enrichment only 5 *stx* gene copies/5 ul was detected. After 18 hours of enrichment only 8 *stx* gene copies/5 ul was detected, demonstrating negligible growth of the STEC during enrichment. Due to the small number of potential STEC bacteria in the enrichment media, the low number of mauve colonies detected upon plating was not surprising (Figure 10 A). The four mauve colonies that could be seen in the lawn of background flora were subsequently picked and re-plated. From these four re-plated mauve picks, a single *stx* positive mauve colony was isolated (Figure 10 B). While sample #346 is an example where, despite poor growth during enrichment, the viable STEC present in the drinking water sample could still be recovered on CA-STEC, it is possible that increased competition from background flora prevented viable STEC in other samples from being detected on CA-STEC.

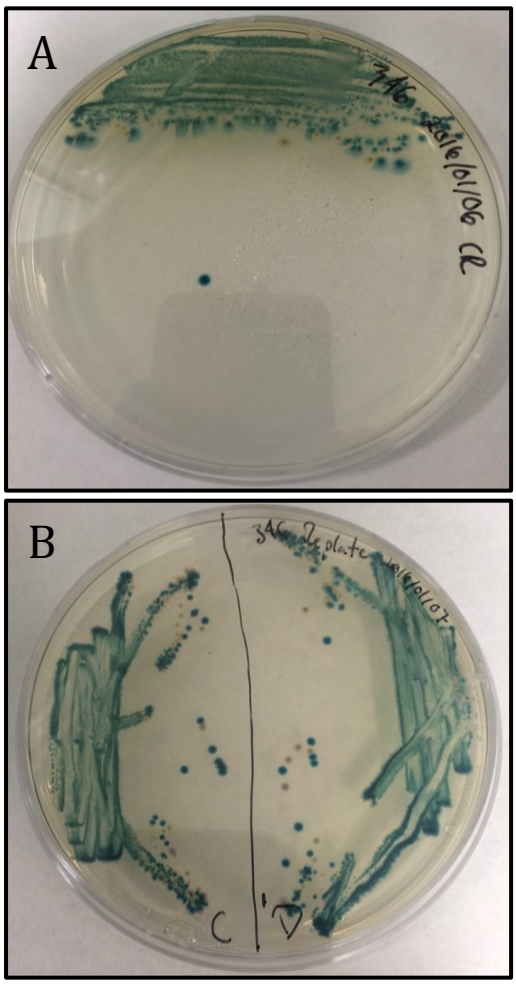

Figure 10 – An example of STEC recovery despite inhibited growth due to background competition during the enrichment phase of the CA-STEC isolation protocol. Pictures depict overnight growth of 100µl of MAC enriched *stx* positive Colilert® drinking water samples on CHROMagar™ STEC agar (sample #346), incubated at 37° C. Four non-isolated mauve coloured colonies were present upon visual inspection of the plate represented in Panel A and were picked and re-plated for isolation at 37° C (picks are not marked in the photo). The plate represented in Panel B has both the re-plated non-isolated mauve coloured colonies #3 (left side of the plate, marked C) and #4 (right side of the plate, marked D) from Panel A. A single mauve coloured colony was isolated from the right side of the plate represented in Panel B and subsequently confirmed as *stx* positive via *stx1*/*stx2* qPCR. No mauve colonies were detected in the re-plates of pick #1 (not pictured), pick #2 (not pictured), or pick #3 from in Panel A.

Overall these data demonstrated that temperature and enrichment were key variables for STEC isolation from CA-STEC plates, largely due to: i) variations in STEC optimal growth temperatures, and ii) competitive background growth. To try and address these challenges, the original protocol for STEC isolation was subsequently revised to include enrichment in a selective broth (MacConkey [MAC]) to reduce competitive growth during the enrichment phase, followed by plating on CA-STEC and incubation at 44°C. Forty-four degrees centigrade (44°C) was adopted as the initial CA-STEC incubation temperature in order to reduce background growth and enhance the selective isolation of mauve colonies. In samples where few or no mauve colonies were observed at 44°C, the sample was subsequently re-plated and incubated at 37°C to identify any STEC isolates favoring growth at lower temperatures.

Using the revised protocol, 850 presumptive mauve STEC isolates and an additional 55 non-mauve isolates were archived. These 905 isolates were tested for the presence of *stx1* and *stx2* by colony qPCR to confirm the presence or absence of STEC-related toxin genes in the isolates. Two hundred and fifty-three (253) of the mauve/non-mauve isolates were positive for *stx* via qPCR. These 253 isolates originated from 59 of the 152 (38.8%) *stx* positive Colilert® drinking water samples. Surprisingly, none of the mauve isolates represented in the other 93 *stx* positive Colilert® drinking water samples were found to be *stx* positive, even though some samples were found to have high *stx* copy numbers per qPCR reaction during screening (i.e., >800,000 copies/5 µl), suggesting that a proportion of environmentally-derived STEC may not grow well on CA-STEC plates. Four (4) of

the 55 non-mauve isolates were initially found to be *stx* positive. However, upon a second round of colony isolation and selection on CA-STEC and *stx* qPCR analysis, these non-mauve isolates were found to be *stx* negative. Accordingly, 249 of the original 850 mauve colonies tested positive for *stx* via qPCR analysis, with 176 isolates *stx1* positive (176/249 [70.7%]), 57 isolates *stx2* positive (59/249 [23.7]), and 14 isolates *stx1* and *stx2* positive (14/249 [5.6%]) (Figure 11).

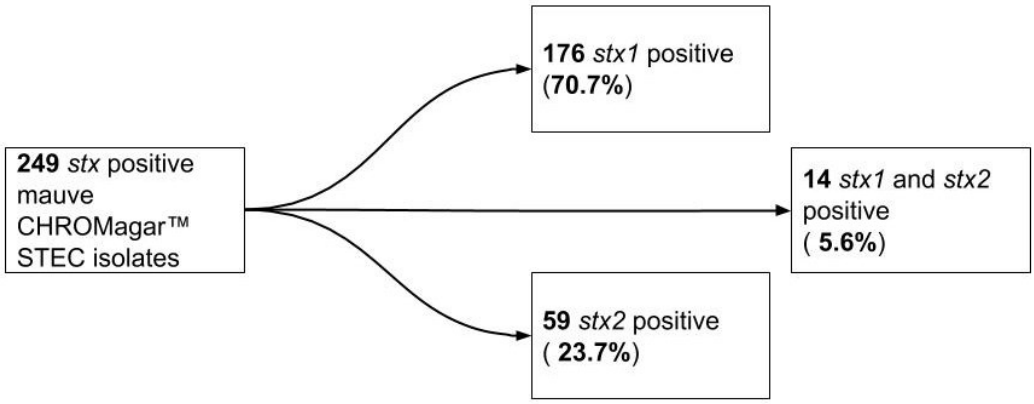

Figure 11 – *stx* positive mauve coloured CHROMagar™ isolates separated by their respective *stx* positivity as determined by *stx1*/*stx2* qPCR.

These 249 *stx* positive mauve isolates represented 57 STEC positive drinking water samples, with 36 of these samples positive for *stx1* (36/57 [64.9%]), 16 for *stx2* (16/57 [29.8%]), and 5 for *stx1* and *stx2* (5/57 [8.8%]) (Figure 12). If multiple *E. coli* isolates from one sample were *stx* positive via qPCR, *stx* positivity from each respective isolate was aggregated and attributed to its corresponding sample.

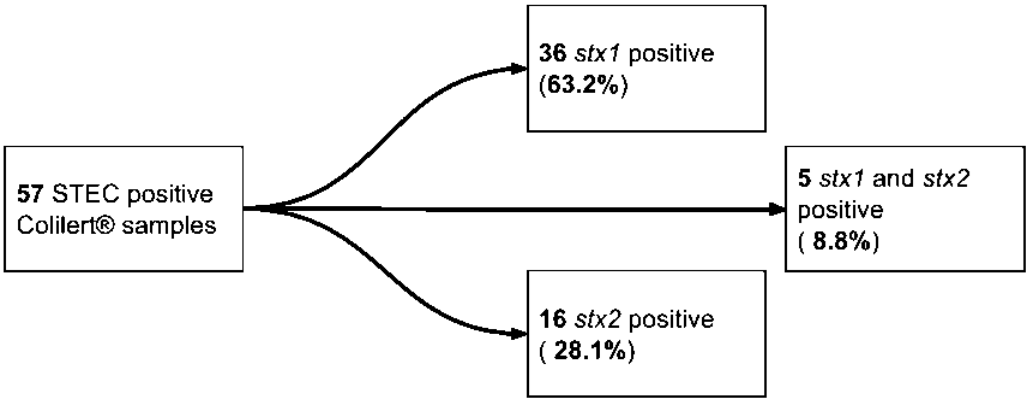

Figure 12 - STEC positive Colilert® Samples separated by their respective *stx* positivity as determined by *stx1*/*stx2* qPCR of all mauve colonies recovered from each sample.

One hundred and forty-three (143) of the 152 *stx* positive drinking water samples (94.1%) yielded mauve colonies: 28 samples yielded only *stx* positive mauve colonies, 29 samples yielded both *stx* positive mauve colonies and *stx* negative mauve colonies, and 86 samples yielded only *stx* negative mauve colonies. Nine (9) of the 152 *stx* positive water samples (5.9%) did not yield any mauve colonies.

Since CA-STEC was used as an isolation screen rather than a confirmatory test for STEC determination, the ability to construct a contingency table and calculate simple statistical performance measures such as sensitivity or specificity for the isolation protocol was limited. As only *stx* positive samples were plated to CA-STEC, no true negatives or false positives could be included in performance calculations of STEC recovery on CA-STEC. With regards to performance calculations of STEC recovery from mauve isolates, since non-mauve isolates were not consistently tested by qPCR, no true negatives or false negatives could be

accurately included in calculations. Due to this lack of data, sensitivity, specificity, positive predictive value, negative predictive value, and other statistical performance measures were not used in order to avoid misinterpretation of comparisons between this and other studies.

Upon visual comparison of the *stx* gene distribution of isolated STEC from *stx* positive Colilert® enriched drinking water samples (Figure 12) and the *stx* gene distribution of the *stx* positive Colilert® enriched drinking water samples themselves (Figure 3), there was a notable difference. The proportion of STEC isolated using CA-STEC that were *stx2* positive was much lower than in *stx* positive Colilert® enriched drinking water samples. This divergence in *stx* distributions was found to be significant when tested with the Pearson's chi-squared test (Chi-squared: 15.45, P-value: <0.0001) and suggests that the CA-STEC isolation protocol disproportionately failed to recover the potentially more pathogenic *stx2*-harboring STEC from environmental water samples.

However, for *stx* positive Colilert® drinking water samples that were both *stx1* and *stx2* positive, this reduced recovery of *stx2* positive STEC may have been a product of the proportion of total bacteria in the sample that were *stx2* positive. In many of the cases where both *stx1* and *stx2* were present in the enriched Colilert® sample, the copy numbers per qPCR reaction of *stx2* genes were noticeably lower than the copy numbers per qPCR reaction of *stx1* genes. For example, in sample #1709 there were 35,000 *stx1* gene copies/5 µl compared to approximately 200 *stx2* gene copies/5 µl, an approximate 175:1 ratio of *stx1: stx2*. A discrepancy in the proportion of *stx* genes in a drinking water sample was not unexpected since

samples that were *stx1* and *stx2* positive could be representative of an STEC carrying both genes, multiple STEC strains carrying each *stx* gene, or a combination of these two possibilities. It is also important to note that due to the ability of STEC to incorporate multiple *stx* genes into their genomes, a single STEC could potentially harbour multiple *stx1* and *stx2* genes, and therefore *stx* copy numbers do not necessarily represent the bacteria that harbour them in a 1:1 ratio [5]. Consequently, the *stx1*/*stx2* ratio of sample #1709 is not an exact representation of STEC organisms in the drinking water sample. Nevertheless, the magnitude of the difference still does suggest that *stx1*-harbouring STEC strains were much more abundant than *stx2*-harbouring STEC strains when the enrichment broth was plated to CA-STEC. Therefore, for samples where copy numbers of *stx2* were noticeably lower than *stx1*, it is reasonable to presume that *stx2* positive STEC strains would be more difficult to isolate simply due to the fact that there were less of these bacteria plated to the media, and as a result, they would be less likely to be randomly picked off the CA-STEC plates. The reverse situation was less common yet had similar results; thus when *stx2* had noticeably higher copy numbers per qPCR reaction than *stx1*, *stx2* and not *stx1* positive STEC were usually recovered from the sample.

4.2.2 Assessing Genetic Diversity of STEC in Drinking Water Samples by (GTG)5 rep-PCR

The results obtained from the CA-STEC isolation revealed that in some drinking water samples multiple STEC isolates were present, raising the possibility that more than one STEC strain might also be present in the same water sample. Consequently, (GTG)5 rep-PCR analysis was performed on the STEC isolates to: i) characterize genetic diversity of isolated strains within each of the water samples; and ii) identify the repertoire of unique STEC strains observed in private well water samples so as to minimize clonal representation of isolates sent to the Canadian National Microbiology Laboratory for serotyping.

Of the 249 *stx* positive mauve isolates collected, 231 isolates underwent (GTG)5 rep-PCR and were subsequently 'fingerprinted' by high-resolution capillary electrophoresis DNA-fragment analysis using the QIAxcel® Advanced system. Six *stx* positive water samples, from which 18 STEC colonies were isolated, were not included in the (GTG)5 PCR analysis due to a laboratory error. Upon comparison of (GTG)5 fingerprints between all isolates from each respective *stx* positive water sample, 65 isolates were determined to be unique, representing 51 *stx* positive water samples. Some water sources contained only a single identifiable STEC clone even though multiple colonies were isolated from the sample (Solid selection – Figure 13), whereas other water samples were shown to be contaminated with multiple genetically unique strains of STEC (Dashed selection – Figure 13). For those water samples which contained multiple unique isolates, a comparison of the (GTG)5 fingerprints of each unique isolate can be found in Appendix B.

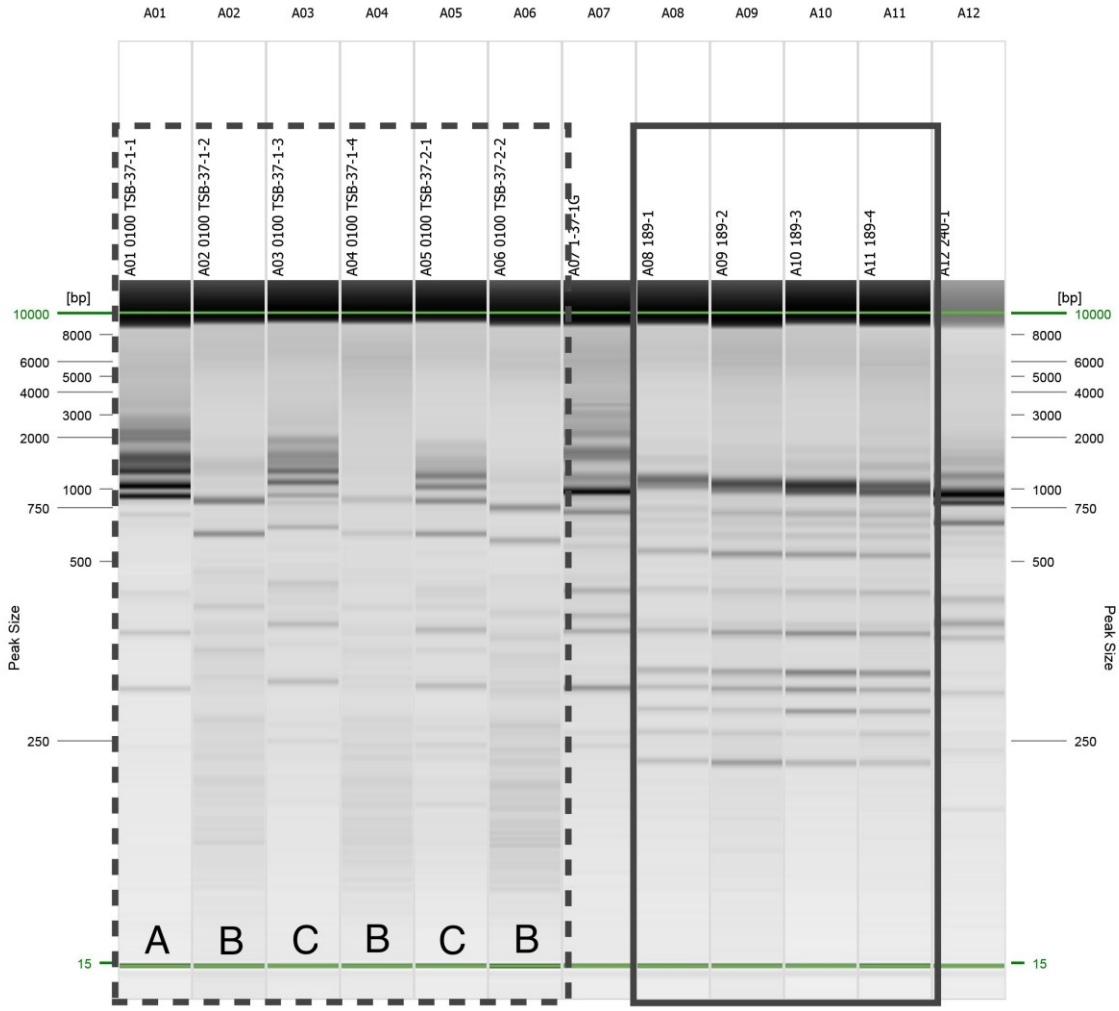

Figure 13 – An example of a virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from 12 mauve CHROMagar™ STEC colonies isolated from 4 *stx* positive Colilert® enriched drinking water samples. The dashed selection highlights the DNA-fragment fingerprints of six *stx* positive mauve colonies isolated from Sample #9100. Three unique (GTG)5 fingerprints (strains) are represented within these six isolates and are labeled (A,B,C) respectively. The solid selection highlights the DNA-fragment fingerprints of four *stx* positive mauve colonies isolated from Sample #189. All four isolates share the same (GTG)5 fingerprint and represent a single STEC strain.

Each of the 65 unique genetic isolates were then tested via a Vitek® Automated Bacterial Identification System (bioMerieux, Marcy-l'Étoile, France) to confirm that the bacterial isolates were truly *E. coli*. Interestingly, one isolate (#981-6) had three distinct colony morphologies when plated onto blood agar during Vitek® analysis, and subsequently one isolate representing each of these respective morphologies was included in the serotype analysis (total of 67 isolates). Sixty-five (65) of the 67 isolates were biochemically confirmed as *E. coli* via Vitek®, and of those isolates not confirmed as *E. coli*, one isolate was identified as *Plesiomonas shigelloides*, and one isolate was identified as *Klebsiella pneumoniae* ssp *pneumonia*. All 67 isolates ran by Vitek® were also analyzed via *stx* qPCR. The presence of *stx* genes were confirmed for all isolates except for the two isolates explicitly identified as non-*E. coli* bacteria. The 65 *E. coli* isolates confirmed to contain *stx* genes were then sent for serotyping.

4.2.3 Serotyping

All 65 STEC isolates were sent to the National Microbiology Laboratory in Guelph, Ontario, Canada, for serotype analysis according to their respective O and H surface antigens. Twenty-one (21) STEC serotypes were identified, the frequencies of which are outlined in Table 20. Five drinking water samples were shown to be contaminated with multiple serotypes (Table 21). Four of the 'big-six' non-O157 serotypes [O26, O103, O121, and O145,] (Table 20) were the most frequently detected serotypes in water, detected in 21 of the 51 STEC positive water samples

(41%), with one sample containing both O26 and O145 serotypes (Sample #1677, Table 21).

Table 20 – The frequency of STEC serotypes recovered from drinking water samples.

Serotype	Number of Samples	Number of Strains
O145:NM[a]	6	7
O168:H8	6	7
O26:H11[a]	6	7
O121:H19[a]	5	5
O5:NM	4	4
O84:NM	4	4
O109:NM	3	3
O156:H25	3	4
O177:NM	3	4
O103:H25 [a]	2	2
O126:H8	2	4
O157:H7[b]	2	2
O26:NM [a]	2	2
O10:NM	1	1
O136:H12	1	1
O146:H21	1	1
O182 (O109):H25[c]	1	1
O182:H25	1	1
O183:H18	1	1
O46:H38	1	1
O8:H19	1	1

[a] Serotypes belonging to the non-O157 'big-six' clinical serotypes.
[b] Serotype commonly associated with clinically-relevant EHEC.
[c] A reaction with the O182 antisera, and a weak reaction with the O109 antisera.
The most clinically relevant serotypes in North America are bolded.

Table 21 – Drinking water samples that contained multiple STEC serotypes and the corresponding STEC serotypes recovered

Sample ID	Serotypes
241	O84:NM ; O177:NM
284	O145:NM ; O109:NM
332	O26:H11 ; O5:NM
1072	O8:H19 ; O126:H8
1677	O26:NM ; O145:NM

Based on a concurrent MSc. project that investigated antimicrobial resistance (AMR) in private well waters, it was determined that 9 of the 65 serotyped isolates were from drinking water samples that had also contained antimicrobial resistant *E. coli* [210]. To determine if any of these 9 drinking water samples contained antibiotic resistant STEC, the STEC isolates from these specific samples were screened for resistance against a panel of 16 different antibiotics based on the National Antimicrobial Resistance Monitoring System (NARMS) as described by Meyer [210]. The AMR profiles of these 9 isolates are represented in Table 22. Interestingly, one of the water samples (#1785) had two STEC strains of the same serotype (O26:H11), but with different resistance patterns. One strain was susceptible to all antibiotics in the panel, whereas the other was resistant to tetracycline and sulfisoxazole. In addition to the differences in (GTG)5 rep-PCR fingerprints, the differences in the antibiograms of these two isolates further confirmed that there

were multiple strains of O26:H11 STEC that had contaminated this drinking water sample.

Table 22 - Antimicrobial resistance profiles of serotyped *E. coli* isolates recovered from AMR-*E. coli* positive drinking water samples

Sample ID	Isolate ID	Serotype	AMR Profile
981	981-6	O121:H19	Susceptible to all
1397	1397-B-4	O156:H25	Susceptible to all
1752	1752-3	O156:H25	Susceptible to all
	1752-4	O156:H25	Susceptible to all
1778	1778-6M-1	O177:NM	Susceptible to all
1785	1785-2	O26:H11	Susceptible to all
	1785-3L	O26:H11	Tetracycline, Sulfisoxazole
1867	1867-2	O168:H8	Tetracycline, Sulfisoxazole
1918	1918-3	O168:H8	Susceptible to all

4.3 Discussion

The objectives of this portion of the thesis were to: i) isolate viable STEC from drinking water samples that screened positive for *stx* genes; ii) evaluate the genetic relatedness between STEC strains recovered from drinking water samples; and iii) serotype genetically unique STEC isolates recovered from drinking water samples.

This study recovered 63 genetically unique STEC strains from well-sourced non-municipal drinking water systems in southern Alberta and to the best of our knowledge, is the first study to recover viable STEC from well-sourced drinking water samples in Canada and one of the only studies to do so globally. The relatively low culturable STEC recovery rate (37.5%) from drinking water samples that screened positive for *stx1* and/or *stx2* genes both highlights, and provides insight

into, the difficulties of studying environmental microbial pathogens. These pathogens are often difficult to detect and recover due to their low numbers in the environment when compared to the surrounding non-pathogenic flora, as well as the phenotypic similarity they often share with the non-pathogenic background bacteria [14,169]. In addition, and unlike clinical samples that strictly represent STEC diversity within human carriers, STEC found in groundwater reflects the potential genetic diversity of STEC strains infecting humans, agricultural animals, as well as wildlife.

Due to the phenotypic similarity of STEC to other environmentally occurring bacteria, and particularly non-pathogenic *E. coli*, a selective agar was used to aid in the isolation of these pathogens from drinking water samples. CA-STEC was chosen because this agar has consistently been suggested as the best option for the isolation of a wide variety of STEC serotypes, particularly for samples that contain significant amounts of background flora, and is ideally suited for protocols that require the recovery of STEC from samples that have already tested positive for the presence of these pathogens [170–175].

An initial protocol for STEC isolation struggled to consistently isolate mauve colonies due to overgrowth of background flora in the *stx* positive drinking water samples. Consequently, a revised isolation protocol was implemented in order to better prevent non-STEC flora from interfering with the growth of presumptive STEC colonies (i.e., mauve colonies). Tzschoppe *et al*. [219] had similar issues with the overgrowth of background environmental flora inhibiting the ability to isolate single mauve colonies on CA-STEC, for which they developed an optimized protocol for

pathogen growth. This optimized protocol involved enriching samples using a growth media selective for gram-negative bacteria, reducing the incubation time of the enrichment, and increasing the incubation temperature for the CA-STEC plates from 37° C to 44° C. Tzschoppe *et al*. [219] found that using the selective growth media: i) resulted in a 50% reduction in natural background flora without decreasing the number of detectable EHEC on the CA-STEC plates; ii) that incubation periods greater than 6 hours were disadvantageous for samples with low levels of pathogen; and iii) that the growth of background flora was strongly inhibited on CA-STEC plates at 44° C. Gill *et al*. [220], although using different selective agar, again highlighted the importance and difficulties of limiting environmental background flora when attempting to isolate STEC and a recent study by Morris *et al*. [169] used CA-STEC to isolate STEC from water samples and similarly experienced issues with identifying and selecting mauve colonies due to high microbial background growth. These two studies also found that increasing incubation temperatures to 42° C helped reduce the amount of background growth. Taking this previous research into consideration, in the revised protocol MAC broth was used during a 4-hour enrichment to inhibit the growth of gram-positive bacteria present in the Colilert® positive drinking water samples and to help prevent the overgrowth of natural background flora prior to plating the samples to CA-STEC. In addition to reducing background growth, enrichment using MAC broth has also been reported to reduce non-specific fluorescence in background flora plated to CA-STEC [170]. To further inhibit the growth of non-STEC background flora in the revised protocol, the overnight CA-STEC incubations were carried out at 44° C. Since coliform bacteria

such as *E. coli* are able to ferment lactose up to temperatures of 44.5 °C, this increased incubation temperature specifically inhibits background growth of environmental bacteria unable to properly function at these higher temperatures [222]. Tzschoppe et al. [219] did however report one strain that grew only at 37° C and not 44° C. To ensure that STEC unable to grow at increased temperatures were not missed, samples with few or no mauve colonies incubated at 44° C were subsequently re-plated and incubated at 37° C.

Results from Gouali *et al*. [171], which investigated optimal enrichment time for samples being plated to CA-STEC agar, conflicted with those from Tzschoppe *et al*. [219], and suggested that longer enrichment times (7 hours as opposed to 3-4 hours) aided in the recovery of STEC from frozen stool samples. However, a notable difference between these methodological examinations was the type of sample plated to CA-STEC, with Gouali *et al*. [171] plating stool samples and Tzschoppe *et al*. [219] plating pathogen-spiked store bought vegetable samples. Arguably, the environmental flora of drinking water would be more similar to the environmental flora experienced by STEC on vegetables, rather than STEC from the human gut. Because of this, a reduction in incubation time was implemented for the enrichment phase in the present study, as suggested by Tzschoppe *et al*. [219]. Supporting this previous assumption about environmental conditions was the observation that studies investigating fecal samples generally did not report issues of background growth, instead reporting that CA-STEC was often successful at reducing the effects of non-STEC during the isolation process, whereas studies investigating non-fecal samples often did report such issues [169–171,173,219,220,223,224]. This observation, along

with the results of the current study, suggest that background flora from environmental sources are better able to grow on CA-STEC than background flora from clinical/direct fecal sources, and this growth can inhibit the proper detection of STEC on this media.

To the best of our knowledge, there are no other studies that used CA-STEC to recover STEC from drinking water samples for direct comparison to the present study and only one other study has used CA-STEC to examine STEC in water [169]. Morris *et al.* [169] recovered STEC from 50% (1/2) of *stx* positive river water samples using a method that paired the filtration of large quantities (20-33L) of river water, PCR, and CA-STEC. The majority of CA-STEC studies have used STEC from archived strain collections, stool samples, or food samples to investigate the effectiveness of this growth media [171,173,174,219,223,224].

Studies that plated STEC strains from archived clinical collections had STEC recovery rates on CA-STEC ranging from 51%-100% and a collective rate of 77.0% [170,172–175,219,224,225]. Verhaegen *et al.* [175] also plated 8 non-clinical STEC strains isolated from cattle carcass swabs (n=6) and cow milk (n=2) with each strain growing as typical mauve colonies on CA-STEC. Studies that used CA-STEC to investigate STEC in human stool samples had STEC recovery rates ranging from 50%-86% and a collective rate of 80.5% [170,171,173,223]. Kase *et al.* [174] spiked foodstuffs (baby spinach, cilantro, alfalfa sprouts, raw cow milk) with a variety of STEC serotypes and 38/105 (36.2%) spiked portions grew *stx* positive mauve colonies on CA-STEC, a recovery rate similar to that of the present study (37.5%).

This similarity is notable since, as mentioned previously, foodstuffs are arguably a more accurate representation of environmental conditions for drinking water samples than stools. These results again suggest that interaction between background flora and STEC during isolation protocols, either during enrichment or on CA-STEC agar, may notably affect the ability of this agar to recover STEC. STEC strains from environmental samples may behave differently than clinically-derived strains on CA-STEC agar, which might explain some of the discrepancy between recovery rates between the present study and those that used clinical STEC strains. The developmental focus of CA-STEC has primarily been on the most clinically relevant STEC (the 'big six' non-O157 and O157), which are not necessarily the most relevant environmental STEC and therefore a difference in recovery was not unexpected [176]. A number of these investigations into the efficacy of CA-STEC for the isolation of a broad range of STEC serotypes have shown that STEC which have different phenotypic characteristics than the most clinically relevant serotypes are more difficult to isolate [170–173,223]. Genetic characteristics such as tellurite resistance (i.e., presence of *ter* gene cluster), the ability to ferment sorbitol, and the presence of the intimin virulence factor (*eae*), rather than serotypes specifically, have been commonly associated with a lack of, or reduced, growth on CA-STEC agar [173,174,219,224–226]. The genetic profiles of STEC commonly found in well-sourced drinking water samples in southern Alberta are unknown and likely highly variable. As a result, any significant differences in the phenotypes of these STEC and the phenotypes of the most clinical relevant STEC may have a substantial impact on the ability of CA-STEC to isolate these environmental pathogens. Another possibility

that may explain why STEC was not recovered from all *stx* positive drinking water samples is that some STEC may not have presented as mauve colonies on CA-STEC and therefore were not picked for *stx* qPCR analysis. Three studies have reported STEC that formed non-mauve colonies on CA-STEC, with white, grey, and blue-grey colonies being subsequently confirmed as STEC [170,172,225]. In addition, the serotypes of the two STEC strains of Canadian origin which produced non-mauve colonies (O121:H19 and O145:NM) match serotypes recovered from water samples of the current study, showing geographical precedence for this behaviour. Finally, despite a drinking water sample testing positive for *stx* via qPCR, this does not necessarily mean that there were viable STEC in the sample, an issue that was explored in the previous chapter. If the *stx* amplified in a water sample by qPCR were from non-viable STEC, STEC in a VBNC state, or due to the presence of *stx* encoding bacteriophages, STEC growth and the resulting *stx* positive mauve colonies would not be expected on CA-STEC.

Although the manufacturer of CA-STEC suggests mauve coloured colonies can be interpreted as STEC, this interpretation would have been incorrect for the majority of mauve colonies isolated from *stx* positive Colilert® enriched drinking water samples. The growth of *stx* negative *E. coli* strains as typical mauve colonies has also been reported by other studies, with a range from 0%-36% of non-STEC strains presenting as presumptive STEC colonies [170,171,173–175,224,225]. Interestingly, strains of other *E. coli* pathotypes (ETEC, EIEC, EPEC, EAEC, and EHEC-like) were commonly found to grow as *stx* negative mauve colonies on CA-STEC [170,171,173–175,219,224,225]. From the perspective of this current project, the results of Cernela *et al.*

[225], which found 83.3% (10/12) of *eae* positive, *stx* negative (i.e., EHEC-like) *E. coli* from cattle fecal sources grew as typical mauve colonies on CA-STEC, are particularly notable. A recent study of STEC in cattle feces collected in western Canada found that 21% (215/1035) of the top seven clinically important STEC serotypes recovered did not harbor *stx* genes, and accordingly these *stx* negative mauve colonies may be *E. coli* of common STEC serotypes without *stx* genes [116]. Importantly, only 5% of recovered STEC serotypes from this study investigating cattle feces lacked any of the four important STEC virulence factors tested (*stx1*, *stx2*, *eae*, and *ehxA*). Strains of most other bacterial species tested for growth on CA-STEC were not found to grow as mauve colonies, although Kase *et al.* [174] did report a *Staphylococcus aureus* that grew as a mauve colony [170,175]. Both the *Plesiomonas shigelloides* and the *Klebsiella pneumoniae ssp pneumoniae* isolated from CA-STEC during the current study formed mauve colonies on CA-STEC however. Relevant studies have found that 4%-16% (12.4% collectively) of *stx* negative stool samples also resulted in the growth of *stx* negative mauve colonies on this selective agar [170,171,173,223]. In the present study only *stx* positive samples were screened for STEC isolates using CA-STEC, and therefore direct comparisons of the number of *stx* negative samples forming mauve colonies cannot be made with the relevant studies mentioned above. Kase *et al.* [174] was the only study to specifically report the percentage of mauve colonies that were *stx* negative (27/319, 8.5%), which was appreciably lower than the percentage of mauve colonies that were *stx* negative in the current study (601/850, 70.7%). Morris *et al.* [169] reported that a "large number"

of mauve colonies were negative for *stx* genes when investigating river water, but did not provide specific counts.

The results of the current study as well as Morris *et al.* [169] suggest that recovering STEC from environmental waters using CA-STEC may be particularly difficult due to increased levels of *stx* negative mauve colonies, and reiterates the suggestion that bacteria isolated from environmental samples likely behave differently than bacteria isolated from human feces on CA-STEC. The frequency of other pathotypes of *E. coli* to grow as typical mauve colonies begs the question of whether these drinking water samples contain multiple *E. coli* pathogens. If so, this could have important public health implications due to the potential development of co-infections or the potential development of new strains of STEC through lateral *stx* gene transfer, both of which could increase the severity of illness associated with exposure to this contaminated water source. These *stx* negative mauve colonies may also be non-pathogenic *E. coli* or other bacterial species, such as the aforementioned *Plesiomonas shigelloides* and *Klebsiella pneumoniae ssp pneumonia*, and further study into the identification of these abnormal colonies is warranted. An alternate possibility that may explain the increased amounts of *stx* negative mauve colonies found in this study, is that the original *stx* genes detected via qPCR were excised from the STEC genome during enrichment or while growing on the CA-STEC agar. The loss of *stx* genes during enrichment, subcultivation, as well as during human infection, has been well documented yet is poorly understood [227–231]. Shiga toxin gene (*stx*) loss is particularly common for *stx* positive bacteria that are not *E. coli* [231–235]. Interestingly, this loss of *stx* genes may explain the negative *stx* qPCR results of

the 4 subcultured non-mauve isolates and the 2 subcultured non-*E. coli* mauve isolates that originally tested positive for *stx* genes.

Twelve of the twenty-one (12/21) serotypes that were recovered from the water samples of this current study (57.1%) have been previously isolated using CA-STEC [170–175,219,224,225]. To the best of our knowledge, STEC serotypes O10:NM, O109:NM, O126:H8, O136:H12, O168:H8, O182 (O109):H25, and O183:H18, have not been examined for their CA-STEC growth efficiency and this is the first time they have been isolated using this media. There have been previous attempts to isolate STEC serotypes O8:H19 and O146:H21 without success, and to the best of our knowledge this is the first time these serotypes have been successfully recovered from CA-STEC [172,173,224]. The recovery of serotype O146:H21 was particularly notable as two previous studies attempted to grow 4 different O146:H21 isolates on CA-STEC at 35-37° C with no growth reported [172,224]. In the present study, when the drinking water sample that was subsequently determined to carry O146:H21 was incubated on CA-STEC at 37° C mauve colonies did grow, although none of the 10 mauve colonies picked were positive for *stx* genes upon qPCR analysis. However, when incubated at 44° C one of the four mauve isolates picked at this temperature was confirmed to be *stx* positive and this isolate was subsequently characterized as O146:H21. This result further suggests that incubation temperature may play an important role in the ability for STEC to be recovered from this agar and that multiple incubation temperatures could be important to incorporate in CA-STEC isolation protocols.

The data collected also suggests that the CA-STEC isolation protocol recovered disproportionately less *stx2*-harboring STEC from environmental water samples, although the reasoning behind the apparent selection against this genetic variant of STEC is unknown. At least a portion of the difference in *stx* gene recovery was likely due to disparities in the amount of each STEC strain present in the aliquots plated to CA-STEC. As noted in the results, strains harbouring *stx2* were often found in lower quantities in samples that likely contained multiple STEC strains, and therefore were less likely to be recovered. It is possible however, that one or more of the genetic factors influencing growth on CA-STEC are correlated with *stx2* for waterborne STEC in southern Alberta, but without a better understanding of both the genetic profiles of these STEC and the factors influencing STEC growth on CA-STEC, it is difficult to provide any potential hypotheses. Any bias towards non-*stx2* STEC by this protocol may prevent an accurate estimation of the STEC population in non-municipal drinking water in southern Alberta. Of particular concern is that the more pathogenic *stx2* carrying STEC were less likely to be isolated using this protocol, which could lead to an overly conservative risk estimate with regards to STEC in non-municipal drinking water in this region.

As outlined previously, there has been a limited amount of research investigating STEC in environmental waters, which makes the comparison of recovered STEC serotypes more challenging. Halabi *et al*. [114] is one of the few studies that has done so, and recovered 10 different EHEC serotypes from rural drinking water in Austria, although only one serotype (O167:H16) carried *stx* genes. One of the serotypes recovered in Austria, O26:H11, was also recovered from the drinking

water samples in the current study. Ramteke and Tewari [113] investigated drinking water sources in India and recovered three different STEC serogroups, two of which (O103 and O157) were also recovered in the current study [113]. It should be noted however that Ramteke and Tewari [113] did not confirm the presence of *stx* genes in these presumptive STEC. Schets *et al*. [108] recovered STEC O157:H7 from 4 private water supplies in the Netherlands and Won *et al*. [117] recovered STEC O157:H7 from 7 private water supplies in Ohio, U.S.A.. Notably, these two studies did not attempt to recover non-O157 STEC and focused solely on O157:H7 [108,117]. While the occurrence of STEC serotypes may not be uniform across all these study areas, these results still demonstrate precedence for these two specific serotypes (O157:H7, O26:H11) and three serogroups (O26, O103, O157) being recovered from drinking water sources.

STEC drinking water outbreaks have also provided information on not only the serotypes present in drinking water sources, but also the potential pathogenicity of waterborne STEC. While the majority of these outbreaks have been attributed to STEC O157:H7, including those in Canada, O26:H11 has also been the cause of drinking water outbreaks and both of these serotypes were recovered from the drinking water samples of the current study [14,55,236–239]. When considering historical occurrences of STEC serotypes, being mindful of potential bias toward the 'big six' and especially O157:H7 serotypes due to the available detection, isolation, and reporting protocols at the time of investigation, is important.

To the best of our knowledge this is the first study to report non-O157 STEC serotypes in drinking water sources in Canada, and as a result, the most similar

Canadian serotype comparisons available are with surface water studies. Both Nadya *et al.* [118] and Johnson *et al*. [51] reported non-O157 STEC serotypes in surface water from Ontario and BC respectively. Twelve (12) of the 21 serotypes recovered from the current study (57.1%) had been previously isolated from Canadian surface water sources: O5:NM, O8:H19, O26:NM, O26:H11, O103:H25, O109:NM, O121:H19, O145:NM, O157:H7, O177:NM, and O168:H8. The STEC O157:H7 serotype was recovered most frequently from surface water sources, unlike the well water sources of the current study. The second most frequent serotype found in surface waters was O26:H11, which was the most frequently recovered serotype from the current study, alongside both O145:NM and O168:H8. This difference in the recovery of O157:H7 and O26:H11 STEC in drinking water sources, despite the relative abundance of both serotypes in surface water, suggests that O26:H11 STEC may be better able to survive transport from surface water into groundwater than O157:H7 STEC. It is reasonable to predict that the STEC that are best able to survive in low nutrient environments, such as water, will be those that are most likely found in drinking water samples.

The majority of the serotypes recovered in this study (17/21; 81.0%) have been previously determined to be clinically relevant, and all 19 serogroups (100%) which were recovered from the drinking water samples have also been previously associated with human illness [32,37,170,172,173,175,224,240–243]. The four serotypes (O10:NM, O109:NM, O136:H12, O168:H8) that have not been specifically linked with human illness, have been previously recovered from bovine fecal samples [244–247]. Of the 51 drinking water samples from which viable STEC was recovered, 40 (78.4%)

contained clinically relevant STEC serotypes and 22/51 samples (43.1%) contained one or more of the most clinically relevant serotypes (the 'big six' or O157 serotypes). These results exemplify the potential of these STEC-contaminated drinking water samples to cause serious human illness and highlight the need to consider drinking water as a potential source of infection when sporadic and outbreak cases of STEC-related illness are being investigated. The remaining 29/51 drinking water samples (56.9%) contained exclusively 'rare' STEC serotypes. The common occurrence of these 'rare' STEC serotypes is important to note due to current diagnostic practices and reporting protocols being primarily focused on only the most clinically relevant serotypes. Even the isolation protocol used for this current study was likely biased towards the recovery of these more well-known STEC serotypes through the use of CA-STEC, which was developed for and is most effective when, isolating O157 and the 'big six' STEC. The potential of missing or underrepresenting these 'rare' STEC during surveillance and epidemiological investigation due to methodological limitations should be concerning, especially considering the frequency of which they were recovered from these water samples. These results provide further evidence that without procedures in place to truly investigate all STEC serotypes our understanding of the epidemiology of this pathogroup will be incomplete, and the risk associated with STEC-disease will likely be underestimated.

Clinical relevance of STEC serotypes is not uniform across the globe however, and therefore it is useful to compare the results of this study to serotypes that have been recovered in Canada and Alberta [221]. Twenty-two out of fifty-one (22/51,

43.1%) drinking water samples from which viable STEC was recovered contained the Canadian top six serogroups or O157 STEC, and 14/19 (73.7%) of serogroups recovered from the drinking water samples in this study have been previously associated with STEC related illness in Canada [221]. When comparing the serotypes recovered from drinking water samples to those of clinical importance within Alberta, seven recovered serotypes (O5:NM, O26:NM, O26:H11, O103:H25, O121:H19, O145:NM and O157:H7) are regularly found within clinical cases in the province, and represent three out of the four most common serogroups (O26, O121, and O157) in clinical cases [35,40,43–45,241,248]. In addition, four out of the top five most abundant serotypes recovered from drinking water samples (O145:NM, O26:H11, O121:H19, O5:NM) are regularly found within clinical cases in Alberta. As the third most common serogroup recovered from STEC patients in Alberta, O111 is unexpectedly absent from the results of this study. As one of the 'big six' it is unlikely that the isolation protocol would have prevented this serogroup from being recovered. However this serogroup has not been commonly recovered from cow feces in western Canada, and therefore there may be lower inputs of these specific STEC into the environment in southern Alberta [116,146,249]. It is also possible that this serogroup is ill adapted to survive in the conditions required for these STEC to be transported to, and survive in, the groundwater sources that these drinking water samples were drawn from. Interestingly serogroup O145 has also rarely been recovered from cattle feces in western Canada and yet it was the second most commonly isolated STEC serogroup from non-municipal drinking water samples (along with O168) [116,249]. This serogroup may be well adapted for environmental

survival or cattle may not be the most important reservoir/source of this serogroup in Alberta. Nonetheless, these results show that the most common and clinically important STEC serotypes in the study region are frequently found within non-municipal drinking water supplies and groundwater consumption may be an overlooked potential source of STEC infection.

The samples that would have the greatest risk of causing STEC related illness are those in which contain particularly virulent STEC, such as O157:H7, or those samples which contain multiple STEC strains and therefore the possibility of co-infection. There were 5 samples in which multiple STEC serotypes were recovered and 7 samples containing multiple strains of the same serotype, suggesting the possibility of STEC co-infection from these drinking water sources. Of the five drinking water samples with multiple serotypes, four also contained at least one of the most common clinical serotypes in Alberta, and one of these four samples contained two of the most clinically relevant serotypes (O26:NM and O145:NM). There were 7 serotypes that had multiple strains within a single drinking water sample, two of which belonged to the most common clinical serotypes in Alberta (O26:H11 and O145:NM). From a public health perspective, these six samples would be of especially high concern due to not only the presence of highly virulent STEC serotypes but also the risk of co-infection.

Another group of samples that were determined to be of particular concern from a public health perspective are those which contained STEC that were anti-microbial resistant (AMR). Although anti-microbial resistance was not the focus of this thesis, the opportunity to further characterize a limited number of isolates was

taken. Two of the nine samples from which STEC isolates were recovered from and were separately determined to contain AMR bacteria, contained AMR STEC. Both AMR positive isolates were also multi-drug resistant, which again, is concerning from a public health perspective [250]. One of the two AMR serotypes represented in these samples (O26:H11) is the second most frequent serotype associated with clinical cases in Canada, has been the causative pathogen of waterborne STEC outbreaks and is clearly a risk to human health [221,251]. This AMR O26:H11 STEC was one of two O26:H11 STEC serotype strains recovered from a single drinking water sample, a result that highlights the potential genetic diversity of environmental STEC even when recovered from the same water sample and belonging to the same serotype.

CHAPTER 5: Seasonality and Spatiotemporal Clustering of Shiga Toxin-producing *Escherichia coli* (STEC) in Non-municipal Well-sourced Drinking Water Samples from Southern Alberta

5.1 Introduction

Understanding temporal and spatial patterns of disease incidence is an important part of determining the etiology of a disease [131]. Studying the relationship between exposure variables and disease distribution in space, time, and space-time, can help determine key risk factors of the disease in question [130,131]. For waterborne disease, and in particular the transmission of disease through non-municipal drinking water sources, identifying patterns of spatiotemporal variability in water contamination can help develop informed management practices and public health programs to help prevent future illness [106,252].

GIS technologies are a valuable tool for investigating spatial patterns of disease, and are commonly used to help identify non-random spatial distributions, or clusters, of disease [253]. Scan statistics are statistical tests that use a flexible scanning window to gradually move through time, space, or space-time settings, and can identify significant clusters of cases in a data set [188]. The ability of these tools to aid in assessing significant clustering effects in space-time is particularly important for the identification of disease outbreaks.

Temporal patterns (i.e., seasonal occurrence) in the incidence of waterborne disease outbreaks, drinking water contamination, and STEC related illness have been previously identified and there are a number of statistical tests to investigate seasonality of disease occurrence [58,74,128,254]. For example, the Edwards' test for seasonality tests the null hypothesis against a harmonic curve with a twelve month period and upon modification, is able to test small and medium sample sizes as well as adjust for months that have notable differences in the number of events being tested [185,186,254].

Many of the risk factors associated with waterborne illness are local to the drinking water source with respect to both space and time, such as the proximity to human septic systems and animal agriculture, or the incidence of an extreme precipitation event [50,60,77,109]. Therefore, determining where and when these pathogens are found in drinking water can aid in the understanding of areas and time-periods at increased risk of contamination. The purpose of this portion of the study was to statistically investigate temporal, spatial and spatiotemporal patterns of *stx* gene and STEC serotype occurrence in non-municipal drinking water samples from southern Alberta and to provide visual representations of the spatial distribution and clustering of these STEC positive drinking water samples. This was done in order to identify any potentially high-risk areas or time-periods of STEC contamination in non-municipal drinking water sources across southern Alberta.

5.2 Results

5.2.1 Edwards' test seasonality analysis

The aggregated time-series of *stx* positivity rates presented in Chapter 3 displayed apparent seasonal patterns of *stx* occurrence in both submitted drinking water samples and *E. coli* positive Colilert® enriched drinking water samples. To determine any statistical significance of these temporal patterns, each monthly time-series presented in Chapter 3 was analyzed by the Edwards' test to detect seasonality within the aggregated data.

Within the monthly time-series of *stx* positivity per 10 000 submitted drinking water samples a statistically significant seasonal pattern was detected by the Edwards' test (Chi-squared = 86.98 [DF = 2]; P-value < 0.001) with the seasonal harmonic curve peaking in early July (Amplitude = 113.4% of peak frequency; Peak angle = 182.8°).

In the early summer of 2005 there was an appreciably higher than average amount of precipitation, and based on the data presented in Chapter 3, a significantly greater occurrence of *stx* in non-municipal drinking water samples was also observed during this year. When 2005 data was treated as an outlier and removed from the seasonality analysis of the monthly time-series of *stx* positivity per 10 000 submitted drinking water samples a statistically significant seasonal pattern was still detected by the Edwards' test (Chi-squared = 19.79 [DF = 2] P-value < 0.001), again with the seasonal harmonic curve peaking in early July (Amplitude = 96.6% of peak frequency; Peak angle = 179.2°). When 2005 data was

removed the significant seasonal harmonic pattern of *stx* occurrence had a reduced peak-amplitude and therefore, the inclusion of the outlying 2005 data within the seasonality analysis for *stx* positivity of submitted drinking water samples effectively increased the strength of the summer peak of *stx* positivity.

Potentially distinct seasonal patterns of *stx* occurrence within submitted drinking water samples and within fecally contaminated drinking water samples were highlighted in Chapter 3 and in lieu of these apparent differences, seasonal analyses of *stx* positivity per 1000 *E. coli* positive Colilert® enriched drinking water samples were also conducted. Within the monthly time-series of *stx* positivity per 1000 *E. coli* positive Colilert® drinking water samples a statistically significant seasonal pattern was detected by the Edwards' test (Chi-squared = 12.10 [DF = 2] P-value = 0.002) with the seasonal harmonic curve peaking in late May (Amplitude = 66.0% of peak frequency; Peak angle = 142.9°).

Again from data presented in Chapter 3, 2005 was determined to be an outlier within the monthly time-series of *stx* positivity per 1000 *E. coli* positive Colilert® drinking water samples. Upon omission of the outlying 2005 data, a statistically significant seasonal pattern of monthly *stx* positivity per 1000 *E. coli* positive Colilert® drinking water samples was not detected by the Edwards' test (Chi-squared = 5.82 [DF = 2] P-value = 0.055). The p-value of this test falls just outside the set significance level of 0.05 however, and caution should be taken when rejecting the presence of a seasonal pattern within this time-series. The seasonal harmonic curve of this Edwards' test peaked in mid-May (Amplitude = 55.6% of peak frequency; Peak angle = 131.9°). The inclusion of the outlying 2005 data within

the seasonality analysis for *stx* positivity of *E. coli* positive Colilert® drinking water samples increased the strength of the seasonal peak to where it became a statistically significant harmonic pattern and also pushed this peak closer towards the summer months by approximately 11 days.

5.2.2 Bernoulli model Kulldorff temporal scan

Building on this temporal investigation of seasonality, *stx* occurrence was then analyzed via a Bernoulli temporal scan statistic in order to determine if there were any time intervals of potentially increased risk of *stx* contamination during the study period. The most likely temporal cluster throughout the entire dataset occurred from May 30 to July 3 of 2005, and as expected, fell within the seasonal peak determined by the Edwards' tests. This time period also corresponded to the June 2005 extreme precipitation event outlined in previous chapters. This temporal cluster included 280 water samples of which 45 were *stx* positive. The relative risk within this cluster was 2.67 with an expected number of cases of 20.96. The log likelihood ratio was 14.22 with a P-value of 0.0003.

5.2.3 Spatial location of *stx* positive water samples and associated STEC serotypes

The geographic locations of the 1607 submitted *E. coli* positive Colilert® drinking water samples that were accompanied by complete ATS information, along with their respective *stx* positivity as determined by *stx1*/*stx2* qPCR analysis, are represented in Figure 14. The geographic location of each of the non-municipal

drinking water samples for which viable STEC was recovered from is provided in Figure 15, along with the corresponding serotype of each STEC.

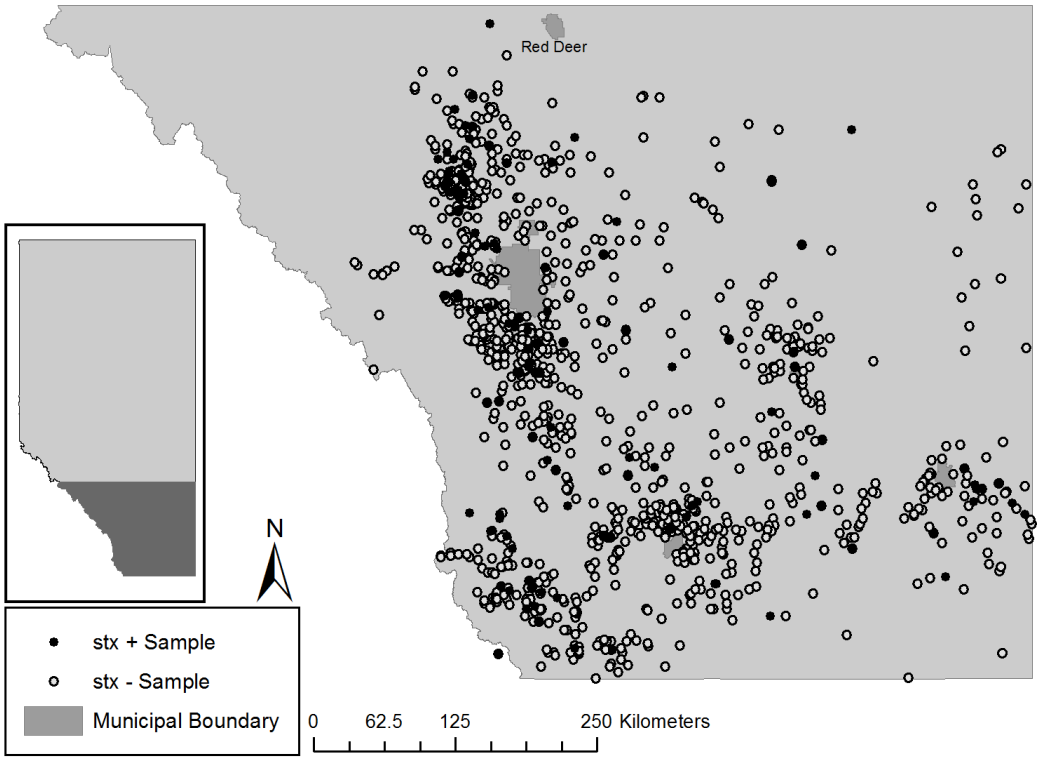

Figure 14 – The geographic location of the corresponding water source for all *E coli* positive drinking water samples having complete ATS information during the study period. Light coloured dots represent *E. coli* positive, *stx* negative, drinking water samples and dark coloured dots represent *E. coli* positive, *stx* positive, drinking water samples.

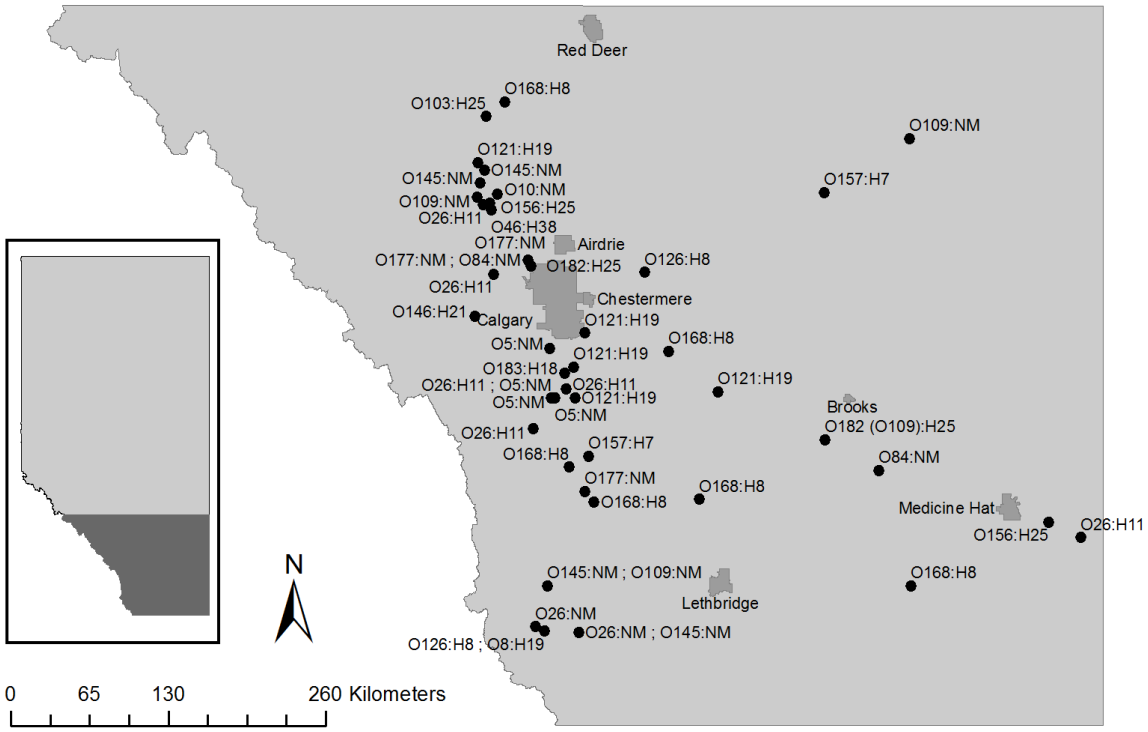

Figure 15 – The geographic location of the water source for each serotyped STEC recovered from drinking water samples and having complete ATS information and the corresponding serotype(s).

5.2.4 Bernoulli model Kulldorff spatial scan

To investigate clustering in space and identify any geographic regions of potentially increased risk of *stx* contamination of groundwater wells in southern Alberta, *stx* occurrence was analyzed via a Bernoulli spatial scan statistic. This test identified a purely spatial cluster in the western portion of southern Alberta (northwest of Calgary, Alberta) when all data was included in the analysis, with no other significant clusters detected (Figure 16).

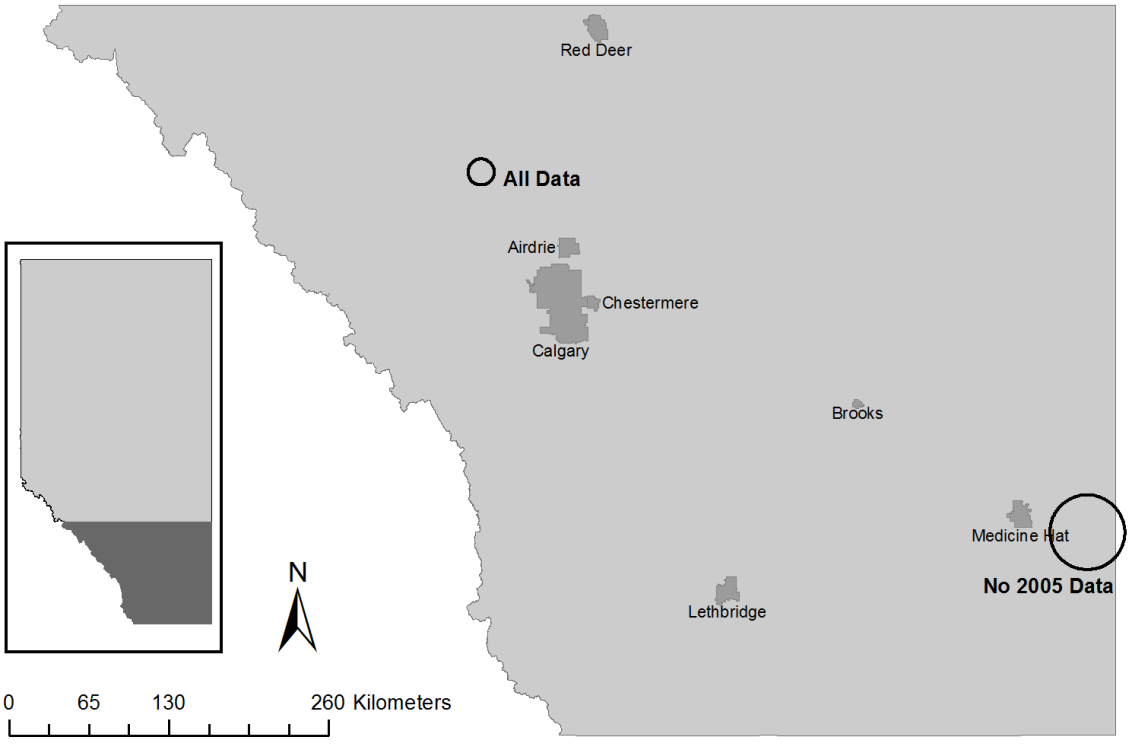

Figure 16 – The geographic location of the most likely spatial cluster of *stx* positive water samples using a Bernoulli model Kulldorff scan for the total study period (circle representing 'All Data') and for a study period with 2005 data removed (circle representing 'No 2005 Data').

The most likely cluster from this test included six total drinking water samples with a radius of 6.58 km. The cluster contained 5 observed *stx* positives compared to an expected 0.48 *stx* positives and had a relative risk of 10.85, a log likelihood ratio of 10.12 and a P-value of 0.047. The STEC serotypes identified in this cluster included O145:NM (3 isolates), and O121:H19.

5.2.5 Space-time permutation model Kulldorff scan

Kulldorff scan statistics are often used as tools to investigate disease outbreaks due to their ability to identify spatiotemporal clustering, which is particularly important when investigating outbreaks due to the potential of important exposures that may only exist over short periods of time or space time [137]. Due to the previously identified clusters in both space and time, a space-time permutation model Kulldorff scan statistic was used instead of a Bernoulli space-time Kulldorff scan statistic since the space-time permutation model automatically adjusts for both purely spatial and purely temporal clusters in the data being scanned [188,193]. The most likely cluster from this test included 160 samples with a radius of 34.53 km during the time frame of 2005/6/20 – 2005/7/3 (Figure 17). The cluster contained 19 observed *stx* positives compared to an expected 4.78 *stx* positives and had a Test statistic of 12.86 and a P-value of 0.0000045. The serotypes observed in this cluster included O10:NM, 026:H11, O46:H38, O109:NM, O103:H25, O121:H19, O145:NM (3 isolates), O156:H25, and O168:H8. The most likely space-time cluster determined by the space-time permutation model encompassed both the purely spatial and purely temporal clusters determined by Bernoulli scan analysis.

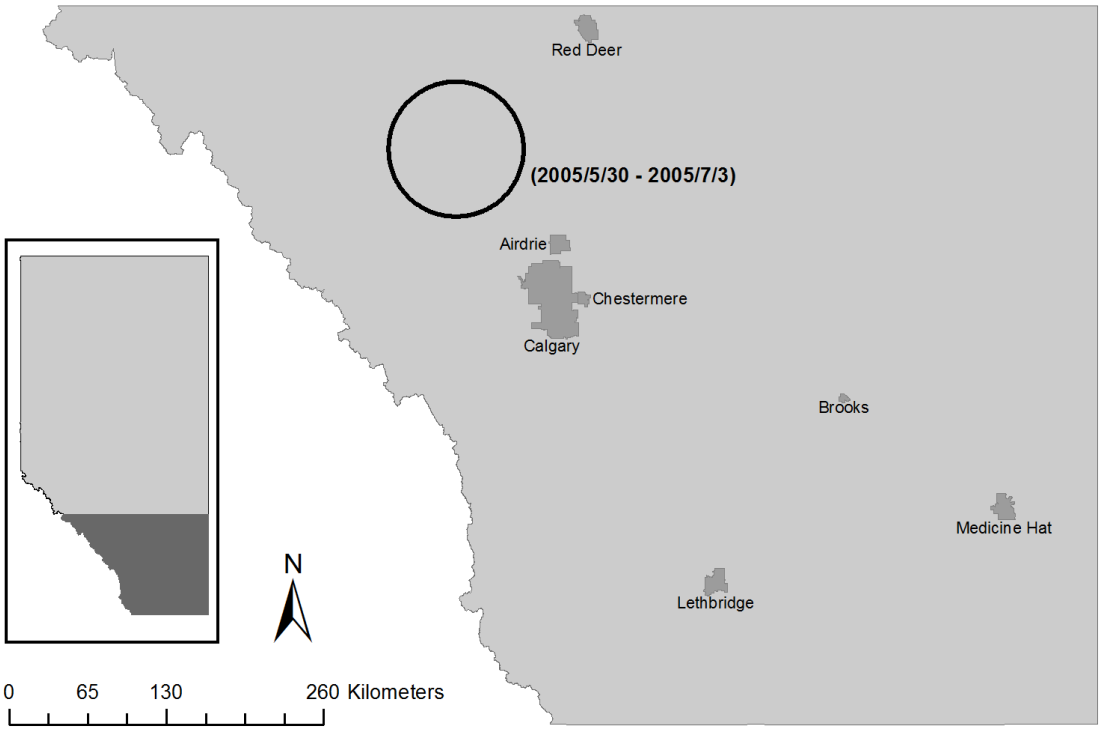

Figure 17 – The geographic location and time frame of the most likely space-time cluster of *stx* positive water samples using a space-time permutation model Kulldorff scan.

5.2.6 Kulldorff scan statistics without 2005 data

The spatiotemporal clustering seen in the results presented suggest that the inclusion of the 2005 data in our analyses may bias results which have assumed independence among cases [137]. Both Bernoulli model temporal and spatial Kulldorff scan statistics, which assume independence among cases, were used again to analyze the adjusted data set with the 2005 data removed. The Bernoulli model Kulldorff temporal scan statistic that was performed on this data set did not identify any statistically significant temporal clusters. The Bernoulli model Kulldorff spatial scan statistic, run without the 2005 data included, did result in a statistically

significant spatial cluster being identified near the eastern edge of southern Alberta (Figure 16). The most likely cluster from this test included sixteen total drinking water samples with a radius of 19.55 km. This cluster of drinking water samples contained 7 observed *stx* positives compared to an expected 1.02 *stx* positives, and had relative risk of 7.59, a log likelihood ratio of 9.22 and a P-value of 0.044. The STEC serotypes identified in this cluster were O26:H11 and O156:H25.

5.3 Discussion

The purpose of this section of the project was to determine if the seasonal patterns of *stx* occurrence in drinking water samples identified in Chapter 3 displayed statistically relevant temporal trends, as well as investigate spatiotemporal clustering of *stx* positive water samples during the study period. The extended duration of the study period, as well as the geographical information that was provided with the majority of water samples, allowed for the investigation into patterns of *stx* occurrence in well-sourced non-municipal drinking water samples over space, time, and space-time.

As discussed in Chapter 3, there appeared to be relatively strong seasonal peaks of *stx* positivity in the early summer throughout the study period. Both the monthly *stx* positivity time-series of submitted drinking water samples and of *E. coli* positive Colilert® enriched drinking water samples had statistically significant seasonal trends as determined by the modified Edwards' test, confirming the significance of what was observed in this previous chapter. Since 2005 could be considered an outlier in study period, seasonality tests were carried out on time-

series where the 2005 samples were omitted. These additional Edwards' tests were done to investigate the extent to which 2005 influenced the seasonal trend in *stx* positivity and particularly if the seasonal trend would remain statistical significant if this outlier year was not included. The Edwards' test of the monthly time-series of *stx* positivity in submitted drinking water samples was still statistically significant when the 2005 data was omitted and the persistence of this significant seasonal trend despite the omission of the 2005 data suggests that the underlying endemic *stx* positivity in non-municipal water supplies peaks during the summer months in southern Alberta. The inclusion of the 2005 data both increased the fit of the model and increased the amplitude of the seasonal trend, suggesting that the extreme precipitation event in June 2005 further reinforced the significant seasonality of *stx* positivity in well-water sources.

The interpretations of the results testing for statistically significant seasonal patterns of *stx* positivity within fecally contaminated well-water sources were less straightforward. The abnormally high levels of *stx* positivity in *E. coli* positive Colilert® enriched drinking water samples during 2005 again reinforced the summer peak of *stx* positivity, increasing the amplitude of the harmonic curve, as well as shifting the peak of the temporal pattern towards the summer months when compared to the results without 2005 data included. This lead to a statistically significant seasonal pattern of *stx* positivity in *E. coli* positive Colilert® enriched drinking water samples during the study period, however the Edwards' test result of the monthly time-series of *stx* positivity in *E. coli* positive Colilert® enriched drinking water samples with 2005 data omitted was not statistically significant. The

p-value of this test was only slightly above the set 0.05 significance level however, and there are some notable limitations of the modified Edwards' test used for the analysis of this data. Firstly, the Edwards' family of seasonality tests can struggle to identify significant seasonal patterns within data sets with smaller sample sizes, such as this time-series [254]. In addition, although this family of tests is suggested for use with more erratic data sets such as the monthly time-series under discussion, data presented in Chapter 3 suggests that there may be a second peak of occurrence in the early spring. Due to the 12-month period of the sinusoidal wave tested by this statistic, the Edwards' test can struggle to identify seasonal patterns that have multiple peaks within the data [183,255]. Both the reduced power of this test with small sample sizes and the potential interference of a secondary peak should be taken into account when evaluating the results of this test and caution is warranted when rejecting the existence of a seasonal trend within this data set. Further statistical testing that can handle both the relatively small sample size of this data set and the potential for multiple peaks in a calendar year, while still being able to adjust for the differences in the amount of events that occur from month to month, is needed to definitively determine if there are statistically significant seasonal patterns in the underlying endemic *stx* positivity of *E. coli*-contaminated water supplies. Nevertheless, the significant differences in monthly IRRs outlined in Chapter 3, the significant seasonal trend identified when all the data is included, and the nearly significant p-value of 0.055, all support the validity of the apparent seasonal patterns of monthly *stx* positivity in *E. coli* positive Colilert® enriched drinking water samples.

Similar to the comparisons done in Chapter 3, there are some notable differences in the statistical patterns of seasonality between *stx* occurrence in submitted drinking water samples and *E. coli* positive Colilert® enriched drinking water samples. For example, the amplitudes of the seasonal harmonic curves were much higher for the Edwards' tests of *stx* positivity in submitted drinking water samples than those for the Edwards' tests of *stx* positivity in *E. coli* positive Colilert® enriched drinking water samples. Additionally, the seasonal peaks of *stx* positivity in submitted drinking water samples were both in the summer (July), while the seasonal peaks of *stx* positivity in *E. coli* positive Colilert® enriched drinking water samples were both in the spring (May). These differences again suggest that the seasonal pattern of *stx* positivity in submitted drinking water samples is distinct from the seasonal pattern of fecally-contaminated drinking water samples. Accordingly, the drivers influencing the contamination of the environment with STEC and the drivers influencing the contamination of well-sourced drinking water with STEC, likely act with at least a certain degree of independence from one another.

These findings of statistically significant seasonal peaks in STEC occurrence were not unexpected, as seasonality is a commonly found pattern when investigating waterborne enteric illness, STEC related illness, as well as environmental drinking water contamination and outbreaks [119,126,128–130,134]. Most importantly, and as mentioned in Chapter 3, the summer peaks of STEC detected in non-municipal well-sourced drinking water in this study matched the seasonal pattern of reported STEC-related illness in Canada and Alberta [44,58,128,130]. In

addition, these spring and summer peaks of *stx* contamination happen to coincide with increased incidence of both STEC shedding from zoonotic reservoirs, STEC occurrence in surface waters, as well as enteric drinking water outbreaks across Canada [51,58,141,143]. This synchronicity between seasonal patterns of environmental STEC contamination and seasonal patterns of STEC infection in Canada outlines a potential pathway linking STEC-contaminated drinking water to STEC-related waterborne illness.

Several factors may play a role in establishing this seasonal occurrence of STEC in groundwater. First, the amount of STEC entering the environment in early spring may increase due to the mobilization of fecal wastes (animal feces or human septic discharge) accrued on the frozen landscape [81,156,160,200]. Later-on, during the summer, well-described seasonal patterns of STEC shedding in zoonotic reservoirs peak, again causing a potential influx of pathogens into the environment [139–143,256]. In the spring and early summer, environmental conditions, such as increased temperatures, spring runoff, ground thaw, groundwater recharge, and periods of rainfall, aid the survival of STEC within water and help transport these pathogens both horizontally and vertically through the surrounding landscape [88,96,98,129]. Aided by these favourable conditions STEC can leach into groundwater and, as demonstrated by the present study, result in the increased summer occurrence of STEC in well water. Consumption of contaminated well water during these high-risk periods provides a direct route for disease transmission, and offers a potentially important explanation for the observed seasonal increases of clinically reported STEC- related illness. To the best of our knowledge, this is the first study to

demonstrate a seasonal pattern of STEC contamination of groundwater that aligns with seasonal pattern of STEC disease in humans.

Clustering can result from both the exposure to stable or transient physical, biological, and social risk factors [137]. Stable risk factors would contribute to clustering over the whole duration of the study period, while transient risk factors would act over a much shorter period of time, such as in the case of an outbreak scenario. The space-time cluster found within this data set suggests that there was a *stx* contamination event during the summer of 2005 in the specific region of southwestern Alberta identified in Figure 17. The overlap of the space-time cluster with both the purely temporal and purely spatial clusters suggests that all three of these clusters were related to a single contamination event. The previously mentioned extreme precipitation event that occurred during this time frame (June 2005), and in the region of the identified space-time cluster, is likely an important contributor to the increased frequency of *stx* occurrence detected by the scan statistics. Further investigating the geographic region identified by the space-time cluster may help reveal stable geographic/hydrogeologic characteristics that increase the susceptibility of this area's drinking water supplies to *stx* contamination during extreme precipitation events (i.e., soil type, land use, etc.).

Since the increase in *stx* positivity during 2005 can be largely attributed to a *stx* contamination event, bias is potentially introduced for the temporal and spatial Bernoulli model Kulldorff scan statistics since they do not account for clustering during analysis. In lieu of this potential bias, purely temporal and spatial Bernoulli model Kulldorff scan statistics analyzing the data set without the 2005 data were

conducted. Tests of this adjusted data set would be more representative of sporadic *stx* positivity within non-municipal drinking water supplies in southern Alberta than if the outlying 2005 data were included. The resulting lack of statistically significant temporal clustering suggests that a consistently stable pattern of underlying *stx* positivity occurred during the study period. The change in location of the most likely spatial cluster after the removal of the 2005 data supports the assumption that the contamination event appreciably influenced the original spatial scan statistic. This additional spatial cluster is more likely representative of an area of increased risk for sporadic *stx* contamination than the previously identified most likely spatial cluster that included 2005 data. Any stable characteristics in the geographic region represented by this additional spatial cluster may provide insight into spatial risk factors for sporadic *stx* contamination of non-municipal water supplies and would merit further study. A spatial cluster geographically identical to this additional 'sporadic' cluster was identified by the original spatial scan of the full data set, but was not found to be statistically significant (data not shown). The inclusion of an additional two negative samples within the scanning window contributed to this cluster's decreased statistical significance during the original analysis. It is unlikely however, that the 2005 contamination event, which was geographically separated from this non-event cluster, would have had any influence on the independence the two negative samples excluded from this second 'sporadic' Bernoulli scan (without 2005 data). In light of this consideration, by excluding the 2005 data from this secondary analysis, other bias may have been introduced and the statistical

significance of this additional 'sporadic' spatial cluster should be interpreted with a degree of caution.

Although there are no directly comparable studies using scan statistics to look at clustering of waterborne STEC, Pearl *et al*. [137] did use scan statistics to investigate clustering of reported cases of *E. coli* O157 across Alberta from 2000-2002. All of the purely temporal clusters from this study occurred during the late spring to early autumn, which coincides with the seasonal peak of *stx* occurrence seen in our results and occurs at a similar time of year as the purely temporal cluster identified in the current project. The purely spatial clusters from Pearl *et al*. [137] were all located in southern Alberta but do not overlap with the purely spatial contamination event cluster identified in the present study. However, there are some clusters identified in the Pearl *et al*. [137]study that do overlap with the 'sporadic' spatial cluster, most notably a 2001 cluster of strictly sporadic O157 cases. The two geolocated water samples from which O157:H7 serotypes were recovered from (Figure 15) were located within overlapping spatial clusters of reported cases of *E. coli* O157 from Pearl *et al*. [137], suggesting that these water samples are from areas in Alberta that may be at a higher risk of *E. coli* O157 infection. So *et al.* [257] conducted a similar study to Pearl *et al*. [137] except there was a focus on clusters of a common *E. coli* O157 molecular subtype. Two space-time clusters from this study geographically overlapped with the 2005 space-time cluster from the current study: a year 2000 cluster identified using a space-time permutation scan statistic and a 2001 cluster identified using a Bernoulli spatial scan statistic. It is difficult to draw any meaningful conclusions from these comparisons due to the substantial

differences between these two studies and the current project, however there is seasonal and geographic overlap between reported cases of *E. coli* 0157 illness in southern Alberta and areas of increased *stx* occurrence in non-municipal drinking water samples. Further investigation into spatial and temporal overlap of reported STEC-related illness and the results of this current study is warranted in order to examine any epidemiological links between STEC-contaminated drinking water and clinical STEC cases.

A recently published study by Valeo *et al*. [106] investigated *E. coli* and total coliform contamination of well water in Alberta during a time period captured by the study period of this thesis project. They found that contamination rates peaked in the summer months and were significantly higher during June and July of 2005, similar to the results for *stx* contamination seen in the current study. A purely spatial Bernoulli scan of their June and July 2005 results identified a statistically significant cluster in the southwest of the province that was focused around certain elevations within the Red Deer River and Bow River basins. Similar to the current study, Valeo *et al*. [106] also noted that these 2005 results occurred during, and just after, a period of extreme rainfall in the Red Deer and Bow River Basins, which likely played a role in the increased contamination rates. The most likely space-time permutation cluster from the present study, as well as three statistically significant clusters identified by a space-time permutation scan statistic of strictly 2005 data (results not shown), overlap with the northern most section of the cluster identified by Valeo *et al*. [106]. This overlap suggests that there is an area within the cluster of increased fecal contamination identified by Valeo *et al*. [106] that also had increased

stx contamination within their drinking water supplies during the same time period. Investigating stable geographic differences within these clusters may help to identify specific characteristics that are present in the STEC cluster, which are not present in the remaining area of the cluster identified by Valeo *et al*. [106], and may contribute to the increased risk of STEC contamination, rather than simply fecal contamination, during extreme precipitation events in this area.

There are some potential sampling limitations that should be taken into account when evaluating these scanning results. Firstly, and as mentioned in previous chapters, the volunteer sampling paradigm used for the collection of samples for this project may bias the results in a variety of ways. For example, when viewing Figure 14, it is apparent that the water sample distribution is not uniform across the study area. This non-random distribution may bias results depending on which areas in southern Alberta are more or less represented by the submitted samples.

An important consideration is that a notable number of the water samples were missing or had incomplete submission information, such as the date of sample collection, corresponding ATS information, or both. Ninety-six percent (96.0%) of samples were available for purely temporal scans, 81% were available for purely spatial scans and 79% were available for space-time scans. There were 9 *stx* positive samples that were excluded from the temporal analyses, 24 from the spatial analyses, and 27 from the space-time analyses. It is unknown how the results would have been affected if the full complement of samples could have been included in these scans. Incomplete data submissions are one of the potential shortfalls of

passive surveillance, yet satisfactory sensitivity can still be achieved for low-prevalence diseases with large enough sample sizes [74,206]. As a descriptive project, the ability of these results to provide an estimate of *stx* occurrence in space, time and space-time in non-municipal water sources across southern Alberta over an extended study period still provides useful information for future public health research and decision-making, despite a potentially lower sensitivity in the results of these statistical scan analyses.

Another notable characteristic of the data set with regards to the results of these scan statistics, were single ATS locations that had multiple submitted water samples associated with them. There are a number of challenges in determining the impact of having multiple water samples from the same geographical coordinates. First, it may be that the samples were collected on the same day from the same location. This may represent multiple samples from a single source or single samples from multiple sources located on the same property. Having two samples submitted from a single source on the same day is not ideal, but because of the varying occurrence of pathogens in drinking water and the limited sample volume tested, even samples taken back to back can be considered relatively unique. The effect of this back-to-back sampling would be to essentially increase the volume of water being analyzed from that location when compared to others in the sample set, which could potentially introduce bias towards those specific locations for that day of the study period. Which way this could potentially skew results would depend on the characteristics of the specific source being sampled, if it is at a particularly high or low risk of *stx* contamination. If the multiple samples were from different sources

on the same plot of land, this data would still be important and valid within the scope of the study. The effect of this sampling situation would be to essentially increase the amount of samples from one geographic area compared to the rest of southern Alberta, which again could potentially skew results if the sampling area is at particular risk of *stx* contamination or vice versa. A similar bias, where results are skewed towards the static risk factors of a single location, could be introduced when multiple samples are analyzed from a single location but over a more extended period of time. For example, some of the locations that were associated with multiple samples had samples that were taken years apart from one another. In this case, due to the continuously changing microbiology of the water source, each sample could still provide important information, particularly due to the influence of temporally dependent risk factors such as precipitation or temperature.

CHAPTER 6: General Discussion

6.1 Key findings

The overall purpose of this project was to determine the occurrence of STEC in non-municipal well-sourced drinking water in southern Alberta and to investigate any spatiotemporal patterns of STEC occurrence. Perhaps the most important finding of this project was that a statistically significant summer peak in STEC occurrence within voluntarily submitted drinking water samples was detected, and that this seasonal pattern aligns with well-established seasonal increases in the reported number of clinical STEC cases, raising the proposition that drinking well water may be an important risk factor for STEC disease in Alberta and Canada.

Statistically significant seasonality of STEC occurrence in the early summer was also detected in *E. coli* positive drinking water samples, with two distinct seasonal peaks in the early spring and early summer. In general, STEC was regularly found in non-municipal drinking water sources throughout the study period. The occurrence of *stx* genes, used as a proxy for STEC occurrence, was determined to be 0.2% (152/ 95,675) in all voluntarily submitted drinking water samples and 8.0% (152/1899) in *E. coli* positive drinking water samples. The 2005 calendar year had a significantly greater *stx* positivity rate in non-municipal well-sourced drinking water samples, and the notably increased *stx* positivity rates in non-municipal well-sourced drinking water samples in 2005 and 2013 may be related to extreme

precipitation events that occurred in southern Alberta. The occurrence of viable STEC in drinking water wells was also confirmed. Sixty-five genetically unique STEC representing 21 different STEC serotypes, including 6 of the 7 most clinically relevant serotypes in Alberta, were successfully recovered from *stx* positive Colilert® enriched drinking water samples. Multiple STEC isolates were recovered from five *stx* positive Colilert® enriched drinking water samples. Statistically significant spatiotemporal clusters of STEC occurrence in non-municipal drinking water wells were identified. Most notably was a STEC contamination event northwest of the city of Calgary during the month of June 2005 that coincided in time and space with an extreme precipitation event.

6.2 Overall interpretation and discussion

6.2.1 The occurrence of STEC in non-municipal well-sourced drinking water

The results of this study demonstrate that STEC was regularly found within non-municipal well-sourced drinking water supplies in southern Alberta. The rate of reported cases of STEC infection from 2011-2015 in Canada was 1.9 per 100,000 population, and a conservative estimate of the average rate of all cases of STEC infection in Canada (reported and non-reported) was 100 per 100,000 population [24,25]. The rate of reported cases of STEC infection from 2011-2016 in Alberta was 3.2 per 100,000 population [27]. By comparison, the *stx* positivity rate of submitted drinking water samples during the study period (March 2004 – July 2016) was 200 per 100,000 drinking water samples. According to the Guidelines for Canadian Drinking Water Quality, if any *E. coli* bacteria, let alone STEC, are detected in 100ml

of water it is deemed unfit for consumption and the Microbial Guidelines for Ready-to-Eat Foods states that if any STEC is detected food it is unfit for consumption [75,258]. Clearly, any STEC positive drinking water sample should be considered a serious threat to the health of the consumer. Furthermore, a notable portion of the STEC-contaminated drinking water samples tested for this study were found to contain STEC bacteria that are more commonly associated with serious human illness, whether that be *stx2*-harbouring STEC, or STEC serotypes belonging to the 'big six' non-O157 and O157 serogroups. The majority of cases of STEC infection go unreported in Canada however, and therefore the impact of the "rare" STEC serotypes recovered from these drinking water samples may still play a meaningful role in the overall burden of disease attributed to this pathotype [24,69]. The consistent occurrence of these notifiable human pathogens in groundwater wells, a water source for which 450,000-600,000 Albertans rely on for drinking water, suggests that non-municipal drinking water may be an important source of STEC-related illness that may be overlooked in the province and potentially across the country [69,71,72].

6.2.2 Spatiotemporal patterns of occurrence of STEC in non-municipal well-sourced drinking water

Another indication that groundwater may be an underrated source of STEC infections is the coordinated seasonal patterns of both STEC occurrence in non-municipal drinking water sources and of STEC related illness in both Canada and the province of Alberta [44,130,138]. Summer increases in cases of STEC are well established

in temperate climates and while potentially seasonal risk factors such as the use of recreational waters and travel to rural areas have been associated with STEC infections, this seasonality has been commonly attributed to poor outdoor cooking practices, particularly the undercooking of ground beef [93,119,128,131,134,195,196,259–262]. There has been a very limited amount of research directly investigating this plausible hypothesis however, likely due to the difficulties of attributing a portion of STEC cases specifically to poor outdoor cooking practices and subsequently determining any seasonal patterns associated with this individual risk factor [259]. Williams *et al.* [259] did investigate this hypothesis and found nearly proportional relationships between seasonal trends in *E. coli* O157:H7 in live cattle, raw ground beef, and humans, which described a logical pathway to human illness. Through comparison however, Williams *et al.* [259] were able to determine that the probability of illness from consumption of a serving of contaminated ground beef remained roughly constant throughout the year, suggesting that seasonal outdoor food practices are not the main driver of the summer peak of O157:H7 infection. Instead, it was the seasonal change in the probability of exposure to a contaminated serving of ground beef that was the driver of this seasonality, which in turn, was likely driven by the seasonal increase in *E. coli* O157:H7 found in live cattle. This proportional relationship between seasonal patterns of *E. coli* O157:H7 occurrence described by Williams *et al.* [259] is similar to the relationship described in this current study between seasonal patterns of STEC occurrence in non-municipal drinking water sources and reported STEC infection in Alberta and Canada, with both relationships highlighting a logical pathway to human illness. Seasonal increases of

STEC shedding and subsequently STEC in the environment, align with seasonal increases in STEC occurrence in groundwater wells that would increase the probability of exposure to a contaminated serving of drinking water, which could very well act as a driver of the seasonal pattern of STEC-related human illness. It is important to note however, that like the Williams *et al.* [259] study, these correlations strictly outline a potential pathway to infection and cannot support direct causality. While untreated drinking water sources and private well systems have been previously identified as risk factors of enteric waterborne disease [61,66,69], to the best of our knowledge this is the first study to suggest a seasonal link between the occurrence of STEC infections and the occurrence of STEC in non-municipal well-sourced drinking water.

It is unknown to what degree increased exposure to contaminated non-municipal drinking water may contribute to the seasonal pattern of STEC-related illness in Alberta. However, the importance of environmental pathways to the development of STEC and enteric infection, particularly in comparison to foodborne pathways, is becoming increasingly apparent [195,196,260,263–265]. For example, recent case-control studies by Denno *et al.* [195] and Jaros *et al.* [196] have found that environmental exposures were equally or more important to the development of sporadic enteric illness in specific populations, and studies from Scotland found that the risk of *E. coli* O157 infection from visiting ruminant pastures is approximately equal to eating undercooked ground beef, and is 100 times higher than eating hamburgers in general [48,264]. Rotariu *et al.* [48] used both regression and quantitative microbial risk assessment (QMRA) models to determine the relative importance of

foodborne and environmental transmission pathways for *E. coli* O157 infection in northeastern Scotland. Environmental pathways were separated into direct environmental exposure (i.e., a day visiting a ruminant pasture) and waterborne exposure (i.e., drinking a glass of water from a private water source), and for each model waterborne exposures were found to be less important than both direct environmental exposures and foodborne exposures. It is important to note however, that the Rotariu *et al.* [48] study did not account for non-O157 STEC infections and occurrence, which are the STEC serotypes responsible for the majority of reported STEC cases in Alberta and accounted for 97% of the STEC isolates recovered from the drinking water samples of this current study [43,45]. Although literature comparing transmission pathways for both O157 and non-O157 is limited, non-O157 STEC seem to be more generally associated with environmental pathways (i.e., animal contact) and less so with foodborne pathways, while O157 STEC show the opposite associations [42,263,266–269]. Geographical differences in common risk factors of STEC infection have been previously reported and the prevalence of non-O157 associated illness in Alberta may indicate that environmental infection pathways are particularly important in this region [263,270].

In addition to potentially greater risk of STEC infection due to the summer increase of STEC contamination in non-municipal drinking water sources, the exposure to this potential source of infection likely increases during the summer as well, due to seasonal activities such as the use of campgrounds and vacation properties that rely on this type of drinking water source. Furthermore, while the risk associated with visiting farms or rural areas has thus far been primarily linked

to direct exposure to animal carriers and/or their excrement, part of this increased risk may be due to the consumption of contaminated drinking water [93,196,260,271]. Rural visitors and new rural residents, particularly those who are coming from an urban environment, may be at greater risk of illness if exposed to STEC-contaminated drinking water due to a lack of protective immunity that is often developed among long-term rural residents [47,76,271–273]. This may be one reason that the most frequently identified settings of enteric outbreaks associated with groundwater in the USA are camps, cabins, or recreational areas; rural areas often visited by urban residents [50]. While the nature of protective STEC immunity is not well understood, this phenomenon is well documented across the globe and commonly occurs in environments at a higher risk of STEC exposure, such as rural settings and areas relying on non-municipal groundwater systems [47,65,273–279]. The hypothesized conditions for the development of protective STEC immunity are low-dose long-term STEC exposures, potentially to less virulent STEC strains, which causes repeated antigenic stimulation and the resulting production of STEC-specific antigens [270,275,277]. The apparent clustering of STEC-antigen seropositivity within families suggests that common exposures (i.e., drinking water source), common practices, or person-to-person transmission, may be important in the development of this immunity [271]. For example, Karmali *et al.* [280] described a Canadian outbreak of STEC O111:NM illness that caused HUS in a 9-month old and uncomplicated diarrhea in the 2-year old sibling and both parents, all of whom tested negative for immune antibodies specific to Shiga toxins. The epidemiological investigation concluded that this outbreak was likely caused by the consumption of

unpasteurized milk while visiting rural family members. None of the rural family members, who regularly consume unpasteurized milk and consumed the same milk that caused the outbreak, became ill. Seven out of the eight rural family members tested positive for anti-STEC antibodies and STEC O111:NM was even recovered from one of these asymptomatic individuals. Licence *et al.* [65] described a waterborne outbreak of six STEC O157 cases in tourists visiting a campsite in Scotland over seven weeks that were linked to a private water supply in an agricultural area. STEC O157 of the same phage type and with identical pulsed field gel electrophoresis profiles were isolated from the cases, the campsite drinking water, sheep feces from around the water source, and drinking water from one of five houses in the area that relied on the same water supply. Despite consuming the water from the same contaminated source as the campsite over a much longer exposure period, none of the local residents of these five houses experienced any gastrointestinal symptoms during this outbreak (testing for STEC-specific antibodies was not done in this study). As exemplified above, a well water purveyor could unknowingly expose immunologically-naïve individuals to poor-quality drinking water because their own protective immunity to STEC would prevent them from developing symptoms and subsequently determining that the drinking water is contaminated. This potential exposure route highlights the importance of regular water testing even in the absence of symptoms indicating poor water quality, particularly during the spring and summer months. In addition, residents who have developed a protective immunity and are regularly exposed to STEC through their drinking water could also become asymptomatic carriers, a state that is repeatedly associated with STEC

outbreaks and sporadic cases, and unknowingly infect others though person to person transmission [1,26,263,281].

Interestingly, the results of the present study not only describe a seasonal pattern in the frequency of occurrence of STEC in well-sourced drinking water, but also suggest that there is a seasonal pattern in the risk profile of general fecal contamination with respect to STEC. This is demonstrated by differences in the seasonal *stx* positivity rates of submitted drinking water samples and seasonal *stx* positivity rates in the subset of fecally-contaminated drinking water samples (i.e., *E. coli* positive). Unlike the single statistically significant summer pulse of STEC occurrence in submitted drinking water samples, fecally-contaminated drinking water samples had two seasonal pulses, one in the early spring and the other in the early summer. Surprisingly, the spring pulse was the more prominent contributor, particularly when outlying data (2005 data) was omitted from the analysis. These seasonal pulses signify increases in the amount of total STEC occurring within fecal inputs to the environment during the early spring and summer, however the statistical significance of these two pulses in *E. coli* positive drinking water samples could not be adequately determined due to limited sample sizes. One possible explanation for this increase in the early spring is the potential mobilization of accrued fecal bacteria during the spring thaw, known as the first flush effect [282]. Southern Alberta is a snowmelt-dominated hydroclimatic regime where peak streamflow and groundwater recharge are primarily controlled by melting-snowpack and glaciers and therefore, peak flow and recharge occur in the spring and early summer [160,283]. While *E. coli* is not expected to grow or thrive during the

winter months in Alberta, *E. coli* have been shown to survive in frozen soils over winter and Baker-Ismali [282] found that in the Canadian prairies, fecal coliforms exhibited a first flush effect with approximately 38-66% of the bacterial load transported in the first 30% of snowmelt runoff volume [106,284,285]. This spring runoff concentrated with fecal bacteria (STEC included) usually occurs during March and April and, either through leaching into groundwater during spring recharge or directly contaminating improperly sealed wells, could explain the greater *stx* positivity rates in fecally-contaminated well water samples during these same months. From a public health perspective this spring increase of STEC in *E. coli* positive drinking water is important because it would suggest that although the early spring may not be the period where general fecal contamination of wells is occurring most frequently, it is a time period where fecally-contaminated wells could be at particular risk of harboring STEC and causing infection. The second seasonal pulse in the summer may be representative of well-established seasonal increases of STEC shedding in live cattle during the summer, and the greater number of STEC entering the landscape would likely also increase the number of fecally-contaminated drinking water samples that would contain STEC [116,139,140,143,145,146,259]. The nadir in *stx* positivity in *E. coli* positive samples that consistently occurred in May could be representative of the period between the STEC loading from the first flush effect and STEC loading from the summer peak of STEC shedding in cattle. It is important to note however that many of the climactic factors that may influence the survival and transport of STEC within the external and aquatic environment have been associated with lag periods between their

influence in the environment and the occurrence of associated enteric illness [158,160,286,287]. For example, Galway *et al.* [200] found that increased streamflow was positively associated with laboratory-confirmed cases of acute gastro-intestinal illness with a one-month lag period for a community within a snowmelt-dominated hydroclimatic regime in British Columbia, Canada. While the existence of lag-periods in the seasonal occurrence of enteric disease is generally accepted, the determining factors influencing the length of these periods are not well understood, and therefore it is difficult to hypothesize the degree to which the occurrence of STEC contamination of well-water in southern Alberta would lag behind the mobilization of these pathogens in the external environment [155].

Similar to two other studies investigating *E. coli* contamination in well water sources in Alberta, the results of this study also indicated the potential of extreme precipitation events to be catalysts for periods of increased *stx* contamination in well-sourced drinking water [74,106]. The periods both during, and following, these events would also likely be times of increased risk of STEC exposure and infection for these drinking water sources. Differences in positivity rates between years with extreme precipitation events suggest that event-specific characteristics, likely associated with the ways in which STEC are mobilized and transported into groundwater, impact the severity of the risk associated with consuming non-municipal drinking water. In 2013 for instance, the potential impact of the extreme precipitation event did not increase the *stx* positivity rate per *E. coli* positive samples, but instead it was the increase in general fecal contamination (i.e., *E. coli* positivity) within submitted drinking water samples that likely drove the overall

increase in *stx* positivity in submitted drinking water samples. Conversely, the potential impact of the extreme precipitation event of 2005 increased both *stx* positivity in fecally-contaminated drinking water samples as well as general fecal contamination of submitted drinking water samples. Together this contamination contributed to a statistically significant increase in STEC occurrence in submitted drinking water samples and likely a corresponding increase in the associated risk of infection.

From a purely spatial perspective there was evidence of geographic areas in the southern part of the province that were at a potentially higher risk of sporadic STEC contamination. Methodological bias may have influenced the statistical significance of this spatial cluster however, and accordingly, this result should be interpreted with a degree of caution. Previous research has identified a number of localized, smaller-scale, spatial risk factors related to the contamination of groundwater wells with fecal pathogens, such as the separation distance between the well and animal yards, the use of septic systems for household waste disposal, and the immediate surrounding hydrogeological conditions of the well [66,77,103,252]. The scarcity of statistically significant purely spatial clusters identified by the Kulldorff models suggest that the geographic risk factors of sporadic STEC contamination are too localized around individual wells to be properly identified at the resolution of these spatial scan statistics, which supports previous findings that well-specific contamination mechanisms are the most influential for bacterial contamination of well-water [66,109]. Following this reasoning, investigating stable

geographic risk factors of sporadic STEC contamination for individual STEC positive wells may be a useful focus of future spatial investigations.

It is extremely important to note that while the present study focused specifically on STEC, these findings warrant a more thorough investigation into the seasonal patterns of other waterborne pathogens in non-municipal well-sourced drinking water. The time, space, and space-time observations from this study, the plausible environmental transport model outlined previously, along with the supporting spatiotemporal evidence of enteric disease prevalence in southern Alberta, together suggest that non-municipal groundwater systems may be an important risk factor not only for STEC, but for a variety of fecal-orally transmitted waterborne diseases. *Cryptosporidium* spp., *Giardia* spp., and *Campylobacter* spp. are pathogens that are commonly found in groundwater, and other pathotypes of *E. coli* such as ETEC and EPEC have also been associated with waterborne disease [1,69,70,126,128,131]. While these other waterborne pathogens may not be associated with the same risk of serious illness as STEC, many are considerably more prevalent in Canada, and accordingly, their occurrence in non-municipal drinking water sources may be greater as well [24,69].

These results highlight the incredible complexity of risk analysis for pathogens within well-sourced drinking water. Taken together, the information provided by these results can help to educate well purveyors about risks associated with non-municipal drinking water and promote proper well management, encouraging not only regular water testing but also water testing during periods of increased risk of STEC contamination and infection.

6.2.3 Recovery of viable STEC

This research has also served to highlight some of the challenges associated with the investigation of environmental pathogens, particularly the difficulties of applying methods and technologies that have been developed and evaluated for clinical use to the study of environmental samples. Samples from environmental settings may contain microbiota that react appreciably different to laboratory methods and conditions when compared to the microbiota of their clinical counterparts. It is important to continue to build upon the limited amount of current research investigating all STEC serotypes in drinking water. Doing so will allow for more meaningful comparisons between studies and the further refinement of specific methods for environmental strains of this *E. coli* pathotype. The present study determined that the use of multiple incubation temperatures during recovery protocols helped to improve isolation and growth of environmental STEC strains on selective agar and more specifically, that the use of an incubation temperature near the upper metabolic limit of these bacteria helped to both reduce ubiquitous growth of background flora and provide growth conditions better suited to certain environmental STEC strains. It is suggested that these methodological considerations be incorporated into future recovery protocols for environmentally sourced STEC.

6.3 Study strengths and limitations

The main strengths of this study included: i) the large sample size of voluntarily submitted drinking water samples that helped to increase the power of

analysis when studying relatively rare pathogens; ii) the extended duration of the study period that allowed investigation of aggregated time-series, iii) the archival of viable environmental samples that allowed the recovery and subsequent characterization of live pathogens; and iv) the ability to geographically locate submitted samples for geospatial analysis and mapping.

There were a number of notable limitations to the study that have been outlined in previous chapters, but briefly: the relatively rare occurrence of *stx* positive samples coupled with smaller sample sizes of *E. coli* positive drinking water samples limited the power of some Poisson-based regression models investigating the relationship of *stx* positivity rates and time, the use of convenience sampling (i.e., voluntarily submitted samples) may introduce bias into the sample set, using a molecular screen for detecting bacteria does not necessarily denote the presence of viable bacteria, and not all submitted samples had corresponding temporal and spatial data available for analysis. While these limitations are important to consider, they should not detract from the importance of the general findings of this project, particularly since the overall effect of these limitations would likely be to provide a more conservative estimate of STEC occurrence in drinking water wells across southern Alberta.

6.4 Future directions

There are a number of avenues of future research related to this thesis that would be worthwhile to investigate. Conducting a more comprehensive examination of other STEC virulence factors (i.e., *eae*, *hlyA*, etc.) in both archived *stx* positive drinking water samples and recovered STEC isolates would be useful. The

identification of detailed virulence profiles would allow for a more thorough assessment of the risk associated with exposures to STEC-contaminated well water in general, as well as each recovered STEC strain specifically. Each STEC isolate could be whole genome sequenced, which would not only provide the complete virulence profile for each strain, but would also allow for the potential identification of genes associated with environmental survival in groundwater or previously identified source tracking markers. As mentioned above, further investigation into the occurrence of additional waterborne pathogens (i.e., additional *E. coli* pathotypes) in the archived Colilert® positive drinking water samples collected for this study would help to better characterize the total risk associated the consumption of non-municipal well-sourced drinking water in southern Alberta. As well, attempting to isolate STEC from archived *stx* positive Colilert® enrichments using alternate recovery methods, such as the STEC-immnoblot method or a labelled-probe analysis, could increase the recovery rate of viable STEC from these samples and provide a more accurate estimation of the STEC population found within the voluntarily submitted drinking water samples [51,288].

One of the most surprising results of this study was the number of *stx* negative mauve coloured colonies that grew on the CA-STEC agar. Identifying these archived *stx* negative mauve CA-STEC isolates would help to increase our understanding of the selective mechanisms of CA-STEC, further refine this STEC isolation protocol, and potentially identify other enteric pathogens (i.e., other *E. coli* pathotypes) occurring in these well-sourced drinking water samples.

Kulldorff scanning statistics are just one of many investigative models and techniques possible using GIS software, and additional analysis, such as the development of a relative risk map of STEC occurrence in non-municipal water sources in southern Alberta, could be completed to better inform future water quality surveillance and management efforts. In addition and as discussed above, the examination of stable spatial characteristics within identified clusters of STEC-positive drinking water samples could also help identify important spatial risk factors associated with increased contamination of non-municipal drinking water supplies. The occurrence of recovered STEC strains in space and time could also be compared to similar cases of reported STEC infection in the province to examine any serotype or virulence profile similarities and investigate potential epidemiological linkages to this infection pathway.

6.5 Conclusions

This project successfully investigated the frequency of STEC occurrence, the spatiotemporal patterns of STEC occurrence, and the characteristics of recovered STEC strains in non-municipal well-sourced drinking water from southern Alberta. As a descriptive study, it fills an important knowledge gap surrounding not only STEC in groundwater and in drinking water wells, but also waterborne enteric pathogens more generally. STEC was found to be a relatively common contaminant in voluntarily submitted non-municipal well-sourced drinking water samples from southern Alberta and represents an important public health risk to the hundreds of thousands of Albertans, and potentially the millions of Canadians, that rely on groundwater wells for their drinking water supplies. This project provided evidence

to support the seasonal influence that contaminated non-municipal drinking water sources may have on patterns of STEC infection in the province and outlined the potential of these drinking water sources to act as an important infection pathway for waterborne enteric pathogens, particularly in the summer months. Both the time-periods and geographical areas that were identified as having potentially increased risk of STEC contamination in well-sourced drinking water can be leveraged to improve individual management and public health surveillance of non-municipal drinking water sources. This study also provides a starting point for many areas of future investigation within the fields of environmental microbiology and waterborne enteric illness.

References Cited

1. Croxen, M. A. *et al.* Recent advances in understanding enteric pathogenic *Escherichia coli*. *Clin. Microbiol. Rev.* **26,** 822–880 (2013).
2. Nataro, J. P. & Kaper, J. B. Diarrheagenic *Escherichia coli*. *Clin. Microbiol. Rev.* **11,** 142–201 (1998).
3. DebRoy, C., Roberts, E. & Fratamico, P. M. Detection of O antigens in *Escherichia coli*. *Anim. Heal. Res. Rev.* **12,** 169–185 (2011).
4. Lee, M. S., Cherla, R. P. & Tesh, V. L. Shiga toxins: Intracellular trafficking to the ER leading to activation of host cell stress responses. *Toxins (Basel).* **2,** 1515–1535 (2010).
5. Kruger, A. & Lucchesi, P. M. A. Shiga toxins and *stx* phages: highly diverse entities. *Microbiology* **161,** 451–462 (2014).
6. Gamage, S. D., Patton, A. K., Hanson, J. F. & Weiss, A. A. Diversity and host range of Shiga toxin-encoding phage. *Infect. Immun.* **72,** 7131–7139 (2004).
7. Boerlin, P. *et al.* Associations between virulence factors of Shiga toxin-producing *Escherichia coli* and disease in humans. *J. Clin. Microbiol.* **37,** 497–503 (1999).
8. Bettelheim, K. A. & Goldwater, P. N. Serotypes of non-O157 Shigatoxigenic *Escherichia coli* (STEC). *Adv. Microbiol.* **4,** 377–389 (2014).
9. Mora, A. *et al.* Seropathotypes, phylogroups, stx subtypes, and intimin types of wildlife-carried, Shiga toxin-producing *Escherichia coli* strains with the same characteristics as human-pathogenic isolates. *Appl. Environ. Microbiol.* **78,** 2578–2585 (2012).
10. Melton-Celsa, A. R. Shiga toxin (Stx) classification, structure, and function. *Microbiol. Spectr.* **2,** 1–13 (2014).
11. Tarr, P. I., Gordon, C. A. & Chandler, W. L. Shiga-toxin-producing *Escherichia coli* and haemolytic uraemic syndrome. *Lancet* **365,** 1073–1086 (2005).
12. Friedrich, A. W. W. *et al. Escherichia coli* harboring Shiga toxin 2 gene variants: frequency and association with clinical symptoms. *J. Infect. Dis.* **185,** 74–84

(2002).

13. Herold, S., Karch, H. & Schmidt, H. Shiga toxin-encoding bacteriophages - Genomes in motion. *Int. J. Med. Microbiol.* **294,** 115–121 (2004).
14. Muniesa, M., Jofre, J. & García-Aljaro, C. Occurrence of *Escherichia coli* O157: H7 and other enterohemorrhagic *Escherichia coli* in the environment. *Sci. Technol.* **40,** 7141–7149 (2006).
15. Beutin, L. & Martin, A. Outbreak of Shiga toxin–producing *Escherichia coli* (STEC) O104: H4 infection in Germany causes a paradigm shift with regard to human pathogenicity of STEC strains. *J. Food Prot.* **75,** 408–418 (2012).
16. Rasko, D. A. *et al.* Origins of the *E. coli* strain causing an outbreak of hemolytic–uremic syndrome in Germany. *N. Engl. J. Med.* **365,** 709–717 (2011).
17. Mellmann, A. *et al.* Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. *PLoS One* **6,** 1–9 (2011).
18. Smith, J. L., Fratamico, P. M. & Gunther, N. W. in *Advances in Applied Microbiology* **86,** 145–197 (Elsevier Inc., 2014).
19. Johnson, K. E., Thorpe, C. M. & Sears, C. L. The emerging clinical importance of non-O157 Shiga toxin-producing *Escherichia coli*. *Clin. Infect. Dis.* **43,** 1587–1595 (2006).
20. Hunter, P. R. Drinking water and diarrhoeal disease due to *Escherichia coli*. *J. Water Health* **1,** 65–72 (2003).
21. Strachan, N. J. C., Fenlon, D. R. & Ogden, I. D. Modelling the vector pathway and infection of humans in an environmental outbreak of *Escherichia coli* O157. *FEMS Microbiol. Lett.* **203,** 69–73 (2001).
22. Crockett, C. S., Haas, C. N., Fazil, A., Rose, J. B. & Gerba, C. P. Prevalence of shigellosis: In the U.S. consistency with dose-response information. *Int. J. Food Microbiol.* **30,** 87–99 (1996).
23. Majowicz, S. E. *et al.* Global incidence of human Shiga toxin–producing *Escherichia coli* infections and deaths: a systematic review and knowledge synthesis. *Foodborne Pathog. Dis.* **11,** 447–455 (2014).

24. Thomas, M. K. *et al.* Estimated numbers of community cases of illness due to *Salmonella*, *Campylobacter* and verotoxigenic *Escherichia coli*: Pathogen-specific community rates. *Can. J. Infect. Dis. Med. Microbiol.* **17,** 229–234 (2006).
25. Government of Canada. Reported cases from 1991 to 2015 in Canada - Notifiable diseases on-line. (2017). at <http://diseases.canada.ca/notifiable/charts?c=yl>
26. Galanis, E. *et al. Investigation of an E. Coli O157:H7 outbreak in Brooks, Alberta, June-July 2002: the role of occult cases in the spread of infection within a daycare setting*. *Canada Communicable Disease Report* **29,** (2003).
27. Government of Alberta. Alberta Interactive Health Data Application - Notifiable Diseases - Age-Sex Specific Incidence Rate (common diseases). (2017). at <http://www.ahw.gov.ab.ca/IHDA_Retrieval/selectSubCategoryParameters.do>
28. Hrudey, S. E., Payment, P., Huck, P. M., Gillham, R. W. & Hrudey, E. J. A fatal waterborne disease epidemic in Walkerton, Ontario: Comparison with other waterborne outbreaks in the developed world. *Water Sci. Technol.* **47,** 7–14 (2003).
29. Livernois, J. *The Economic Costs of the Walkerton Water Crisis*. *The Walkerton Inquiry* (2002).
30. Huovinen, E. *et al.* Excess healthcare costs of a large waterborne outbreak in Finland. *Scand. J. Public Health* **41,** 761–766 (2013).
31. Luna-Gierke, R. E. *et al.* Multiple-aetiology enteric infections involving non-O157 Shiga toxin-producing *Escherichia coli* - FoodNet, 2001-2010. *Zoonoses Public Health* **61,** 492–498 (2014).
32. Bettelheim, K. A. The non-O157 Shiga-toxigenic (verocytotoxigenic) *Escherichia coli*; under-rated pathogens. *Crit. Rev. Microbiol.* **33,** 67–87 (2007).
33. Frank, C. *et al.* Epidemic profile of Shiga-toxin–producing *Escherichia coli* O104: H4 outbreak in Germany. *N. Engl. J. Med.* **365,** 1771–80 (2011).
34. Luna-Gierke, R. E. *et al.* Outbreaks of non-O157 Shiga toxin-producing

Escherichia coli infection: USA. *Epidemiol. Infect.* **142,** 2270–2280 (2014).

35. Chui, L. *et al.* Comparison between ImmunoCard STAT!® and real-time PCR as screening tools for both O157:H7 and non-O157 Shiga toxin-producing *Escherichia coli* in Southern Alberta, Canada. *Diagn. Microbiol. Infect. Dis.* **77,** 8–13 (2013).
36. Gould, L. H. *et al.* 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,** 453–60 (2013).
37. Brooks, J. T. *et al.* Non-O157 Shiga toxin-producing *Escherichia coli* infections in the United States, 1983-2002. *J. Infect. Dis.* **192,** 1422–1429 (2005).
38. Farfan, M. J. & Torres, A. G. Molecular mechanisms that mediate colonization of Shiga toxin-producing *Escherichia coli* strains. *Infect. Immun.* **80,** 903–913 (2012).
39. Karmali, M. A., Gannon, V. & Sargeant, J. M. Verocytotoxin-producing *Escherichia coli* (VTEC). *Vet. Microbiol.* **140,** 360–370 (2010).
40. Chui, L. *et al.* Molecular profiling of *Escherichia coli* O157: H7 and non-O157 strains isolated from humans and cattle in Alberta, Canada. *J. Clin. Microbiol.* **53,** 986–990 (2015).
41. Frank, C., Kapfhammer, S., Werber, D., Stark, K. & Held, L. Cattle density and Shiga toxin-producing *Escherichia coli* infection in Germany: increased risk for most but not all serogroups. *Vector borne zoonotic Dis.* **8,** 635–643 (2008).
42. McPherson, M. *et al.* Serogroup-specific risk factors for Shiga toxin-producing *Escherichia coli* infection in Australia. *Clin. Infect. Dis.* **49,** 249–256 (2009).
43. Chui, L. *et al.* Prevalence of Shiga toxin-producing *Escherichia coli* as detected by enzyme-linked immunoassays and real-time PCR during the summer months in Northern Alberta, Canada. *J. Clin. Microbiol.* **49,** 4307–4310 (2011).
44. Couturier, M. R., Lee, B., Zelyas, N. & Chui, L. Shiga-toxigenic *Escherichia coli* detection in stool samples screened for viral gastroenteritis in Alberta, Canada. *J. Clin. Microbiol.* **49,** 574–578 (2011).
45. Chui, L., Patterson-Fortin, L., Kuo, J., Li, V. & Boras, V. Evaluation of enzyme

immunoassays and real-time PCR for detecting Shiga toxin-producing *Escherichia coli* in Southern Alberta, Canada. *J. Clin. Microbiol.* **53,** 1019–1023 (2015).

46. Vally, H. *et al.* Epidemiology of Shiga toxin producing *Escherichia coli* in Australia, 2000-2010. *BMC Public Health* **12,** 63 (2012).
47. Olsen, S. J. *et al.* A waterborne outbreak of *Escherichia coli* O157:H7 infections and hemolytic uremic syndrome: Implications for rural water systems. *Emerg. Infect. Dis.* **8,** 370–375 (2002).
48. Rotariu, O. *et al.* Combining risk assessment and epidemiological risk factors to elucidate the sources of human *E. coli* O157 infection. *Epidemiol. Infect.* **140,** 1414–1429 (2012).
49. Chekabab, S. M., Paquin-Veillette, J., Dozois, C. M. & Harel, J. The ecological habitat and transmission of *Escherichia coli* O157:H7. *FEMS Microbiol. Lett.* **341,** 1–12 (2013).
50. Wallender, E. K., Ailes, E. C., Yoder, J. S., Roberts, V. a. & Brunkard, J. M. Contributing factors to disease outbreaks associated with untreated groundwater. *Groundwater* **52,** 886–897 (2013).
51. Johnson, R. P. *et al.* Detection and prevalence of verotoxin-producing *Escherichia coli* O157 and Non-O157 serotypes in a Canadian watershed. *Appl. Environ. Microbiol.* **80,** 2166–2175 (2014).
52. Pebody, R. G. *et al.* An international outbreak of Vero cytotoxin-producing *Escherichia coli* O157 infection amongst tourists; a challenge for the European infectious disease surveillance network. *Epidemiol. Infect.* **123,** 217–223 (1999).
53. Auld, H., MacIver, D. & Klaassen, J. Heavy rainfall and waterborne disease outbreaks: The Walkerton example. *J. Toxicol. Environ. Heal. - Part A* **67,** 1879–1887 (2004).
54. Garvey, P., McKeown, P., Carroll, A. & McNamara, E. Epidemiology of verotoxigenic *E. coli* in Ireland, 2008. *Epi Insight* **10,** 63–69 (2009).
55. Saxena, T., Kaushik, P. & Krishna Mohan, M. Prevalence of *E. coli* O157:H7 in water sources: an overview on associated diseases, outbreaks and detection

methods. *Diagn. Microbiol. Infect. Dis.* **82,** 249–264 (2015).

56. Rogan, W. J. & Brady, M. T. Drinking water from private wells and risks to children. *Pediatrics* **123,** e1123–e1137 (2009).
57. Flint, J. A., Doré, K., Majowicz, S. E., Edge, V. L. & Sockett, P. From stool to statistics: Reporting of acute gastrointestinal illnesses in Canada. *Can. J. Public Heal.* **95,** 309–313 (2004).
58. Schuster, C. J. *et al.* Infectious disease outbreaks related to drinking water in Canada, 1974-2001. *Can. J. Public Heal.* **96,** 254–258 (2005).
59. Uhlmann, S. *et al.* Where's the pump? Associating sporadic enteric disease with drinking water using a geographic information system, in British Columbia, Canada, 1996-2005. *J. Water Health* **7,** 692–698 (2009).
60. Richardson, H. Y., Nichols, G., Lane, C., Lake, I. R. & Hunter, P. R. Microbiological surveillance of private water supplies in England - The impact of environmental and climate factors on water quality. *Water Res.* **43,** 2159–2168 (2009).
61. Beer, K. D. *et al.* Surveillance for waterborne disease outbreaks associated with drinking water—United States, 2011–2012. *MMWR. Morb. Mortal. Wkly. Rep.* **64,** 842–848 (2015).
62. Craun, G. F. *et al.* Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. *Clin. Microbiol. Rev.* **23,** 507–528 (2010).
63. Pearl, D. L. *et al.* Epidemiological characteristics of reported sporadic and outbreak cases of *E. coli* O157 in people from Alberta, Canada (2000-2002): methodological challenges of comparing clustered to unclustered data. *Epidemiol. Infect.* **136,** 483–91 (2008).
64. Craun, G. F., Calderon, R. L. & Nwachku, N. in *Drinking Water and Infectious Disease: Establishing the Links* (eds. Hunter, P., Waite, M. & Rochi, E.) 105–117 (CRC Press and IWA Publishing, 2003).
65. Licence, K., Oates, K. R., Synge, B. A. & Reid, T. M. S. An outbreak of *E. coli* O157 infection with evidence of spread from animals to man through contamination of a private water supply. *Epidemiol. Infect.* **126,** 135–8 (2001).
66. Hynds, P. D., Thomas, M. K. & Pintar, K. D. M. Contamination of groundwater

systems in the US and Canada by enteric pathogens, 1990-2013: a review and pooled-analysis. *PLoS One* **9,** e93301 (2014).

67. Corkal, D., Schutzman, W. C. & Hilliard, C. R. Rural water safety from the source to the on-farm tap. *J. Toxicol. Environ. Health* **67,** 1619–1642 (2004).
68. Risebro, H. L., Breton, L., Aird, H., Hooper, A. & Hunter, P. R. Contaminated small drinking water supplies and risk of infectious intestinal disease: A prospective cohort study. *PLoS One* **7,** e42762 (2012).
69. Murphy, H. M. *et al.* Estimating the burden of acute gastrointestinal illness due to *Giardia*, *Cryptosporidium*, *Campylobacter*, *E. coli* O157 and norovirus associated with private wells and small water systems in Canada. *Epidemiol. Infect.* **144,** 1355–1370 (2016).
70. Butler, A. J., Pintar, K. D. M. & Thomas, M. K. Estimating the relative role of various subcategories of food, water, and animal contact transmission of 28 enteric diseases in Canada. *Foodborne Pathog. Dis.* **13,** 57–64 (2016).
71. Summers, R. J. *Alberta Water Well Survey. A report prepared for Alberta Environment*. (2010).
72. Alberta Health. *Domestic Well Water Quality in Alberta, Fact Sheets*. (2014).
73. Health Canada. *Guidance on waterborne bacterial pathogens*. (Water, Air and Climate Change Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, 2013). at <https://www.canada.ca/content/dam/canada/health-canada/migration/healthy-canadians/publications/healthy-living-vie-saine/water-pathogens-pathogenes-eau/alt/water-pathogens-pathogenes-eau-eng.pdf>
74. Invik, J., Barkema, H. W., Massolo, A., Neumann, N. F. & Checkley, S. Total coliform and *Escherichia coli* contamination in rural well water: analysis for passive surveillance. *J. Water Health* **15,** 729–740 (2017).
75. Health Canada. *Guidelines for Canadian drinking water quality: guideline technical document - Escherichia coli*. (Water, Air and Climate Change Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, 2012). at <https://www.canada.ca/content/dam/canada/health-

canada/migration/healthy-canadians/publications/healthy-living-vie-saine/water-ecoli-eau/alt/water-ecoli-eau-eng.pdf>

76. Raina, P. S. *et al.* The relationship between *E. coli* indicator bacteria in well-water and gastrointestinal illness in rural famililties. *Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique* **90,** 172–175 (2015).
77. Goss, M. J., Barry, D. A. J. & Rudolph, D. L. Contamination in Ontario farmstead domestic wells and its association with agriculture: 1. Results from drinking water wells. *J. Contam. Hydrol.* **32,** 267–293 (1998).
78. Rudolf, D. & Goss, M. *Ontario farm groundwater survey. Prepared for Agriculture Canada under the Federal-Provincial Land Management Assistance Program, 1992*. (Agriculture and Agri-Food Canada, 1993).
79. Sketchell, J. & Shaheen, N. Ground water quality in rural Saskatchewan–emerging issues for drinking water. in *Maintaining drinking water quality–lessons from the prairies and neyond. Proceedings of the 9th National Conference on Drinking Water'. Regina. Saskatchewan. Canada.(Ed. W. Robertson.) pp* 242–258 (2001).
80. Alberta Environment and Parks. *Taking Water Samples*. (2017).
81. Thomas, K. M. *et al.* A role of high impact weather events in waterborne disease outbreaks in Canada, 1975 - 2001. *Int. J. Environ. Health Res.* **16,** 167–180 (2006).
82. O'Dwyer, J., Downes, M. M. & Adley, C. C. The impact of meteorology on the occurrence of waterborne outbreaks of vero cytotoxin-producing *Escherichia coli* (VTEC): A logistic regression approach. *J. Water Health* **14,** 39–46 (2016).
83. Renter, D. G., Sargeant, J. M., Hygnstorm, S. E., Hoffman, J. D. & Gillespie, J. R. *Escherichia coli* O157:H7 in free-ranging deer in Nebraska. *J. Wildl. Dis.* **37,** 755–760 (2001).
84. Cooley, M. B. *et al.* Development of a robust method for isolation of Shiga toxin-positive *Escherichia coli* (STEC) from fecal, plant, soil and water samples from a leafy greens production region in California. *PLoS One* **8,** e65716 (2013).
85. Norman, K. N., Strockbine, N. A. & Bono, J. L. 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,** 6689–6703 (2012).

86. Persad, A. K. & Lejeune, J. T. Animal reservoirs of Shiga toxin-producing *Escherichia coli*. *Microbiol. Spectr.* **2,** EHEC-0027-2014 (2014).
87. Besser, T. E., Davis, M. A. & Walk, S. T. in *Population Genetics of Bacteria: a Tribute To Thomas S. Whittam* (eds. Walk, S. T. & Feng, P. C. H.) 303–324 (ASM Press, 2011).
88. Fremaux, B., Prigent-Combaret, C. & Vernozy-Rozand, C. Long-term survival of Shiga toxin-producing *Escherichia coli* in cattle effluents and environment: An updated review. *Veterinary Microbiology* **132,** 1–18 (2008).
89. Fremaux, B. *et al.* Survival and spread of Shiga toxin-producing *Escherichia coli* in alpine pasture grasslands. *J. Appl. Microbiol.* **108,** 1332–1343 (2010).
90. Fukushima, H., Hoshina, K. & Gomyoda, M. Long-term survival of Shiga toxin-producing *Escherichia coli* O26, O111, and O157 in bovine feces. *Appl. Environ. Microbiol.* **65,** 5177–5181 (1999).
91. Ogden, I. D. *et al.* Long-term survival of *Escherichia coli* O157 on pasture following an outbreak associated with sheep at a scout camp. *Lett. Appl. Microbiol.* **34,** 100–104 (2002).
92. Bolton, D. J. *et al.* Incidence and survival of non-O157 verocytotoxigenic *Escherichia coli* in soil. *J. Appl. Microbiol.* **111,** 484–490 (2011).
93. Money, P. *et al.* Cattle, weather and water: Mapping *Escherichia coli* O157:H7 infections in humans in England and Scotland. *Environ. Microbiol.* **12,** 2633–2644 (2010).
94. Jiang, X., Morgan, J. & Doyle, M. P. Fate of *Escherichia coli* O157: H7 in manure-amended soil. *Appl. Environ. Microbiol.* **68,** 2605–2609 (2002).
95. McGechan, M. & Lewis, D. Transport of particulate and colloid-sorbed contaminants through soil, Part 1: general principles. *Biosyst. Eng.* **83,** 255–273 (2002).
96. Tyrrel, S. F. & Quinton, J. N. Overland flow transport of pathogens from agricultural land receiving faecal wastes. *J. Appl. Microbiol.* **94,** 87S–93S

(2003).

97. Ogden, I. D., Fenlon, D. R., Vinten, A. J. A. & Lewis, D. The fate of *Escherichia coli* O157 in soil and its potential to contaminate drinking water. *Int. J. Food Microbiol.* **66,** 111–117 (2001).
98. Gagliardi, J. V. & Karns, J. S. Leaching of *Escherichia coli* O157:H7 in diverse soils under various agricultural management practices. *Appl. Environ. Microbiol.* **66,** 877–883 (2000).
99. Artz, R. R. E. & Killham, K. Survival of *Escherichia coli* O157 : H7 in private drinking water wells : influences of protozoan grazing and elevated copper concentrations. *FEMS Microbiol. Lett.* **216,** 117–122 (2002).
100. Watterworth, L., Rosa, B., Schraft, H., Topp, E. & Leung, K. T. Survival of various ERIC-genotypes of Shiga toxin-producing *Escherichia coli* in well water. *Water. Air. Soil Pollut.* **177,** 367–382 (2006).
101. Wang, G. D. & Doyle, M. P. Survival of enterohemorrhagic *Escherichia coli* O157 : H7 in water. *J. Food Prot.* **61,** 662–667 (1998).
102. Rice, E., Johnson, C., Wild, D. & Reasoner, D. Survival of *Escherichia coli* O157:H7 in drinking water associated with a waterborne disease outbreak of hemorrhagic colitis. *Lett. Appl. Microbiol.* **15,** 38–40 (1992).
103. Ritter, L. *et al.* Sources, pathways, and relative risks of contaminants in surface water and groundwater: a perspective prepared for the Walkerton inquiry. *J. Toxicol. Environ. Heal. Part A* **65,** 1–142 (2002).
104. Schaffter, N. & Parriaux, A. Pathogenic-bacterial water contamination in mountainous catchments. *Water Res.* **36,** 131–139 (2002).
105. Artz, R. R. E., Townend, J., Brown, K., Towers, W. & Killham, K. Soil macropores and compaction control the leaching potential of *Escherichia coli* O157:H7. *Environ. Microbiol.* **7,** 241–248 (2005).
106. Valeo, C., Checkley, S., He, J. & Neumann, N. Rainfall and microbial contamination in Alberta well water. *J. Environ. Eng. Sci.* **11,** 18–28 (2016).
107. Reynolds, K. A., Mena, K. D. & Gerba, C. P. Risk of waterborne illness via drinking water in the United States. *Rev. Environ. Contam. Toxicol.* **192,** 117–158 (2008).

108. Schets, F. M. *et al. Escherichia coli* O157:H7 in drinking water from private water supplies in the Netherlands. *Water Res.* **39,** 4485–4493 (2005).
109. Hynds, P. D., Misstear, B. D. & Gill, L. W. Development of a microbial contamination susceptibility model for private domestic groundwater sources. *Water Resour. Res.* **48,** W12504 (2012).
110. Jackson, S. G. *et al. Escherichia coli* O157:H7 diarrhoea associated with well water and infected cattle on an Ontario farm. *Epidemiol. Infect.* **120,** 17–20 (1998).
111. Lascowski, K. M. S. *et al.* Shiga toxin-producing *Escherichia coli* in drinking water supplies of north Paraná State, Brazil. *J. Appl. Microbiol.* **114,** 1230–1239 (2013).
112. Ram, S., Vajpayee, P. & Shanker, R. Contamination of potable water distribution systems by multiantimicrobial-resistant enterohemorrhagic *Escherichia coli*. *Environ. Health Perspect.* **116,** 448–452 (2008).
113. Ramteke, P. W. & Tewari, S. Serogroups of *Escherichia coli* from drinking water. *Environ. Monit. Assess.* **130,** 215–220 (2007).
114. Halabi, M. *et al.* Prevalence of Shiga toxin-, intimin- and haemolysin genes in *Escherichia coli* isolates from drinking water supplies in a rural area of Austria. *Int. J. Hyg. Environ. Health* **211,** 454–457 (2008).
115. Gu, G. *et al.* Occurrence and population density of *Escherichia coli* O157 contamination in irrigation ponds on produce farms in the Suwannee River Watershed. *Can. J. Microbiol.* **59,** 339–46 (2013).
116. Stanford, K., Johnson, R. P., Alexander, T. W., McAllister, T. A. & Reuter, T. 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 (2016).
117. Won, G., Gill, A. & LeJeune, J. T. Microbial quality and bacteria pathogens in private wells used for drinking water in northeastern Ohio. *J. Water Health* **11,** 555–562 (2013).
118. Nadya, S. *et al.* Phenotypic and genotypic characteristics of Shiga toxin-producing *Escherichia coli* isolated from surface waters and sediments in a

Canadian urban-agricultural landscape. *Front. Cell. Infect. Microbiol.* **6,** 36 (2016).

119. Gannon, V. P. J. *et al.* Bacterial pathogens in rural water supplies in Southern Alberta, Canada. *J. Toxicol. Environ. Health* **67,** 1643–1653 (2004).
120. Johnson, J. Y. M. *et al.* Prevalence of *Escherichia coli* O157: H7 and *Salmonella* spp. in surface waters of southern Alberta and its relation to manure sources. *Canadian Journal of Microbiology* **49,** 326–335 (2003).
121. Edge, T. A. *et al.* Investigation of an environmental benchmark for waterborne pathogens in agricultural watersheds in Canada. *J. Environ. Qual.* **41,** 21–30 (2012).
122. Walters, S. P., Gannon, V. P. J. & Field, K. G. Detection of *Bacteroidales* fecal indicators and the zoonotic pathogens *E. coli* 0157:H7, *Salmonella*, and *Campylobacter* in river water. *Environ. Sci. Technol.* **41,** 1856–1862 (2007).
123. Jokinen, C. C. *et al.* Spatial and temporal drivers of zoonotic pathogen contamination of an agricultural watershed. *J. Environ. Qual.* **41,** 242–252 (2012).
124. Jokinen, C. *et al.* Molecular subtypes of *Campylobacter* spp., *Salmonella enterica*, and *Escherichia coli* O157:H7 isolated from faecal and surface water samples in the Oldman River watershed, Alberta, Canada. *Water Res.* **45,** 1247–1257 (2011).
125. Wilkes, G. *et al.* Associations among pathogenic bacteria, parasites, and environmental and land use factors in multiple mixed-use watersheds. *Water Res.* **45,** 5807–5825 (2011).
126. Wilkes, G. *et al.* Seasonal relationships among indicator bacteria, pathogenic bacteria, *Cryptosporidium* oocysts, *Giardia* cysts, and hydrological indices for surface waters within an agricultural landscape. *Water Res.* **43,** 2209–2223 (2009).
127. Jokinen, C. C. *et al.* The occurrence and sources of *Campylobacter* spp., *Salmonella enterica* and *Escherichia coli* O157:H7 in the Salmon River, British Columbia, Canada. *J. Water Health* **8,** 374–386 (2010).
128. Lal, A., Hales, S., French, N. & Baker, M. G. Seasonality in human zoonotic

enteric diseases: A systematic review. *PLoS One* **7,** e31883 (2012).

129. Naumova, E. N. Mystery of seasonality: Getting the rhythm of nature. *J. Public Health Policy* **27,** 2–12 (2006).
130. Michel, P. *et al.* Temporal and geographical distributions of reported cases of *Escherichia coli* O157:H7 infection in Ontario. *Epidemiol. Infect.* **122,** 193–200 (1999).
131. Valcour, J. E. *et al.* A descriptive analysis of the spatio-temporal distribution of enteric diseases in New Brunswick, Canada. *BMC Public Health* **16,** 204 (2016).
132. Fleury, M., Charron, D. F., Holt, J. D., Allen, O. B. & Maarouf, A. R. A time series analysis of the relationship of ambient temperature and common bacterial enteric infections in two Canadian provinces. *Int. J. Biometeorol.* **50,** 385–391 (2006).
133. Varga, C. *et al.* Incidence, distribution, seasonality, and demographic risk factors of *Salmonella Enteritidis* human infections in Ontario, Canada, 2007–2009. *BMC Infect. Dis.* **13,** 212 (2013).
134. Philipsborn, R., Ahmed, S. M., Brosi, B. J. & Levy, K. Climatic drivers of diarrheagenic *Escherichia coli* incidence: a systematic review and meta-analysis. *J. Infect. Dis.* **214,** 6–15 (2016).
135. Heiman, K. E., Mody, R. K., Johnson, S. D., Griffin, P. M. & Hannah Gould, L. *Escherichia coli* O157 outbreaks in the United States, 2003–2012. *Emerg. Infect. Dis.* **21,** 1293–1301 (2015).
136. Griffin, P. M. & Tauxe, R. V. The epidemiology of infections caused by *Escherichia coli* O157: H7, other enterohemorrhagic *E. coli*, and the associated hemolytic uremic syndrome. *Epidemiol. Rev.* **13,** 60–98 (1991).
137. Pearl, D. L. *et al.* The use of outbreak information in the interpretation of clustering of reported cases of *Escherichia coli* O157 in space and time in Alberta, Canada, 2000-2002. *Epidemiol. Infect.* **134,** 699–711 (2006).
138. Waters, J. R., Sharp, J. C. & Dev, V. J. Infection caused by *Escherichia coli* O157:H7 in Alberta, Canada, and in Scotland: a five-year review, 1987-1991. *Clin Infect Dis* **19,** 834–843 (1994).

139. Hancock, D. D., Besser, T. E., Rice, D. H., Herriott, D. E. & Tarr, P. I. A longitudinal study of *Escherichia coli* O157 in fourteen cattle herds. *Epidemiol. Infect.* **118,** 193–195 (1997).
140. Barkocy-Gallagher, G. a *et al.* Seasonal prevalence of Shiga toxin-producing *Escherichia coli*, including O157:H7 and non-O157 serotypes, and *Salmonella* in commercial beef processing plants. *J. Food Prot.* **66,** 1978–1986 (2003).
141. Stanford, K. *et al.* Ecology of *Escherichia coli* O157: H7 in commercial dairies in southern Alberta. *J. Dairy Sci.* **88,** 4441–4451 (2005).
142. Smith, D. R. *et al.* Use of rope devices to describe and explain the feedlot ecology of *Escherichia coli* O157:H7 by time and place. *Foodborne Pathog. Dis.* **2,** 50–60 (2005).
143. Dewsbury, D. M. A. *et al.* Summer and winter prevalence of Shiga toxin–producing *Escherichia coli* (STEC) O26, O45, O103, O111, O121, O145, and O157 in feces of feedlot cattle. *Foodborne Pathog. Dis.* **12,** 726–732 (2015).
144. Ogden, I. D., MacRae, M. & Strachan, N. J. C. Is the prevalence and shedding concentrations of *E. coli* O157 in beef cattle in Scotland seasonal? *FEMS Microbiol. Lett.* **233,** 297–300 (2004).
145. Edrington, T. S. *et al.* Seasonal shedding of *Escherichia coli* O157: H7 in ruminants: a new hypothesis. **3,** 413–421 (2006).
146. Pearce, M. C. *et al.* Prevalence and virulence factors of *Escherichia coli* serogroups O26, O103, O111 and O145 shed by cattle in Scotland. *Appl. Environ. Microbiol.* **72,** 653–659 (2006).
147. Kudva, I. T., Hatfield, P. G. & Hovde, C. J. Characterization of *Escherichia coli* O157:H7 and other Shiga toxin-producing *E. coli* serotypes isolated from sheep. *J. Clin. Microbiol.* **35,** 892–899 (1997).
148. Milnes, A. S. *et al.* Intestinal carriage of Verocytotoxigenic Escherichia coli O157, Salmonella, thermophilic Campylobacter and Yersinia enterocolitica, in cattle, sheep and pigs at slaughter in Great Britain during 2003. *Epidemiol. Infect.* **136,** 739–751 (2008).
149. Franco, A. *et al.* Prevalence and concentration of Verotoxigenic Escherichia coli O157:H7 in adult sheep at slaughter from Italy. *Zoonoses Public Health* **56,**

215–220 (2009).

150. Fraser, E. M., MacRae, M., Ogden, I. D., Forbes, K. J. & Strachan, N. J. C. Effects of seasonality and a diet of brassicas on the shedding of Escherichia coli O157 in sheep. *Foodborne Pathog. Dis.* **10,** 649–654 (2013).
151. Chapman, P. A., Siddons, C. A., Cerdan Malo, A. T. & Harkin, M. A. A 1-year study of *Escherichia coli* O157 in cattle, sheep, pigs and poultry. *Epidemiol. Infect.* **119,** 245–250 (1997).
152. Sánchez, S. *et al.* Variation in the prevalence of non-O157 Shiga toxin-producing *Escherichia coli* in four sheep flocks during a 12-month longitudinal study. *Small Rumin. Res.* **93,** 144–148 (2010).
153. Orden, J. A. *et al.* A longitudinal study of verotoxin-producing *Escherichia coli* in two dairy goat herds. *Vet. Microbiol.* **132,** 428–434 (2008).
154. Dunn, J. R., Keen, J. E., Moreland, D. & Thompson, R. A. Prevalence of *Escherichia coli* O157:H7 in White-tailed Deer from Louisiana. *J. Wildl. Dis.* **40,** 361–365 (2004).
155. Guzman Herrador, B. R. *et al.* Analytical studies assessing the association between extreme precipitation or temperature and drinking water-related waterborne infections: a review. *Environ. Heal.* **14,** 29 (2015).
156. Kistemann, T. *et al.* Microbial load of drinking water reservoir tributaries during extreme rainfall and runoff. *Appl. Environ. Microbiol.* **68,** 2188–2197 (2002).
157. Cann, K. F., Thomas, D. R., Salmon, R. L., Wyn-Jones, A. P. & Kay, D. Extreme water-related weather events and waterborne disease. *Epidemiol. Infect.* **141,** 671–686 (2013).
158. Curriero, F. C., Patz, J. A., Rose, J. B. & Lele, S. The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948–1994. *Am. J. Public Health* **91,** 1194–1199 (2001).
159. Bifolchi, N. *et al.* Weather and livestock risk factors for *Escherichia coli* O157 human infection in Alberta, Canada. *Epidemiol. Infect.* **142,** 2302–2313 (2014).
160. Galway, L. P., Allen, D. M., Parkes, M. W. & Takaro, T. K. Seasonal variation of

acute gastro-intestinal illness by hydroclimatic regime and drinking water source: A retrospective population-based study. *J. Water Health* **12,** 122–135 (2014).

161. Alberta Environment. *Alberta Environment's Drinking Water Program: A 'Source to Tap, Multi-Barrier' Approach*. (2009).
162. Barrell, R. A., Hunter, P. R. & Nichols, G. Microbiological standards for water and their relationship to health risk. *Commun. Dis. public Heal.* **3,** 8–13 (2000).
163. Ahmed, W., Gyawali, P. & Toze, S. Quantitative PCR measurements of *Escherichia coli* including Shiga toxin-producing *E. coli* (STEC) in animal feces and environmental waters. *Environ. Sci. Technol.* **49,** 3084–3090 (2015).
164. Clark, S. T. *et al.* Evaluation of low-copy genetic targets for waterborne bacterial pathogen detection via qPCR. *Water Res.* **45,** 3378–3388 (2011).
165. Hacker, J., Blum-Oehler, G., Mühldorfer, I. & Tschäpe, H. Pathogenicity islands of virulent bacteria: Structure, function and impact on microbial evolution. *Mol. Microbiol.* **23,** 1089–1097 (1997).
166. Chui, L. *et al.* Comparison of Shiga toxin-producing *Escherichia coli* detection methods using clinical stool samples. *J. Mol. Diagn.* **12,** 469–475 (2010).
167. StataCorp. in *Stata 13 Base Reference Manual* (Stata Press, 2013). at <http://www.stata.com/manuals13/rnbreg.pdf>
168. Osgood, D. W. Poisson-based regression analysis of aggregate crime rates. *J. Quant. Criminol.* **16,** 21–43 (2000).
169. Morris, D., Kavanagh, S., Carney, K., MacDomhnaill, B. & Cormican, M. CapE (capture, amplify, extract): A rapid method for detection of low level contamination of water with Verocytotoxigenic *Escherichia coli* (VTEC). *Sci. Total Environ.* **563–564,** 267–272 (2016).
170. Zelyas, N. *et al.* Assessment of commercial chromogenic solid media for the detection of non-O157 Shiga toxin-producing *Escherichia coli* (STEC). *Diagn. Microbiol. Infect. Dis.* **85,** 302–308 (2016).
171. Gouali, M., Ruckly, C., Carle, I., Lejay-Collin, M. & Weill, F. X. Evaluation of CHROMagar STEC and STEC O104 chromogenic agar media for detection of Shiga toxin-producing *Escherichia coli* in stool specimens. *J. Clin. Microbiol.* **51,**

894–900 (2013).

172. Gill, A., Huszczynski, G., Gauthier, M. & Blais, B. Evaluation of eight agar media for the isolation of Shiga toxin-Producing *Escherichia coli*. *J. Microbiol. Methods* **96,** 6–11 (2014).
173. Wylie, J. L. *et al.* Evaluation of a new chromogenic agar medium for detection of Shiga toxin-producing *Escherichia coli* (STEC) and relative prevalences of O157 and non-O157 STEC in Manitoba, Canada. *J. Clin. Microbiol.* **51,** 466–471 (2013).
174. Kase, J. a, Maounounen-Laasri, A., Son, I., Lin, A. & Hammack, T. S. Comparison of eight different agars for the recovery of clinically relevant non-O157 Shiga toxin-producing Escherichia coli from baby spinach, cilantro, alfalfa sprouts and raw milk. *Food Microbiol.* **46,** 280–287 (2015).
175. Verhaegen, B., De Reu, K., Heyndrickx, M. & De Zutter, L. Comparison of six chromogenic agar media for the isolation of a broad variety of non-O157 Shigatoxin-producing *Escherichia coli* (STEC) serogroups. *Int. J. Environ. Res. Public Health* **12,** 6965–6978 (2015).
176. *CHROMagar STEC Instructions For Use*. (CHROMagar, 2017).
177. Anderson, K. M., Abbott, J., Zhao, S. & Liu, E. Molecular subtyping of Shiga toxin-producing *Escherichia coli* using a commercial repetitive sequence-based PCR assay. *J. Food Prot.* **78,** 902–911 (2015).
178. Mohapatra, B. R. & Mazumder, A. Comparative efficacy of five different rep-PCR methods to discriminate *Escherichia coli* populations in aquatic environments. *Water Sci. Technol.* **58,** 537–547 (2008).
179. Versalovic, J., Schneider, M., de Bruijn, F. J. & Lupski, J. R. Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction. *Methods Mol. Cell. Biol.* **5,** 25–40 (1994).
180. Pincus, D. H. in *Encyclopedia of Rapid Microbiological Methods* (ed. Miller, M.) 1–32 (2006).
181. Ewing, E. H. in *Identification of Enterobacteriaceae* (eds. Ewing, W. H. & Edwards, P. R.) 96–134 (Elsevier Science, 1986).
182. Edwards, J. H. Seasonal incidence of congenital disease in Birmingham. *Ann.*

Hum. Genet. **25,** 89–93 (1961).

183. Leger, A. S. Comparison of two tests for seasonality in epidemiological data. *J. R. Stat. Soc. Ser. C Appl. Stat.* **25,** 280–286 (1976).
184. Abramson, J. H. WinPepi. (2016). at <http://www.brixtonhealth.com/pepi4windows.html>
185. Roger, J. H. A significance test for cyclic trends in incidence data. *Biometrika* **64,** 152–155 (1977).
186. Abramson, J. H. in *WINPEPI Manual* 1–140 (2016).
187. Costa, M. A. & Kulldorff, M. in *Scan Statistics: Methods and Applications* (eds. Glaz, J., Pozdnyakov, V. & Wallenstein, S.) 129–152 (Birkhäuser Boston, a part of Springer Science+Business Media, 2009). doi:10.1007/978-0-8176-4749-0
188. Kulldorff, B. M. *SaTScan User Guide for version 9.4*. (2015).
189. Kulldorff, M. & Information Management Services Inc. SaTScan™: Software for the spatial and space-time scan statistics. (2015). at <http://www.satscan.org/>
190. Kulldorff, M. A spatial scan statistic. *Commun. Stat. - Theory Methods* **26,** 1481–1496 (1997).
191. Invik, J. Total coliform and *Escherichia coli* positivity in rural well water in Alberta, Canada: spatiotemporal analysis and risk factor assessment. (Univeristy of Calgary, 2015).
192. Alberta Environment and Parks. Alberta Township Survey System. (2010). at <http://aep.alberta.ca/recreation-public-use/recreation-on-agricultural-public-land/alberta-township-survey-system.aspx>
193. Kulldorff, M., Heffernan, R., Hartman, J., Assunção, R. & Mostashari, F. A space-time permutation scan statistic for disease outbreak detection. *PLoS Med.* **2,** e59 (2005).
194. AltaLIS Ltd. AltaLIS Virtual Order Desk. (2017). at <http://www.altalis.com/altalisnet/main.asp>
195. Denno, D. M. *et al.* Tri-county comprehensive assessment of risk factors for sporadic reportable bacterial enteric infection in children. *J. Infect. Dis.* **199,** 467–476 (2009).

196. Jaros, P. *et al.* 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 (2013).
197. Naumova, E. N. *et al.* Use of passive surveillance data to study temporal and spatial variation in the incidence of giardiasis and cryptosporidiosis. *Public Health Rep.* **115,** 436–447 (2000).
198. Environment and Climate Change Canada. Top ten weather stories for 2005: story one. at <https://www.ec.gc.ca/meteo-weather/default.asp?lang=En&n=A4DD5AB5-1>
199. Environment and Climate Change Canada. Canada's top ten weather stories of 2013. at <https://www.ec.gc.ca/meteo-weather/default.asp?lang=En&n=5BA5EAFC-1&offset=2&toc=show>
200. Galway, L. P., Allen, D. M., Parkes, M. W., Li, L. & Takaro, T. K. Hydroclimatic variables and acute gastro-intestinal illness in British Columbia, Canada: A time series analysis. *Water Resour. Res.* **51,** 885–895 (2015).
201. Patz, J. a *et al.* The potential health impacts of climate variability and change for the United States: executive summary of the report of the health sector of the U.S. National Assessment. *Environ. Health Perspect.* **108,** 367–376 (2000).
202. Charron, D. F. *et al.* Vulnerability of waterborne diseases to climate change in Canada: a review. *J. Toxicol. Environ. Heal. Part A* **67,** 1667–1677 (2004).
203. Lake, I. R., Bentham, G., Kovats, R. S. & Nichols, G. L. Effects of weather and river flow on cryptosporidiosis. *J. Water Health* **3,** 469–474 (2005).
204. Britton, E., Hales, S., Venugopal, K. & Baker, M. G. The impact of climate variability and change on cryptosporidiosis and giardiasis rates in New Zealand. *J. Water Health* **8,** 561–571 (2010).
205. Eisenberg, J. N. S. *et al.* Environmental determinants of infectious disease: A framework for tracking causal links and guiding public health research. *Environ. Health Perspect.* **115,** 1216–1223 (2007).
206. Craighead, L., Gilbert, W., Subasinghe, D. & Häsler, B. Reconciling surveillance systems with limited resources: An evaluation of passive surveillance for rabies in an endemic setting. *Prev. Vet. Med.* **121,** 206–214 (2015).

207. Osgood, D. W. & Chambers, J. M. Social disorganization outside the metropolis: an analysis of rural youth violence. *Criminology* **38,** 81–116 (2000).
208. Parsons, B. D., Zelyas, N., Berenger, B. M. & Chui, L. Detection, characterization, and typing of Shiga toxin-producing *Escherichia coli*. *Front. Microbiol.* **7,** 478 (2016).
209. McKay, A. M. Viable but non-culturable forms of potentially pathogenic bacteria in water. *Lett. Appl. Microbiol.* **14,** 129–135 (1992).
210. Meyer, K. Antimicrobial resistant *Escherichia coli* in Alberta's rural well water. (University of Calgary, 2017).
211. *Standard methods for the examination of water and wastewater*. (American Public Health Association, American Water Works Association, Water Environment Federation, 2012).
212. Grande, L. *et al.* Whole-genome characterization and strain comparison of VT2f-producing *Escherichia coli* causing hemolytic uremic syndrome. *Emerg. Infect. Dis.* **22,** 2078–2086 (2016).
213. Feng, P. C. H., Jinneman, K., Scheutz, F. & Monday, S. R. Specificity of PCR and serological assays in the detection of *Escherichia coli* Shiga toxin subtypes. *Appl. Environ. Microbiol.* **77,** 6699–6702 (2011).
214. Prager, R., Fruth, A., Siewert, U., Strutz, U. & Tschäpe, H. *Escherichia coli* encoding Shiga toxin 2f as an emerging human pathogen. *Int. J. Med. Microbiol.* **299,** 343–353 (2009).
215. van Duynhoven, Y. T. H. P. *et al.* Prevalence, characterisation and clinical profiles of Shiga toxin-producing *Escherichia coli* in The Netherlands. *Clin. Microbiol. Infect.* **14,** 437–445 (2008).
216. Friesema, I. *et al.* Emergence of *Escherichia coli* encoding Shiga toxin 2f in human Shiga toxin-producing *E. coli* (STEC) infections in the Netherlands, January 2008 to December 2011. *Eurosurveillance* **19,** 26–32 (2014).
217. Friesema, I. H. M. *et al.* Hemolytic uremic syndrome associated with *Escherichia coli* O8:H19 and Shiga toxin 2f gene. *Emerg. Infect. Dis.* **21,** 168–169 (2015).
218. Schmidt, H. *et al.* A new Shiga toxin 2 variant (Stx2f) from *Escherichia coli*

isolated from pigeons. *Appl. Environ. Microbiol.* **66,** 1205–1208 (2000).

219. Tzschoppe, M., Martin, A. & Beutin, L. A rapid procedure for the detection and isolation of enterohaemorrhagic *Escherichia coli* (EHEC) serogroup O26, O103, O111, O118, O121, O145 and O157 strains and the aggregative EHEC O104:H4 strain from ready-to-eat vegetables. *Int. J. Food Microbiol.* **152,** 19–30 (2012).
220. Gill, A., Martinez-Perez, A., McIlwham, S. & Blais, B. Development of a method for the detection of verotoxin-producing *Escherichia coli* in food. *J. Food Prot.* **75,** 827–37 (2012).
221. Catford, A. *et al.* Risk profile on non-O157 verotoxin-producing *Escherichia coli* in produce, beef, milk and dairy products in Canada. *Int. Food Risk Anal. J.* **4,** 21 (2014).
222. Durso, L. M. Primary isolation of Shiga toxigenic *Escherichia coli* from environmental sources. *J. Environ. Qual.* **42,** 1295–307 (2013).
223. McCallum, C., McGregor, A. & Vanniasinkam, T. Prevalence of Shiga toxin-producing *Escherichia coli* (STEC) in Tasmania, Australia. *Pathology* **45,** 681–688 (2013).
224. Hirvonen, J. J., Siitonen, A. & Kaukoranta, S. S. Usability and performance of CHROMagar STEC medium in detection of Shiga toxin-producing *Escherichia coli* strains. *J. Clin. Microbiol.* **50,** 3586–3590 (2012).
225. Cernela, N., Zweifel, C. & Stephan, R. Evaluation of the CHROMagar (TM) STEC medium for the detection of Shiga toxin-producing *Escherichia coli*. *J. Food Saf. Food Qual. Fur Leb.* **66,** 28–32 (2015).
226. Kerangart, S. *et al.* Variable tellurite resistance profiles of clinically-relevant Shiga toxin-producing *Escherichia coli* (STEC) influence their recovery from foodstuffs. *Food Microbiol.* **59,** 32–42 (2016).
227. Karch, H., Meyer, T., Rüssmann, H. & Heesemann, J. Frequent loss of Shiga-like toxin genes in clinical isolates of *Escherichia coli* upon subcultivation. *Infect. Immun.* **60,** 3464–3467 (1992).
228. Tozzoli, R. *et al.* Shiga toxin-converting phages and the emergence of new pathogenic *Escherichia coli*: a world in motion. *Front. Cell. Infect. Microbiol.* **4,**

80 (2014).

229. Bielaszewska, M. *et al.* Shiga toxin gene loss and transfer in vitro and in vivo during enterohemorrhagic *Escherichia coli* O26 infection in humans. *Appl. Environ. Microbiol.* **73,** 3144–3150 (2007).
230. Friedrich, A. W. *et al.* Prevalence, virulence profiles, and clinical significance of Shiga toxin–negative variants of enterohemorrhagic *Escherichia coli* O157 infection in humans. *Clin. Infect. Dis.* **45,** 39–45 (2007).
231. Khalil, R. K. S., Skinner, C., Patfield, S. & He, X. Phage-mediated Shiga toxin (Stx) horizontal gene transfer and expression in non-Shiga toxigenic *Enterobacter* and *Escherichia coli* strains. *Pathog. Dis.* **74,** ftw037 (2015).
232. Paton, J. C. & Paton, A. W. Instability of a Shiga toxin type 2 gene in *Enterobacter cloacae*. *J. Clin. Microbiol.* **35,** 1917 (1997).
233. Alperi, A. & Figueras, M. J. Human isolates of *Aeromonas* possess Shiga toxin genes (stx1 and stx2) highly similar to the most virulent gene variants of *Escherichia coli*. *Clin. Microbiol. Infect.* **16,** 1563–1567 (2010).
234. Grotiuz, G., Sirok, A., Gadea, P., Varela, G. & Schelotto, F. Shiga toxin 2-producing *Acinetobacter haemolyticus* associated with a case of bloody diarrhea. *J. Clin. Microbiol.* **44,** 3838–3841 (2006).
235. Schmidt, H., Montag, M., Bockemuhl, J., Heesemann, J. & Karch, H. Shiga-like toxin II-related cytotoxins in *Citrobacter freundii* strains from humans and beef samples. *Infect. Immun.* **61,** 534–543 (1993).
236. Rangel, J. M., Sparling, P. H., Crowe, C., Griffin, P. M. & Swerdlow, D. L. Epidemiology of *Escherichia coli* O157:H7 outbreaks, United States, 1982-2002. *Emerg. Infect. Dis.* **11,** 603–609 (2005).
237. Caprioli, A., Morabito, S. & Scavia, G. in *Encyclopedia of Food Safety* (eds. Motarjemi, Y., Moy, G. & Todd, E. C. D.) **1,** 417–423 (Academic Press, 2014).
238. McCall, B. J. *et al.* An outbreak of Shiga toxin-producing *Escherichia coli* infection associated with a school camp. *Commun. Dis. Intell. Q. Rep.* **34,** 54–56 (2010).
239. Hrudey, S. E. & Hrudey, E. J. Published case studies of waterborne disease outbreaks- evidence of a recurrent threat. *Water Environ. Res.* **79,** 233–45

(2007).

240. Käppeli, U., Hächler, H., Giezendanner, N., Beutin, L. & Stephan, R. Human infections with non-O157 Shiga toxin-producing *Escherichia coli*, Switzerland, 2000-2009. *Emerg. Infect. Dis.* **17,** 180–185 (2011).
241. Woodward, D. L., Clark, C. G., Caldeira, R. A., Ahmed, R. & Rodgers, F. G. Verotoxigenic *Escherichia coli* (VTEC): A major public health threat in Canada. *Can. J. Infect. Dis.* **13,** 321–330 (2002).
242. Beutin, L. & Fach, P. Detection of Shiga toxin-producing *Escherichia coli* from nonhuman sources and strain typing. *Microbiol. Spectr.* **2,** EHEC-0001-2013 (2014).
243. Centers for Disease Control and Prevention. *National Enteric Disease Surveillance: Shiga toxin-producing Escherichia coli (STEC) Annual Report, 2014*. (2017).
244. Diarra, M. S. *et al.* Genotype, serotype, and antibiotic resistance of sorbitol-negative *Escherichia coli* isolates from feedlot cattle. *J. Food Prot.* **72,** 28–36 (2009).
245. Girardeau, J. P. *et al.* Association of virulence genotype with phylogenetic background in comparison to different seropathotypes of Shiga toxin-producing *Escherichia coli* isolates. *J. Clin. Microbiol.* **43,** 6098–6107 (2005).
246. Tanaro, J. D. *et al.* Non-O157: H7 Shiga toxin–producing *Escherichia coli* in bovine rectums and surface water streams on a beef cattle farm in Argentina. *Foodborne Pathog. Dis.* **9,** 878–884 (2012).
247. Nyholm, O. *et al.* Hybrids of Shigatoxigenic and enterotoxigenic *Escherichia coli* (STEC/ETEC) among human and animal isolates in Finland. *Zoonoses Public Health* **62,** 518–524 (2015).
248. Gilmour, M. W. *et al.* Isolation and detection of Shiga toxin-producing *Escherichia coli* in clinical stool samples using conventional and molecular methods. *J. Med. Microbiol.* **58,** 905–911 (2009).
249. Conrad, C. C., Stanford, K., McAllister, T. A., Thomas, J. & Reuter, T. 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 (2014).

250. Tanwar, J., Das, S., Fatima, Z. & Hameed, S. Multidrug resistance: An emerging crisis. *Interdiscip. Perspect. Infect. Dis.* **2014,** (2014).
251. Hoshina, K. *et al.* Enterohemorrhagic *Escherichia coli* O26 outbreak caused by contaminated natural water supplied by facility owned by local community. *Jpn. J. Infect. Dis.* **54,** 247–248 (2002).
252. Krolik, J. *et al.* A spatial analysis of private well water *Escherichia coli* contamination in Southern Ontario. *Geospat. Health* **8,** 65–75 (2013).
253. Musa, G. J. *et al.* Use of GIS mapping as a public health tool—from cholera to cancer. *Heal. Serv. Insights* **6,** 111–116 (2013).
254. Rau, R. in *Seasonality in Human Mortality: A Demographic Approach* 39–81 (Springer Science & Business Media, 2006).
255. Wallenstein, S., Weinberg, C. R. & Gould, M. Testing for a pulse in seasonal event data. *Biometrics* **45,** 817–830 (1989).
256. García, A., Fox, J. G. & Besser, T. E. Zoonotic Enterohemorrhagic *Escherichia coli*: A One Health Perspective. *Inst. Lab. Anim. Res. J.* **51,** 221–232 (2010).
257. So, H. C. *et al.* Spatio-temporal scan statistics for the detection of outbreaks involving common molecular subtypes: Using human cases of *Escherichia coli* O157: H7 provincial PFGE pattern 8 (National Designation Ecxai.0001) in Alberta as an example. *Zoonoses Public Health* **60,** 341–348 (2013).
258. Health Canada. *Microbial Guidelines for Ready-to-Eat Foods- A Guide for the Conveyence Industry and Environment Health Officers (EHO)*. (2010). at <http://publications.gc.ca/collections/collection_2014/sc-hc/H164-167-2013-eng.pdf>
259. Williams, M. S. *et al.* Determining relationships between the seasonal occurrence of *Escherichia coli* O157: H7 in live cattle, ground beef, and humans. *Foodborne Pathog. Dis.* **7,** 1247–1254 (2010).
260. Kassenborg, H. D. *et al.* Farm visits and undercooked hamburgers as major risk factors for sporadic *Escherichia coli* O157: H7 infection: data from a case-control study in 5 FoodNet sites. *Clin. Infect. Dis.* **38,** S271–S278 (2004).
261. Shipman, A. R., Jones, S. E., Smith, G., Stewart, B. & Mccarthy, N. A case of

verocytotoxin-producing *Escherichia coli* O157 from a private barbecue in South East England. *Eurosurveillance* **14,** pii=19273 (2009).

262. MacDonald, C. *et al.* Outbreak of *Escherichia coli* O157:H7 leading to the recall of retail ground beef - Winnipeg, Manitoba, May 1999. *Canada Commun. Dis. Rep.* **26,** 109–111 (2000).
263. Kintz, E., Brainard, J., Hooper, L. & Hunter, P. Transmission pathways for sporadic Shiga-toxin producing *E. coli* infections: A systematic review and meta-analysis. *Int. J. Hyg. Environ. Health* **220,** 57–67 (2017).
264. Strachan, N. J. C., Dunn, G. M., Locking, M. E., Reid, T. M. S. & Ogden, I. D. *Escherichia coli* O157: Burger bug or environmental pathogen? *Int. J. Food Microbiol.* **112,** 129–137 (2006).
265. Voetsch, A. C. *et al.* Risk factors for sporadic Shiga toxin-producing *Escherichia coli* O157 infections in FoodNet sites, 1999–2000. *Epidemiol. Infect.* **135,** 993–1000 (2007).
266. Byrne, L. *et al.* Epidemiology and microbiology of Shiga toxin-producing *Escherichia coli* other than serogroup O157 in England, 2009-2013. *J. Med. Microbiol.* **63,** 1181–1188 (2014).
267. Friesema, I. H. M., Schotsborg, M., Heck, M. E. O. C. & Van Pelt, W. Risk factors for sporadic Shiga toxin-producing *Escherichia coli* O157 and non-O157 illness in The Netherlands, 2008-2012, using periodically surveyed controls. *Epidemiol. Infect.* **143,** 1360–1367 (2015).
268. Rivas, M. *et al.* Risk factors for sporadic Shiga toxin-producing *Escherichia coli* infections in children, Argentina. *Emerg. Infect. Dis.* **14,** 763–771 (2008).
269. Wang, X. *et al.* Comparison of clinical and epidemiological features of Shiga toxin-producing *Escherichia coli* O157 and non-O157 infections in British Columbia, 2009 to 2011. *Can. J. Infect. Dis. Med. Microbiol.* **24,** e102-106 (2013).
270. Piérard, D. *et al.* A case-control study of sporadic infection with O157 and non-O157 verocytotoxin-producing *Escherichia coli*. *Epidemiol. Infect.* **122,** 359–365 (1999).
271. Wilson, J. B. *et al.* Vero cytotoxigenic *Escherichia coli* infection in dairy farm

families. *J. Infect. Dis.* **174,** 1021–1027 (1996).

272. Teschke, K. *et al.* Water and sewage systems, socio-demographics, and duration of residence associated with endemic intestinal infectious diseases: A cohort study. *BMC Public Health* **10,** 767 (2010).
273. Strauss, B., King, W., Ley, A. & Hoey, J. R. A prospective study of rural drinking water quality and acute gastrointestinal illness. *BMC Public Health* **1,** 8 (2001).
274. Stephan, R., Ragettli, S. & Untermann, F. Prevalence and characteristics of verotoxin-producing *Escherichia coli* (VTEC) in stool samples from asymptomatic human carriers working in the meat processing industry in Switzerland. *J. Appl. Microbiol.* **88,** 335–341 (2000).
275. Takeda, T. *et al.* High prevalence of serum antibodies to Verotoxins 1 and 2 among healthy adults in Japan. *Journal of Infection* **27,** 211–213 (1993).
276. Hong, S. *et al.* Anti-Shiga toxin immunoglobulin G antibodies in healthy South Korean slaughterhouse workers. *Scand. J. Infect. Dis.* **44,** 168–173 (2012).
277. Belongia, E. A. *et al.* Diarrhea incidence and farm-related risk factors for *Escherichia coli* O157:H7 and *Campylobacter jejuni* antibodies among rural children. *J. Infect. Dis.* **187,** 1460–1468 (2003).
278. Karmali, M. A. *et al.* Age-specific frequencies of antibodies to *Escherichia coli* verocytotoxins (Shiga toxins) 1 and 2 among urban and rural populations in southern Ontario. *J. Infect. Dis.* **188,** 1724–1729 (2003).
279. Reymond, D. *et al.* Neutralizing antibodies to *Escherichia coli* vero cytotoxin 1 and antibodies to O157 lipopolysaccharide in healthy farm family members and urban residents. *J. Clin. Microbiol.* **34,** 2053–2057 (1996).
280. Karmali, M. A. *et al.* Enzyme-linked immunosorbent assay for detection of immunoglobulin G antibodies to *Escherichia coli* vero cytotoxin 1. *J. Clin. Microbiol.* **32,** 1457–1463 (1994).
281. Eklund, M., Leino, K. & Siitonen, A. Clinical *Escherichia coli* strains carrying *stx* genes: *stx* Variants and *stx*-positive virulence profiles. *J. Clin. Microbiol.* **40,** 4585–4593 (2002).
282. Baker-Ismail, S. Survival and overland transport of fecal coliform under Canadian prairie conditions. (University of Regina, 2016).

283. Eaton, B. & Moore, R. D. in *Compendium of Forest Hydrology and Geomorphology in British Columbia, Volume 1 of 2* (ed. R.G. Pike, T.E. Redding, R.D. Moore, R.D. Winkler, and K. D. B.) **2,** 446 (2010).
284. Adhikari, H., Barnes, D. L., Schiewer, S. & White, D. M. Total coliform survival characteristics in frozen soils. *J. Environ. Eng.* **133,** 1098–1105 (2007).
285. Texier, S. *et al.* Persistence of culturable *Escherichia coli* fecal contaminants in dairy alpine grassland soils. *J. Environ. Qual.* **37,** 2299–2310 (2008).
286. Naumova, E. N. *et al.* Seasonality in six enterically transmitted diseases and ambient temperature. *Epidemiol. Infect.* **135,** 281–292 (2007).
287. Aramini, J. *et al.* Drinking water quality and health care utilization for gastrointestinal illness in Greater Vancouver. *Canada Commun. Dis. Rep.* **26,** 211–214 (2000).
288. Atalla, H. N., Johnson, R. P., McEwen, S. A., Usborne, W. R. & Gyles, C. L. Use of a Shiga toxin (Stx)-enzyme-linked immunosorbent assay and immunoblot for detection and isolation of Stx-producing *Escherichia coli* from naturally contaminated beef. *J. Food Prot.* **63,** 1167–1172 (2000).

Appendix A

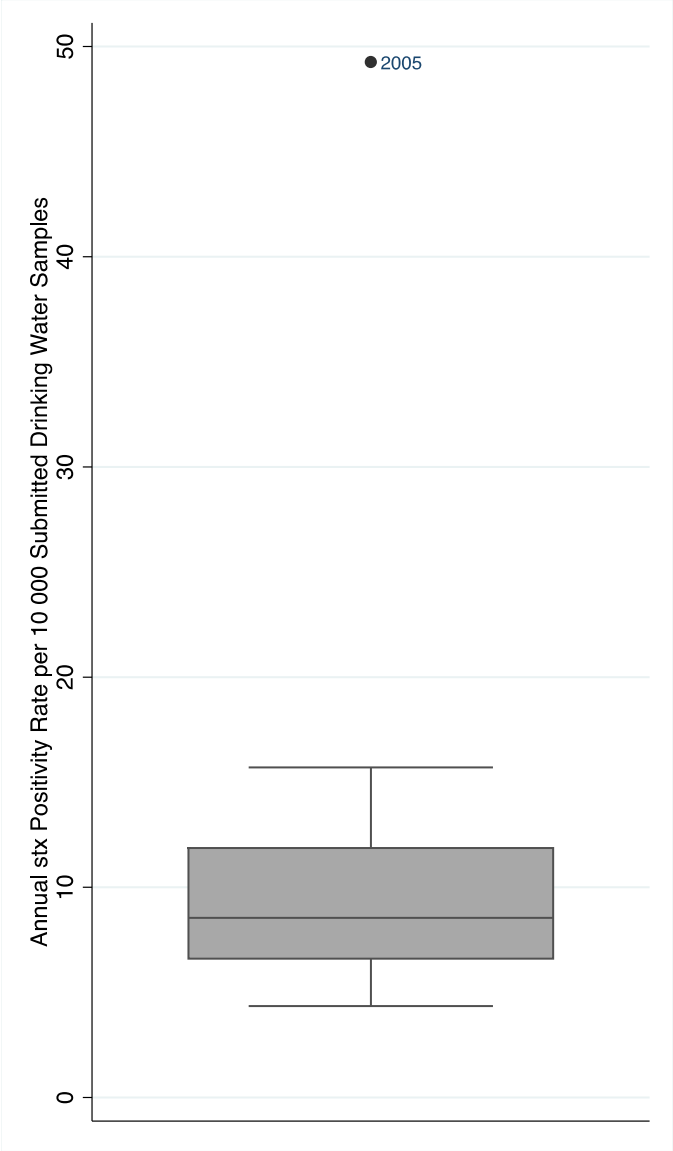

Figure A 1 – A box plot of annual *stx* positivity rates per 10 000 submitted drinking water samples. Outlying values (those outside of the upper or lower fences of the box plot) are labeled and represented by a black dot.

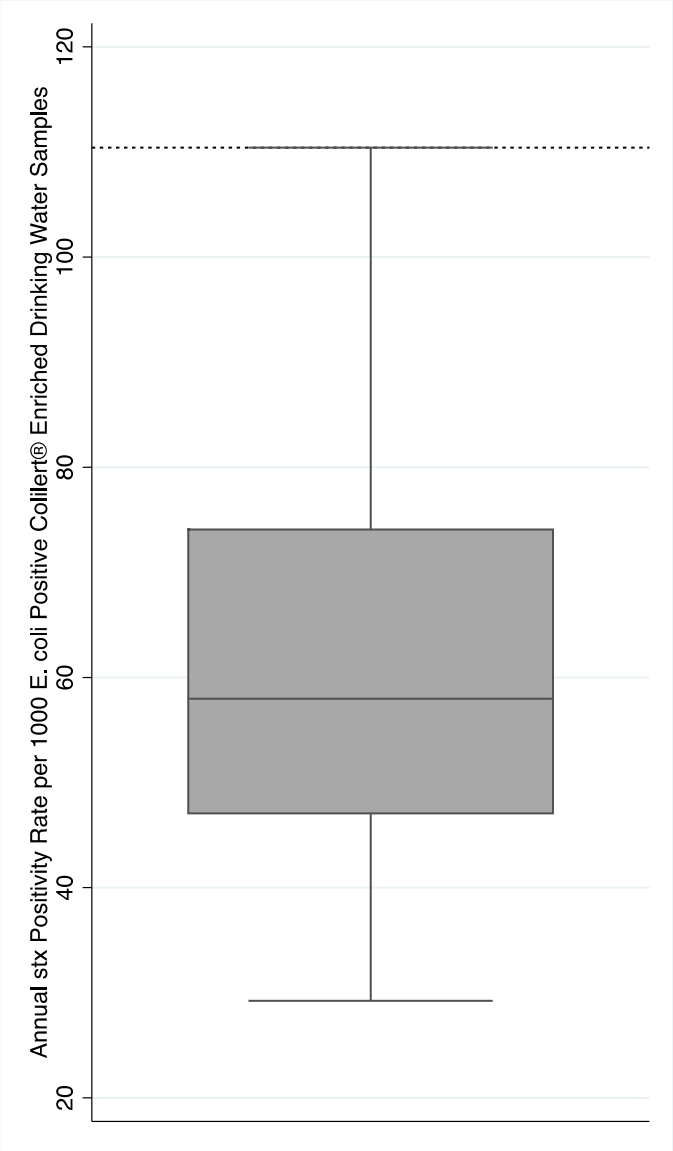

Figure A 2 – A box plot of annual *stx* positivity rates per 1000 *E. coli* positive Colilert® enriched drinking water samples. The dashed reference line represents the value of the 2005 annual *stx* positivity rate, which fell just within the upper fence of the box plot.

Appendix B

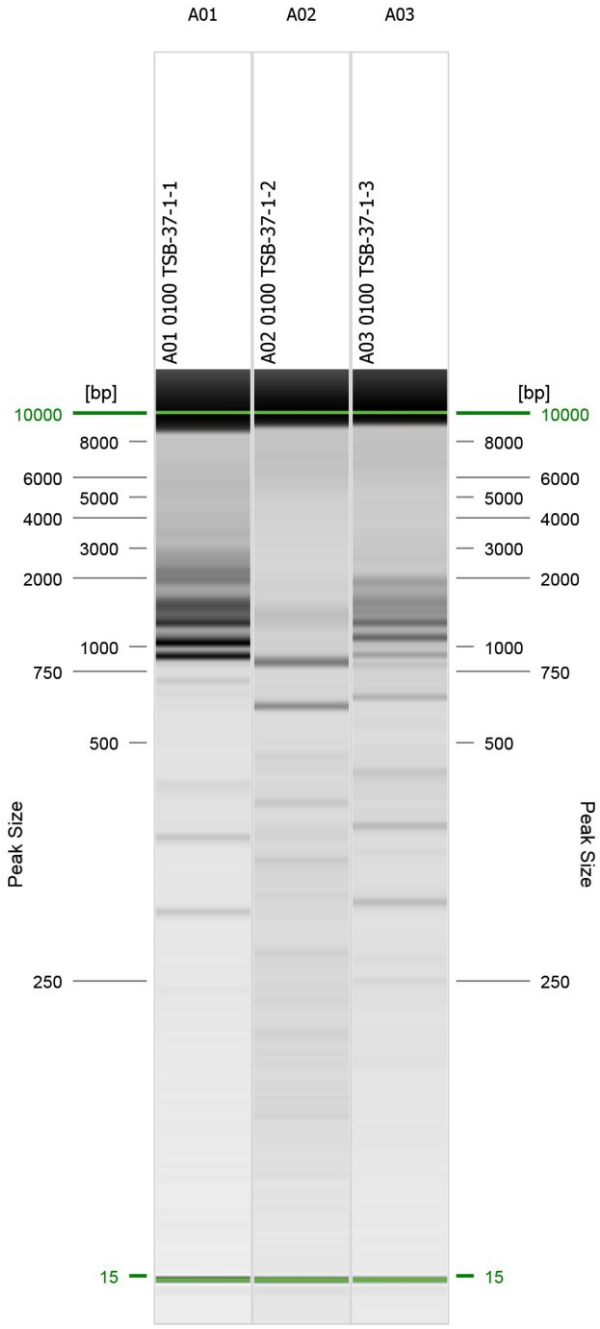

Figure B 1 – A virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 9100.

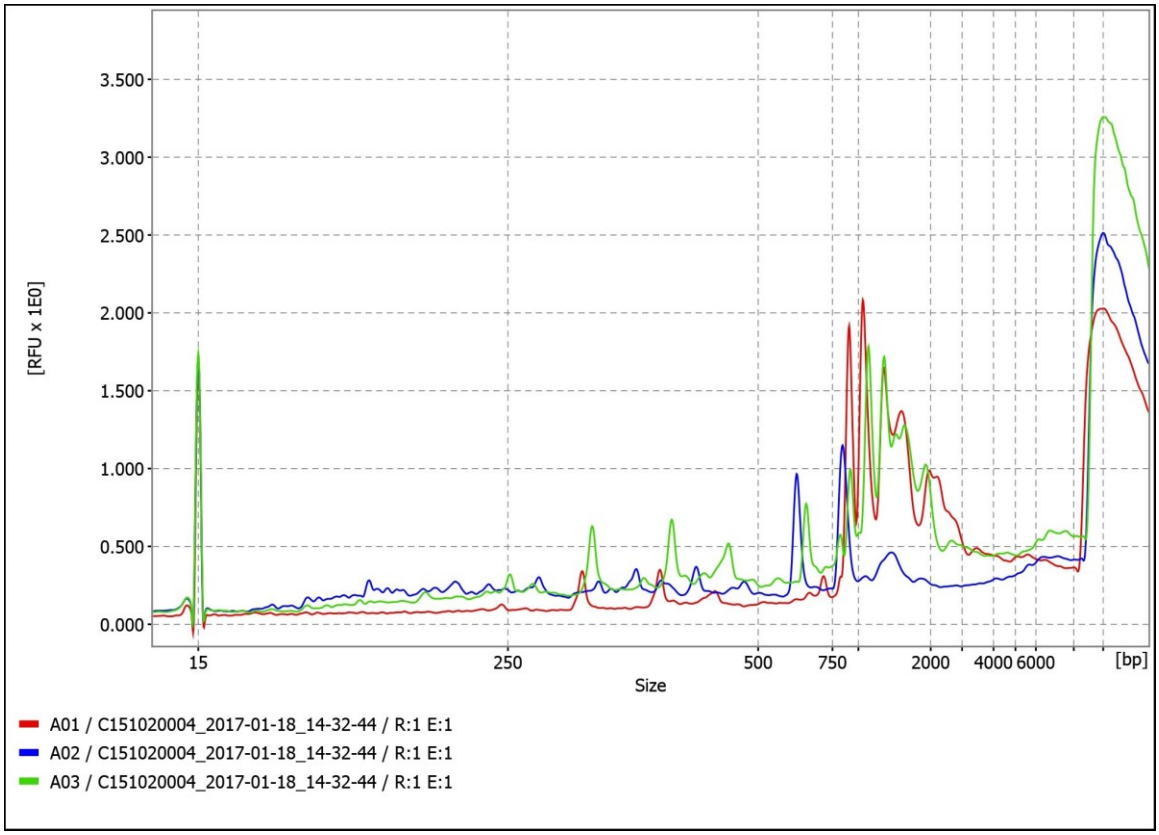

Figure B 2 – An electropherogram of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 9100.

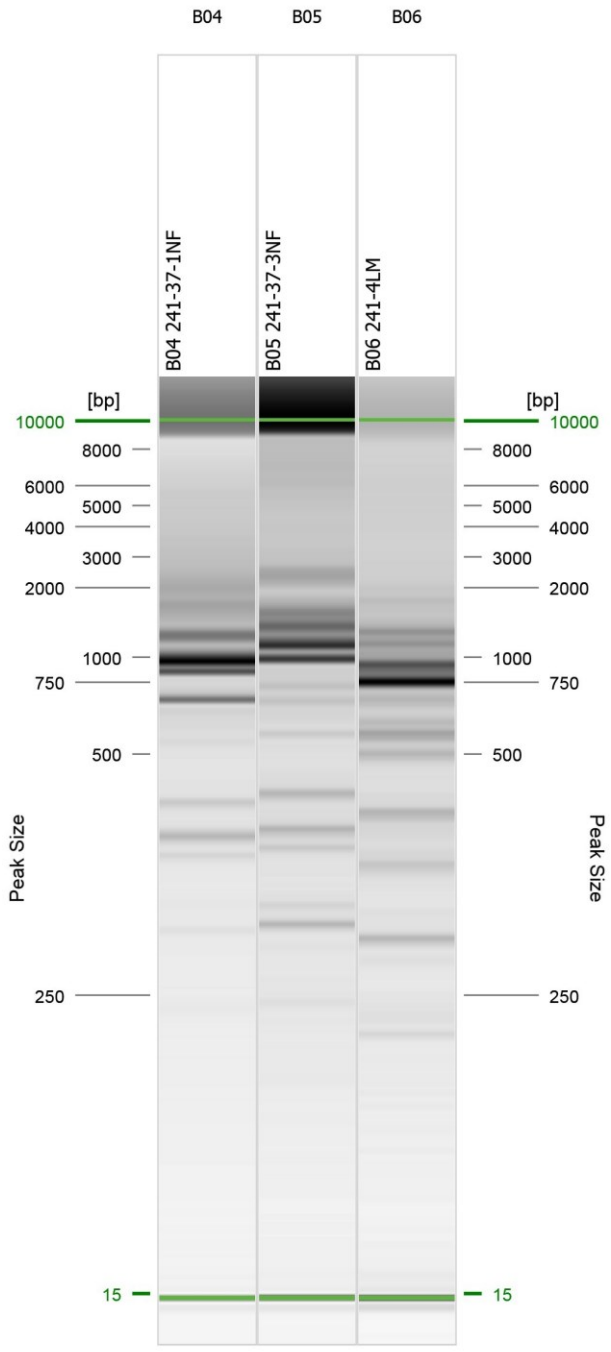

Figure B 3 – A virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 241.

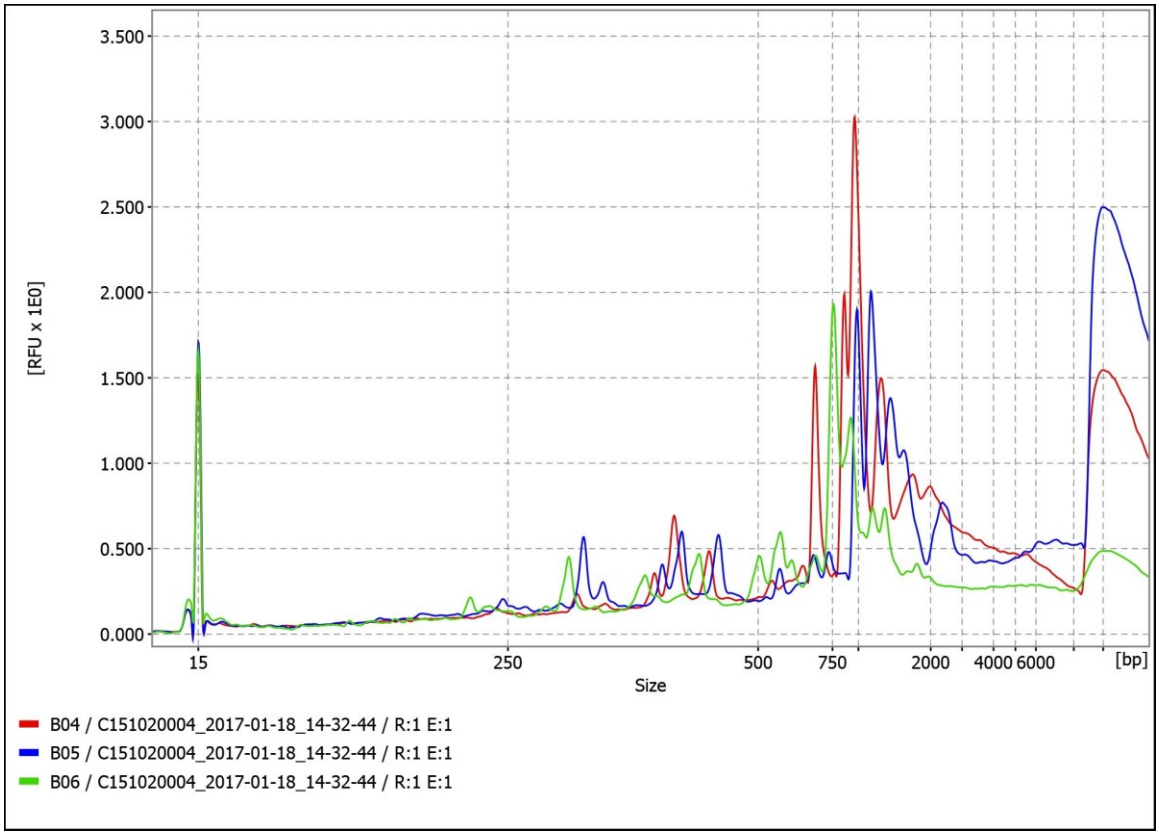

Figure B 4 – An electropherogram of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 241.

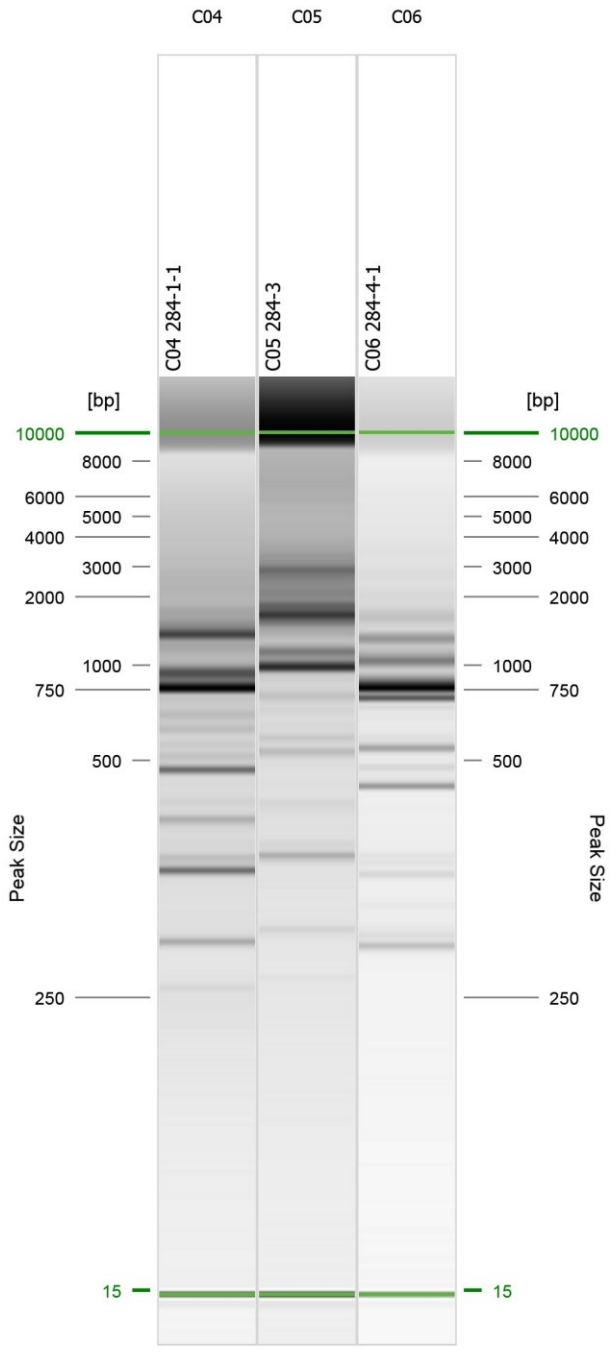

Figure B 5 – A virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 284.

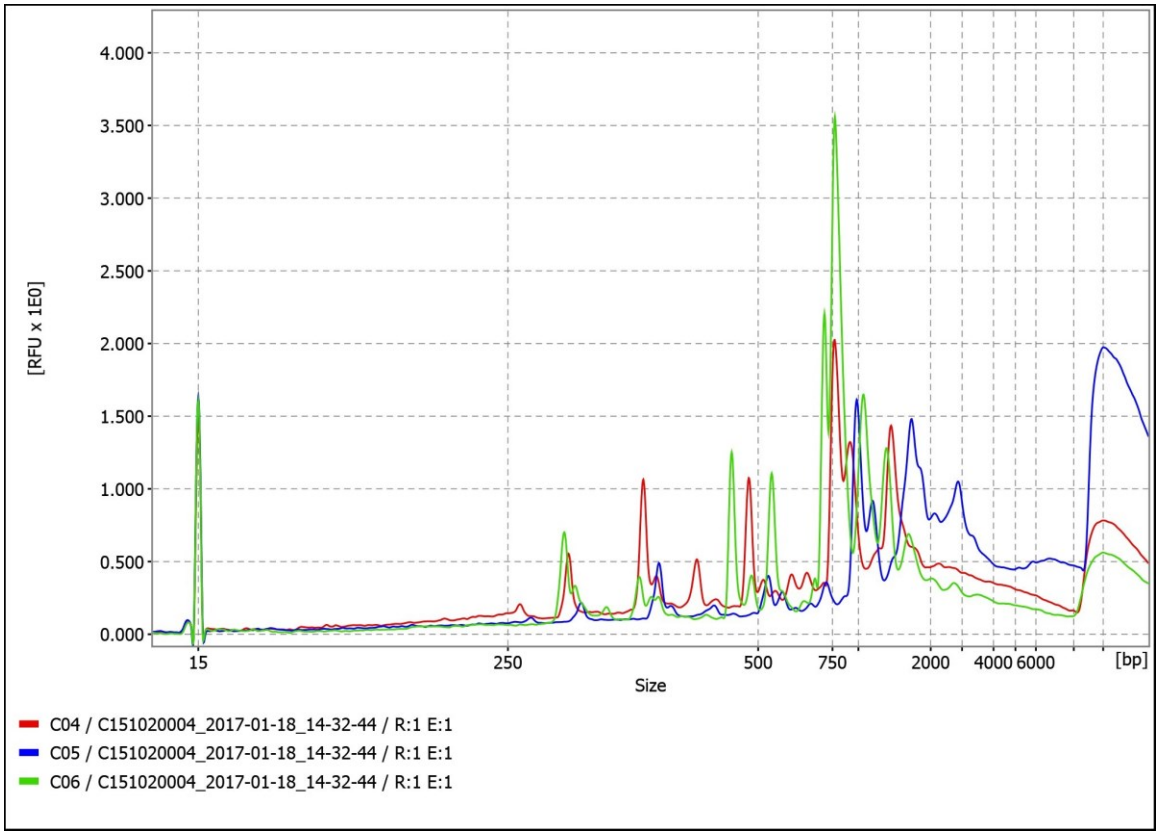

Figure B 6 – An electropherogram of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 284.

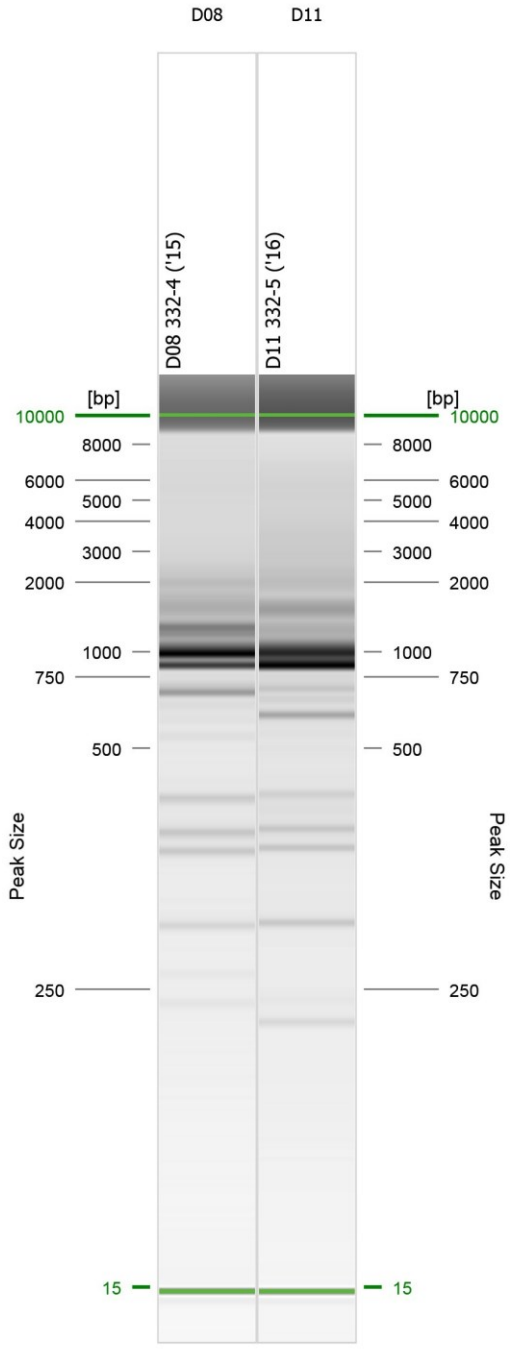

Figure B 7 – A virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the two unique STEC strains recovered from sample # 332.

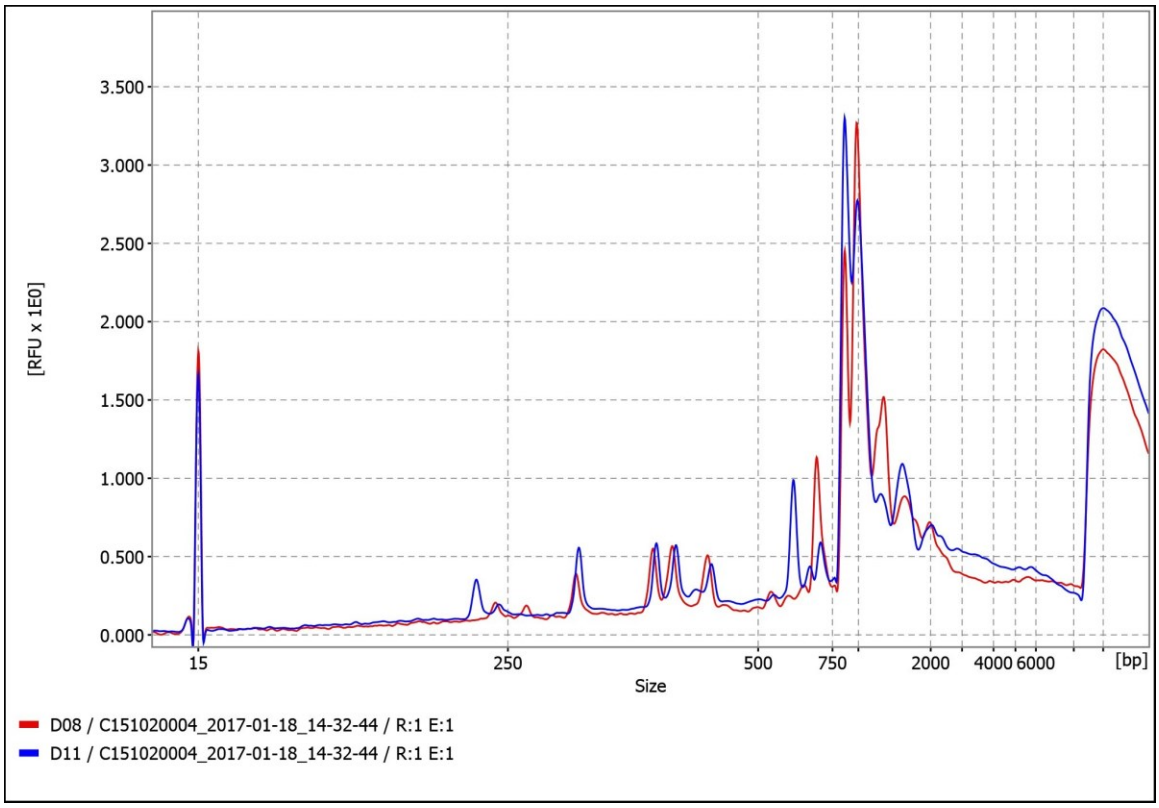

Figure B 8 – An electropherogram of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the two unique STEC strains recovered from sample # 332.

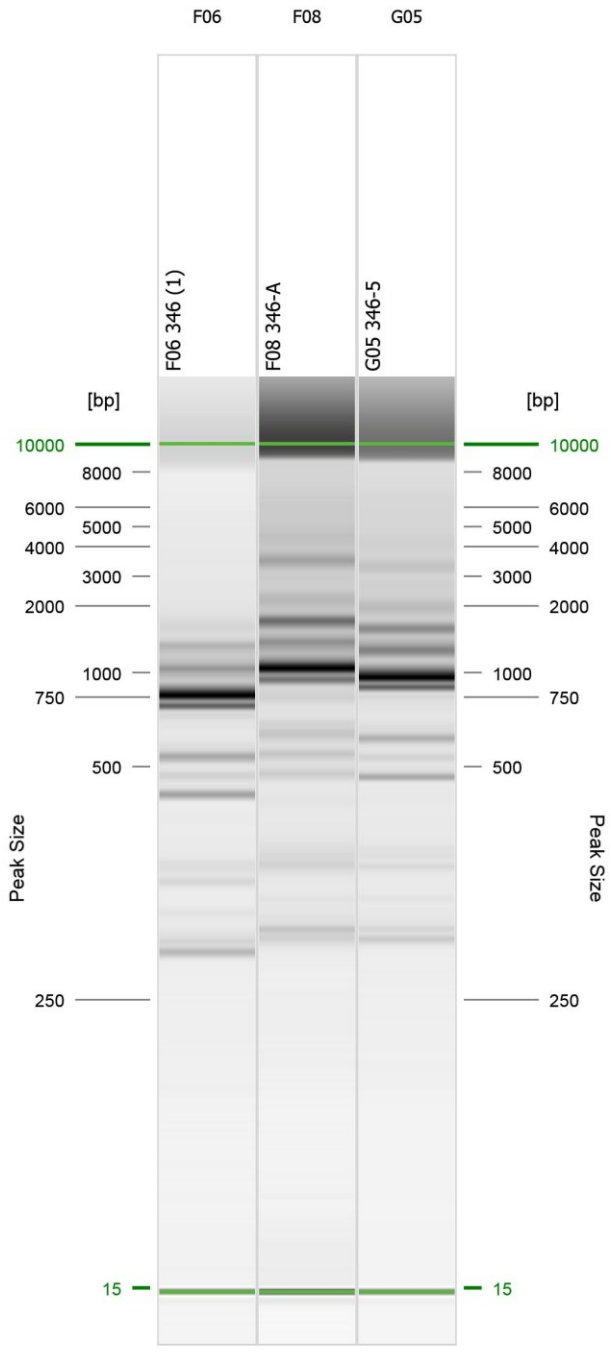

Figure B 9 – A virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 346.

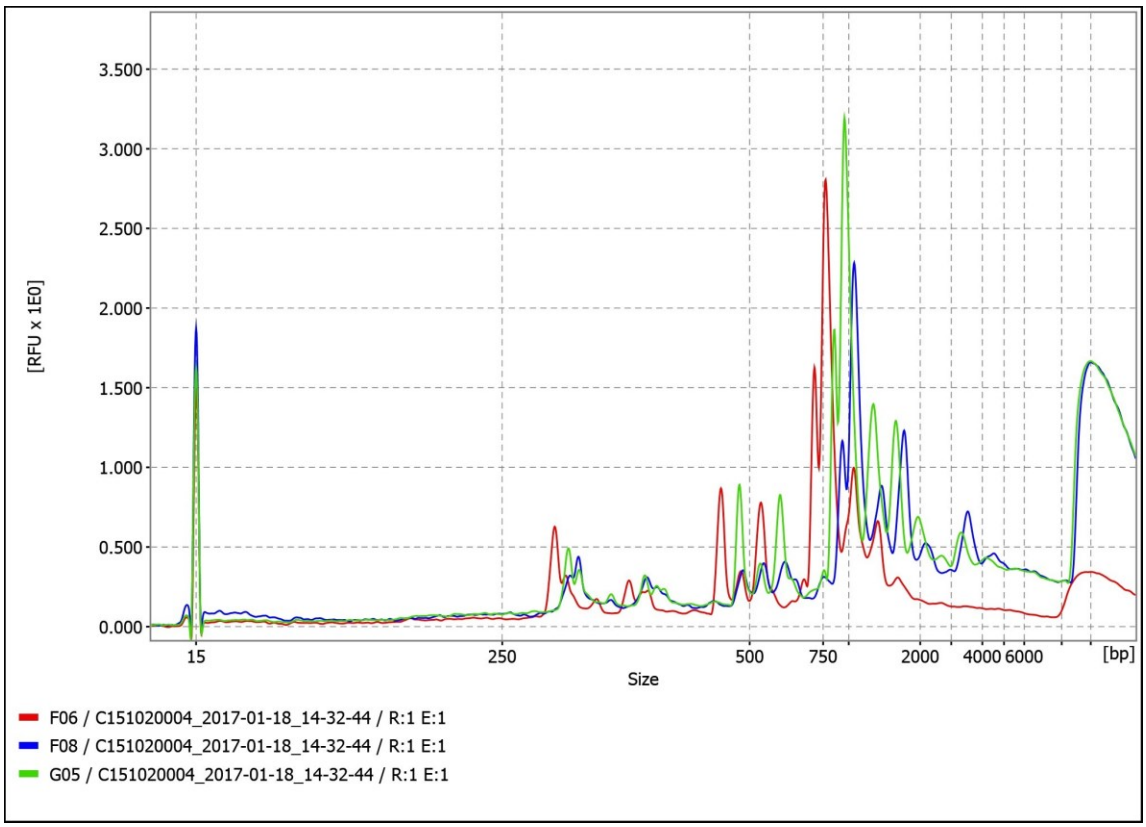

Figure B 10 – An electropherogram of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 346.

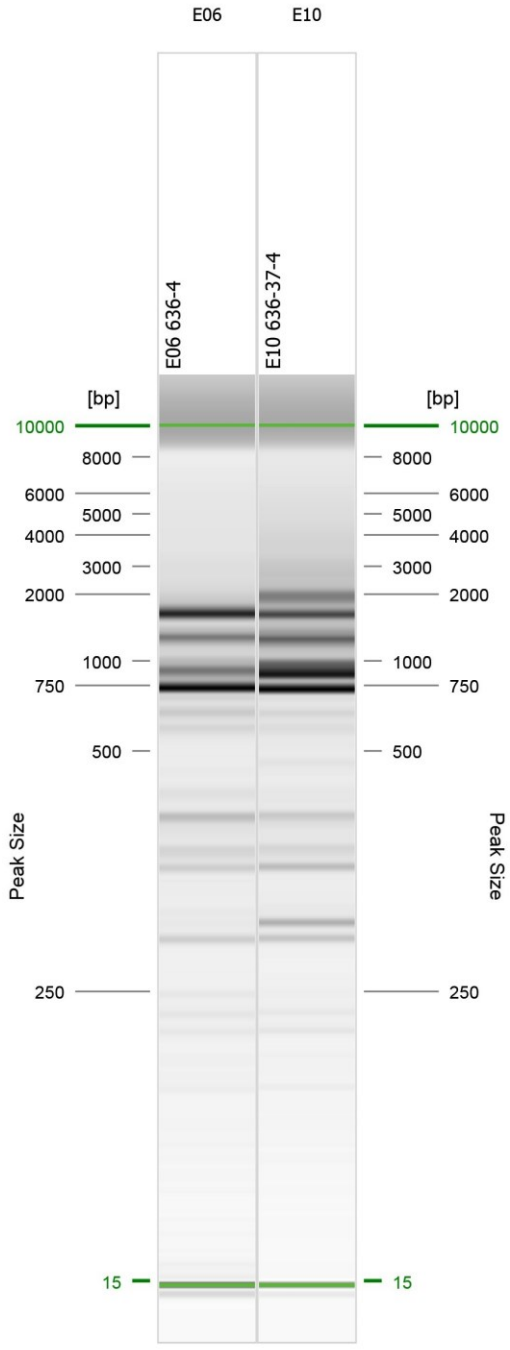

Figure B 11 – A virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the two unique STEC strains recovered from sample # 636.

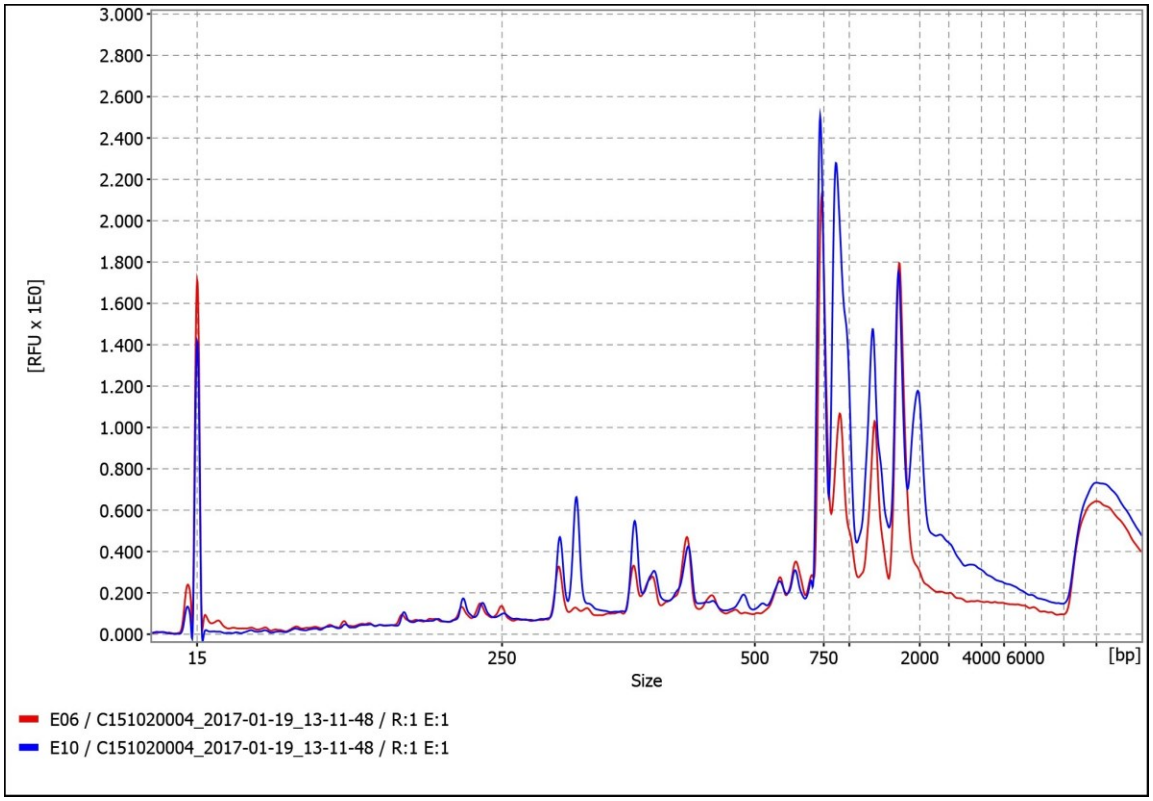

Figure B 12 – An electropherogram of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the two unique STEC strains recovered from sample # 636.

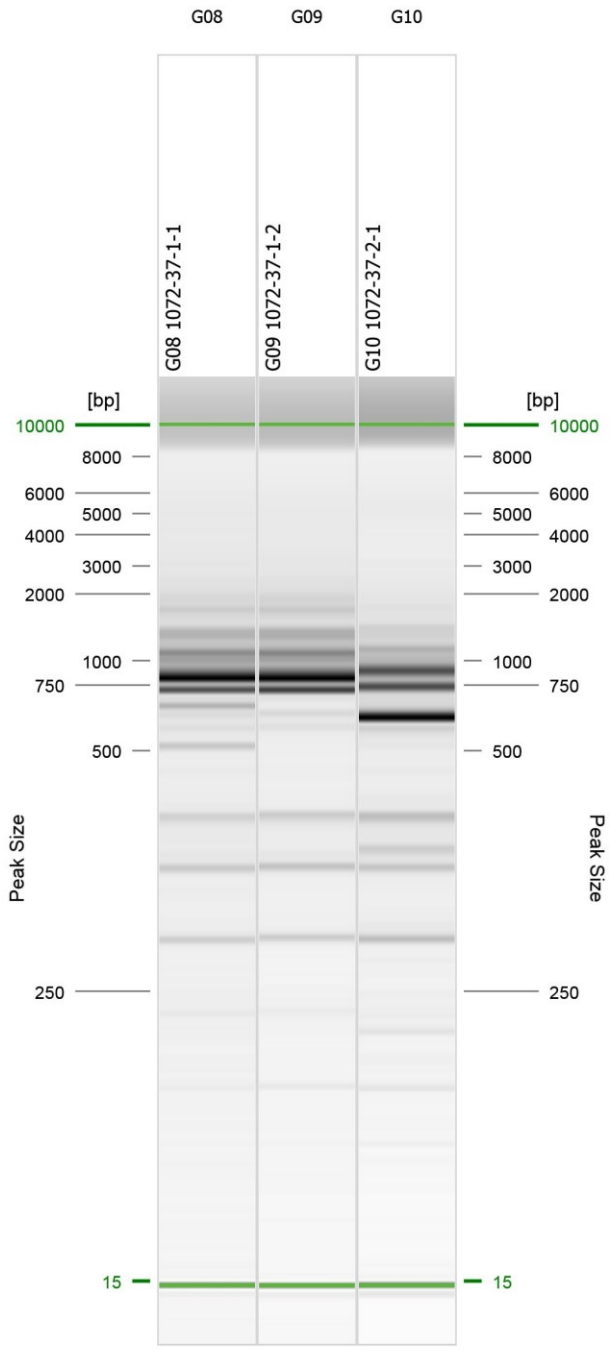

Figure B 13 – A virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 1072.

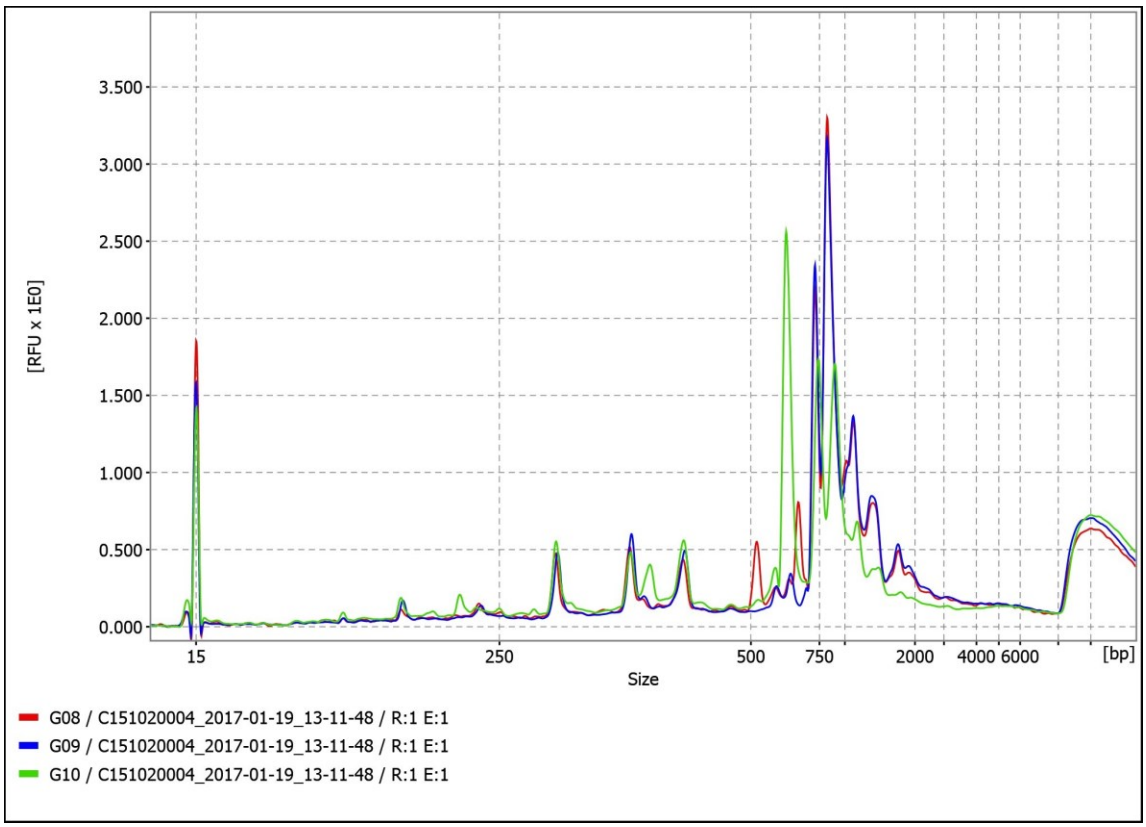

Figure B 14 – An electropherogram of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the three unique STEC strains recovered from sample # 1072.

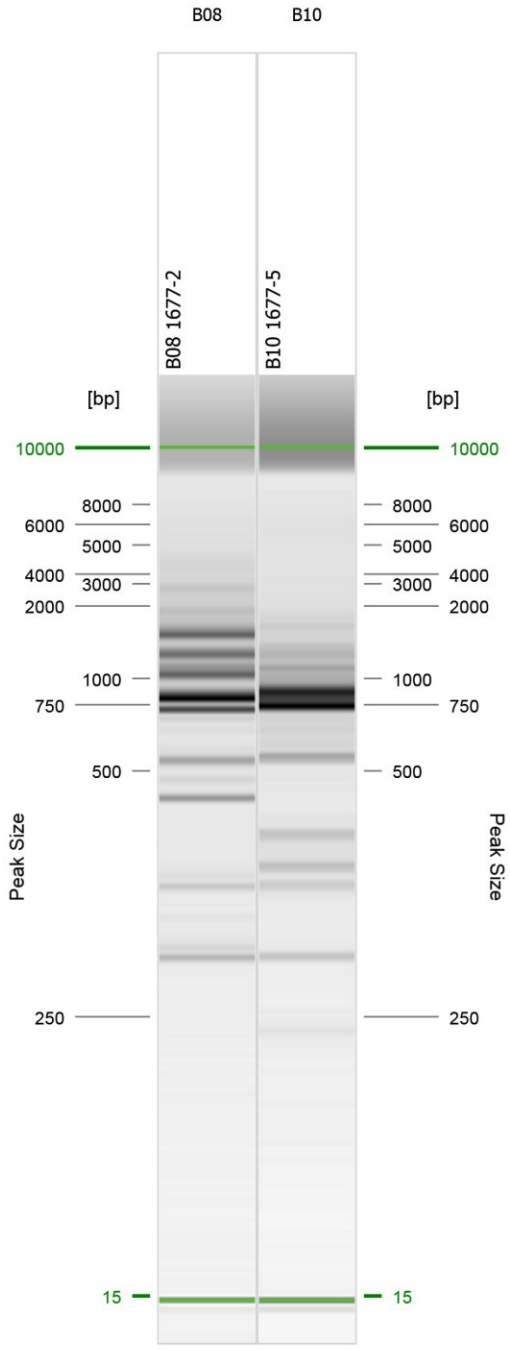

Figure B 15 – A virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the two unique STEC strains recovered from sample # 1677.

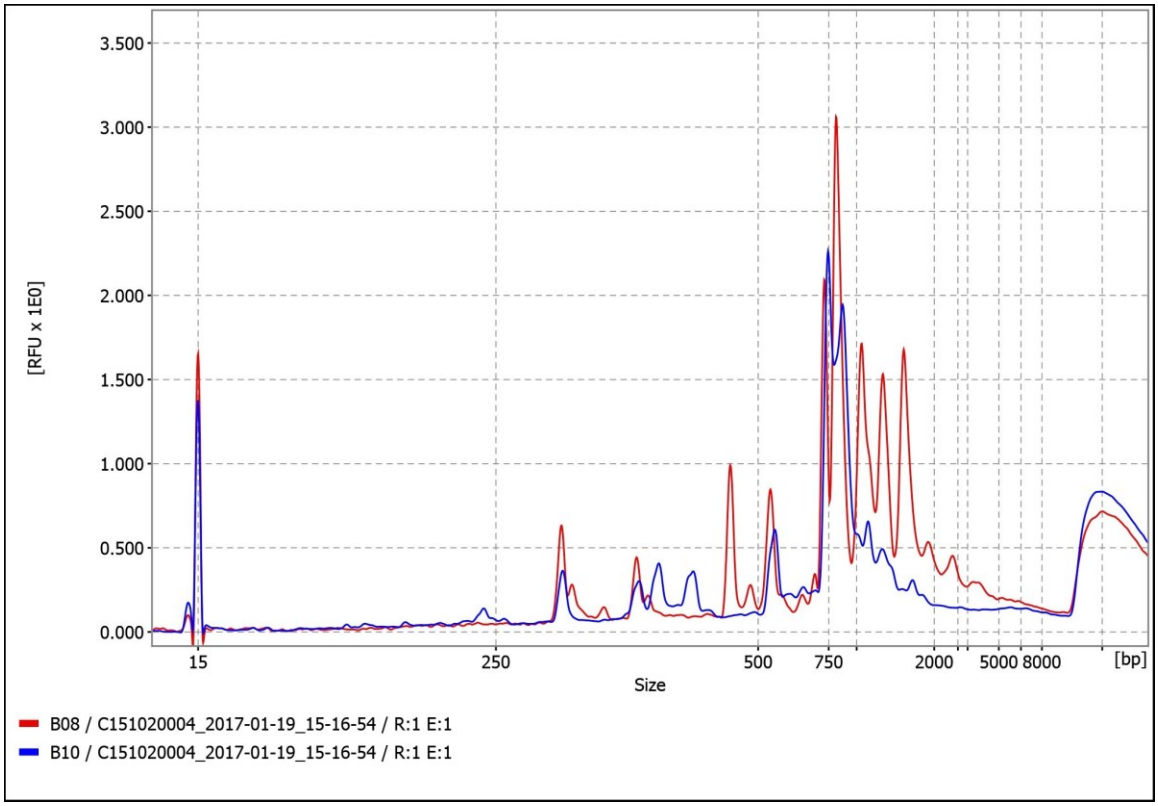

Figure B 16 – An electropherogram of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the two unique STEC strains recovered from sample # 1677.

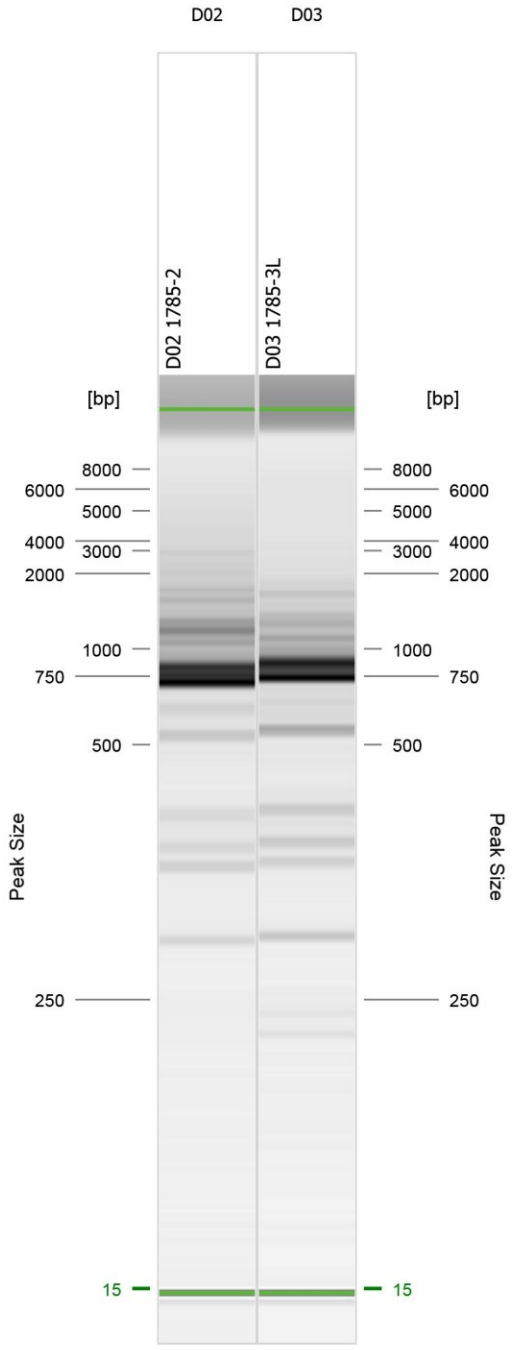

Figure B 17 – A virtual gel image of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the two unique STEC strains recovered from sample # 1785.

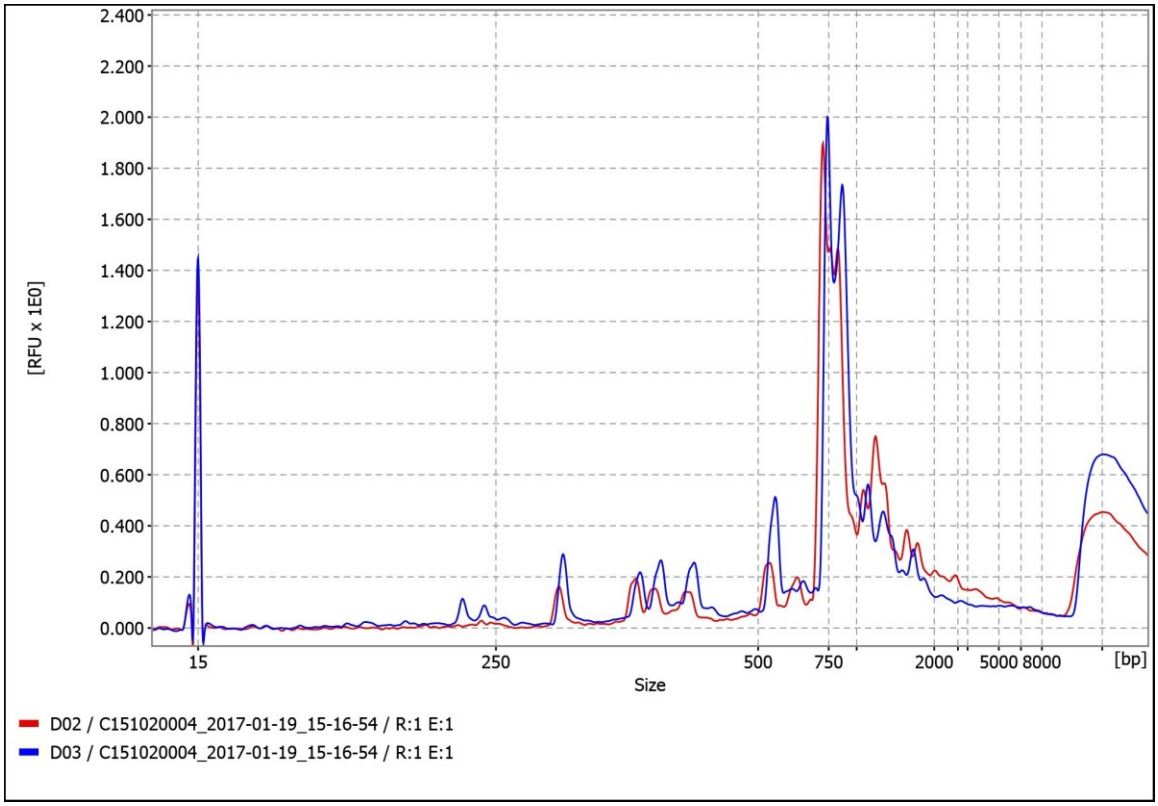

Figure B 18 – An electropherogram of the high-resolution capillary electrophoresis DNA-fragment analysis (QIAxcel® Advanced system) of (GTG)5 PCR products from the two unique STEC strains recovered from sample # 1785.