Molecular Epidemiology of Waterborne Zoonoses in the North Island of New Zealand

A thesis presented in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Veterinary Science
(Epidemiology and Public Health)
at
Institute of Veterinary, Animal and Biomedical Sciences (IVABS),
Massey University, Palmerston North,
New Zealand

Rima Devi Shrestha

2016

(Submitted October 2015)

Abstract

Campylobacter, *Cryptosporidium* and *Giardia* species are three important waterborne zoonotic pathogens of global public health concern. This PhD opens with an interpretive overview of the literature on *Campylobacter*, *Cryptosporidium* and *Giardia* spp. in ruminants and their presence in surface water (Chapter 1), followed by five epidemiological studies of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. in cattle, sheep and aquatic environment in New Zealand (Chapters 2-6).

The second chapter investigated four years of retrospective data on *Campylobacter* spp. (n=507) to infer the source, population structure and zoonotic potential of *Campylobacter jejuni* from six high-use recreational rivers in the Wanganui-Manawatu region of New Zealand through the generalised additive model, generalised linear/logistic regression model, and minimum spanning trees. This study highlights the ubiquitous presence of *Campylobacter* spp. in both low and high river flows, and during winter months. It also shows the presence of *C. jejuni* in 21% of samples containing highly diverse strains, the majority of which were associated with wild birds only. These wild birds-associated *C. jejuni* have not been detected in human, suggesting they may not be infectious to human. However, the presence of some poultry and ruminant-associated strains that are potentially zoonotic suggested the possibility of waterborne transmission of *C. jejuni* to the public. Good biosecurity measures and water treatment plants may be helpful in reducing the risk of waterborne *Campylobacter* transmission

In the third study, a repeated cross-sectional study was conducted every month for four months to investigate the source of drinking source-water contamination. A total of 499 ruminant faecal samples and 24 river/stream water samples were collected from two rural town water catchments (Dannevirke and Shannon) in the Manawatu-Wanganui region of New Zealand, and molecular analysis of those samples was performed to determine the occurrence of *Campylobacter*, *Cryptosporidium*, and *Giardia* spp. and their zoonotic potential. The major pathogens found in faecal samples were *Campylobacter* (n=225 from 7/8 farms), followed by *Giardia* (n=151 from 8/8 farms), whereas *Giardia* cysts were found in many water samples (n=18), followed by *Campylobacter* (n=4). On the contrary, *Cryptosporidium* oocysts were only detected in a few faecal (n=18) and water (n=3) samples. *Cryptosporidium* and *Giardia* spp. were detected in a higher number of faecal samples from young animals

(≤ 3 months) than juvenile and adult animals, whereas *Campylobacter* spp. were highly isolated in the faecal samples from juvenile and adult ruminants. PCR-sequencing of the detected pathogens indicated the presence of potentially zoonotic *C. jejuni* and *C. coli*, *Cryptosporidium parvum* (gp60 allelic types IIA18G3R1 and IIA19G4R1) and *Giardia duodenalis* (assemblages AII, BII, BIII, and BIV) in cattle and sheep. In addition, potentially zoonotic *C. jejuni* and *Giardia duodenalis* assemblages AII, BI, BII, and BIV were also determined in water samples. These findings indicate that these three pathogens of public health significance are present in ruminant faecal samples of farms and in water, and may represent a possible source of human infection in New Zealand.

In the fourth study, PCR-sequencing of *Cryptosporidium* spp. isolates obtained from the faeces of 6-week- old dairy calves (n=15) in the third study were investigated at multiple loci (18S SSU rDNA, HSP70, Actin and gp60) to determine the presence of mixed *Cryptosporidium* spp. infections. *Cryptosporidium parvum* (15/15), *C. bovis* (3/15) and *C. andersoni* (1/15), and two new genetic variants were determined along with molecular evidence of mixed infections in five specimens. Three main *Cryptosporidium* species of cattle, *C. parvum, C. bovis and C. andersoni,* were detected together in one specimen. Genetic evidence of the presence of *C. Anderson* and two new *Cryptosporidium* genetic variants are provided here for the first time in New Zealand. These findings provided additional evidence that describes *Cryptosporidium* parasites as genetically heterogeneous populations and highlighted the need for iterative genotyping at multiple loci to explore the genetic makeup of the isolates.

The *C. jejuni* and *C. coli* isolates (n=96) obtained from cattle, sheep and water in the third study were subtyped to determine their genetic diversity and zoonotic potential using a modified, novel multi-locus sequence typing method ("massMLST"; Chapter 5). Primers were developed and optimised, PCR-based target-MLST alleles' amplification were performed, followed by next generation sequencing on an Illumina MiSeq machine. A bioinformatics pipeline of the sequencing data was developed to define *C. jejuni* and *C. coli* multi-locus sequence types. This study demonstrated the utility and potential of this novel typing method, massMLST, as a strain typing method. In addition to identifying the possible *C. jejuni/coli* clonal complexes or sequence types of 68/96 isolates from ruminant faeces and water samples, this study reported three new *C. jejuni* strains in cattle in New Zealand, along with many strains, such as CC-61, CC-828 and CC-21, that have also been found in humans, indicating the public health significance of these isolates circulating on the farms in the two water catchment areas. Automation of the massMLST method and

may allow a cost-effective high-resolution typing method in the near future for multi-locus sequence typing of large collections of *Campylobacter* strains.

In the final study (Chapter 6), a pilot metagenomic study was carried out to obtain a snapshot of the microbial ecology of surface water used in the two rural towns of New Zealand for drinking purposes, and to identify the zoonotic pathogens related to waterborne diseases. Fresh samples collected in 2011 and 2012, samples from the same time that were frozen, and samples that were kept in the preservative RNA*later* were sequenced using whole-genome shotgun sequencing on an Illumina MiSeq machine. *Proteobacteria* was detected in all the samples characterised, although there were differences in the genus and species between the samples. The microbial diversity reported varied between the grab and stomacher methods, between samples collected in the year 2011 and 2012, and among the fresh, frozen and RNA*later* preserved samples. This study also determined the presence of DNA of potentially zoonotic pathogens such as *Cryptosporidium, Campylobacter* and *Mycobacterium* spp. in water. Use of metagenomics could potentially be used to monitor the ecology of drinking water sources so that effective water treatment plans can be formulated, and for reducing the risk of waterborne zoonosis.

As a whole, this PhD project provides new data on *G. duodenalis* assemblages in cattle, sheep and surface water, new information on mixed *Cryptosporidium* infections in calves, a novel "massMLST" method to subtype *Campylobacter* species, and shows the utility of shotgun metagenomic sequencing for drinking water monitoring. Results indicate that ruminants (cattle and sheep) in New Zealand shed potentially zoonotic pathogens in the environment and may contribute to the contamination of surface water. A better understanding of waterborne zoonotic transmission would help in devising appropriate control strategies, which could reduce the shedding of *Campylobacter*, *Cryptosporidium*, and *Giardia* spp. in the environment and thereby reduce waterborne transmission.

Preface

This PhD thesis aimed to study the molecular epidemiology of waterborne zoonosis in New Zealand, focussing on top three notifiable diseases: campylobacteriosis, cryptosporidiosis, and giardiasis. The project aimed to determine the presence of *Campylobacter*, *Cryptosporidium*, and *Giardia* spp. in ruminants (cattle and sheep) on farms and surface water in two catchment areas in the North Island: Dannevirke and Shannon. In addition to providing relevant epidemiological data, this project also developed a novel typing method, "massMLST" and applied state of the art metagenomic approaches using next generation sequencing technology.

Acknowledgements

"Cultivate the habit of being grateful for every good thing that comes to you, and to give thanks continuously. And because all things have contributed to your advancement, you should include all things in your gratitude."

— Ralph Waldo Emerson

First and foremost, I would like to express enormous gratitude to my supervisors for their excellent help, guidance, and encouragement to accomplish the research work presented in this thesis. Thank you Dr. Alexandro Grinberg for your inspiring guidance to help me understand the concept of molecular studies, preparing manuscript, including this thesis. I am also enormously thankful to Dr. Patrick Biggs for bringing the concept of a novel "massMLST" method, helping me to understand the concept of next-generation sequencing, and assisting in bioinformatics data analysis. I wish to express my profound gratitude to Prof. Dr. Nigel P. French for helping me to find the research funds, providing amicable suggestions during the study period, particularly in the statistical analysis of chapter 2. Dr. Evelyn J. Pleydell, thank you for your valuable suggestions, especially for the chapter 1, 2 and 3 of this PhD study and supporting me morally when I was on stress. Thanks also go to Dr. Deborah J. Prattley for her assistance in improving my writing techniques.

This thesis would not have been prepared without the help of molecular epidemiology laboratory (mEpiLab) staffs: Dr. Anne Midwinter, Dr. V.S.R Dukkipati, Lynn Rogers, Sarah Moore, Anthony Pita, Niluka Velathanthri, Angy Reynolds, Errol Kwan, Tessy George, Rukhshana Akhtar, and Neville Haack; and Massey genome services staffs Trish McLenachan and Richard Fong. Thank you all for your technical assistance and laboratory expertise.

I am also thankful to Dr. Jonathan Marshall for providing help in analysing the results, Dr. Mark Stevenson, Dr Julie Collins-Emerson and Dr. Jackie Benschop for their valuable suggestions and encouragement during my study period.

I am immensely grateful to farmers involved in this study, Horowhenua and Tararua District Councils for allowing me permission to take faecal/water samples from their premises.

I gratefully acknowledge the support of New Zealand Development Scholarship, International Student Support Office, McGeorge Research Funding Agency, Ministry of Health, and IVABS' Postgraduate Research Fund Committee, without which I could not have undertaken the PhD studies. Thanks also go to mEpiLab and IVABS administrative staffs, Wendey Mahrey, Simon Verschaft, and Debbie Hill.

I am heartily indebted to my fellow students from mEpiLab, Epicentre and IVABS, especially Tessy, Springer, Patricia, Hamid, Fang, Shoukai, Barbara, Anja, Ben, Julanda, Peter, Emilie, Masako, Anou, Nelly, and Late Mbvuto for their friendship, sharing thoughts, widening my horizon, and being kind to me during the moments of stress. I am also thankful to Nepalese in the Palmerston North for their support during the first year of my study.

I am thankful to my parents, in-laws, and all my family members who directly and indirectly provided moral supports for completing my study.

Finally yet importantly, I am indebted to my husband, Saroj Yadav, for his continuous support during my study period, especially for driving vehicle for sampling, cooking delicious food, creating and performing music, and counselling and understanding me when I was in stress.

List of Contents

List of Tables

List of Figures

List of Abbreviations

18S SSU rDNA /18S rRNA	Small Subunit 18S Ribosomal RNA
AIC	Akaike Information Criterion
BA	Blood Agar
bg	Beta-Giardin
bp	Base Pairs
CC	Clonal Complex
CDC	Centre For Disease Control
DAPI	4-,6-Diam,Idino-2-Phenylindole
ELISA	Enzyme Linked Immunosorbent Assay
ESR	The Institute Of Environmental Science And Research
FAO	Food And Agriculture Organisation
GAM	Generalised Additive Model
Gdh	Glutamate Dehydrogenase
GIS	Geographical Information System
GLM	Generalised Linear Model
GP60	Glycoprotein (Or 60-Kda Glycoprotein)
HSP70	70 kDa Heat Shock Protein Gene
IFA	Immunofluorescence Assay
IMS	Immunomagnetic Separation
IVABS	Institute Of Veterinary, Animal And Biomedical Sciences
mCCDA	Modified Charcoal, Cefoperazone Desoxycholate Agar
MEGAN	Metagenome Analyser
MGS	Massey Genome Service
MGW	Molecular Grade Water
MLST	Multilocus Sequence Typing
MSSP	Manawatu Sentinel Surveillance Program
MST	Minimum Spanning Tree

MU	Massey University
mEpiLab	Molecular Epidemiology And Public Health Laboratory
NCBI	National Centre For Biotechnology Information
NGS	Next-Generation Sequencing
NIWA	National Institute Of Water And Atmospheric Research
NZGL	New Zealand Genomics Limited
OPG	Oocysts Per Gram Of Faeces
OR	Odds Ratio
PAUDA	Protein Alignment Using A DNA Aligner
PCoA	Principal Co-Ordinate Analyses
PCR	Polymerase Chain Reaction
pDNA	Pseudo-DNA
PEG	Polyethylene Glycol
PFGE	Pulsed Field Gel Electrophoresis
PRU	Protozoa Research Unit
QC	Quality Check
RFLP	Restriction Fragment Length Polymorphism
SNPs	Single Nucleotide Polymorphisms
spp	Species
ST	Sequence Types
Tpi	Triosephosphate Isomerase
UPGMA	Unweighted Pair Group Method With Arithmetic Mean
USEPA	United States Environmental Protection Agency
UV	Ultra Violet
VBNC	Viable But Non-Culturable
WGS	Whole Genome Shotgun Sequencing
WHO	World Health Organisation
WINZ	Water Information New Zealand

Chapter 1

Campylobacter, Cryptosporidium and *Giardia* spp. in ruminants and their contamination to surface water: A literature review

1.1 Background

In the1860s, the English physician John Snow for the first time studied and traced cholera as a waterborne disease (Okun, 1996). Thereafter, several other waterborne diseases such as typhoid, cryptosporidiosis, giardiasis and campylobacteriosis have been described and detected worldwide, indicating that water can be contaminated with a variety of pathogenic microorganisms (Mackenzie et al., 1994; O'Connor, 2002; Okun, 1996). As a result, water quality and safety have become major concerns due to the risk that continues to be posed to public health worldwide by major life-threatening illnesses such as diarrhoea, haemolytic uremic syndrome (HUS) and Guillain-Barre syndrome. Many studies report direct or accidental ingestion of water contaminated with human faeces as one of the sources of waterborne disease. For example, bathers in coastal waters and in swimming pools often contract waterborne diseases associated with viruses/protozoa related to humans. Therefore, many countries have implemented water source management plans and developed guidelines for recreation and drinking water use to control faecal contamination such as sewage pollution (WHO, 2003; 2011). The guidelines address the protection of drinking water sources, the treatment of drinking water, compliance standards for drinking water quality and the monitoring of aquatic resources. Implementation of the guidelines has sharply reduced the incidence of waterborne enteric diseases in countries such as the USA, Canada, the UK, and New Zealand. Nevertheless, waterborne

diseases persist as one of the major problems in both developed and developing countries, causing more than 3.4 million deaths every year (Gleick, 2002; WHO, 2004; 2011).

1.1.1 Faecal pollution source

Human faecal contamination of water sources is often identified and is adequately controlled in many countries. However, control of non-human faecal contamination is not always adequately addressed, although animal to human waterborne transmission of several pathogens has been recognised (O'Connor, 2002; Wilson et al., 2008; WHO, 2012). For example, in the year 2000, waterborne *E. coli* O157:H7 and *Campylobacter jejuni* outbreaks reported in Walkerton, Canada caused more than 2,300 gastrointestinal disease cases and seven deaths. This outbreak was related to drinking water from a well contaminated with cattle manure after a period of heavy spring rainfall (Hrudey et al., 2002; O'Connor, 2002). Similarly, Wilson *et al.* (2008) reported that 96.6% of *C. jejuni* infections in the UK were attributable to farm livestock. Therefore, the examination of non-human sources of faecal pollution and the factors associated with such pollution are important to control the risks posed to human health by water.

According to the Food and Agriculture Organization (FAO, 2007) report on global human and agricultural sources of faecal pollution, cattle and sheep contribute 65% of total faecal pollution. This high percentage suggested that ruminant faecal contamination of water sources potentially contained zoonotic pathogens such as *Campylobacter*, *Cryptosporidium* and *Giardia* (Muirhead, 2004; WHO, 2012). Nevertheless, several factors such as the type of ruminants and their density, their faecal load in the environment, the characteristics of the catchments, natural events such as precipitations, and the pathogens themselves contribute to variations in water contamination and human infections.

Increasing populations have and always will demand an increased food supply. This demand has driven the dramatic growth in ruminant livestock populations in the last few decades (FAO, 2014). It is predicted that the human population will be ~9 billion in 2050 and that global meat and milk production will nearly double to 465 and 1,043 million tonnes, respectively (FAO, 2006). The rise in ruminant populations has also led to a growing concern over land use change, water pollution and an unacceptable public health risk concerning waterborne diseases (WHO, 2004; Tran et al., 2010). In addition, studies on the distribution of waterborne pathogens conducted in different

countries, time-periods, rainfall events, and pathogens' survival in the environment have shown evidence that the presence of ruminants is a risk factor for water contamination. Therefore, understanding the role of ruminants in water contamination is vital for preventing sporadic cases and outbreaks of waterborne diseases, particularly in a country like New Zealand where there is a large ruminant population living on pasture and around surface water sources.

1.1.2 Pathogens in water and molecular studies

Approximately, 243 of the 616 livestock pathogens (39%) are currently known to infect humans (Cleaveland et al., 2001) and a few of these pathogens are associated with waterborne transmission routes. Public health agencies have classified higher priority zoonotic pathogens and ranked *Cryptosporidium* and *Giardia* into Rank 1 and *Campylobacter* into Rank 2 based on the criteria of zoonotic evidence, confirmed waterborne transmission route, occurrence of disease outbreaks in healthy humans, seriousness of illnesses and their consequences, global distribution and susceptibility to water treatment (Suresh et al., 2012). These pathogens are also found in the gastrointestinal tract of ruminants and contaminated waterways when faeces containing pathogens are either deposited or flushed into the water. In addition, these pathogens contributed to 26% of the total infectious disease burden in low-income countries and 0.7% in high-income countries (Grace et al., 2012), and resulted in death in an estimated 700 000 diarrhoeal episodes in 2011 (Walker et al., 2013). These three pathogens are also relevant to New Zealand, where they accounted for 41% of all notifiable diseases in 2013. In addition, New Zealand has the highest rate of enteric disease notification in humans among the developed countries (Baker et al., 2007a; Snel et al., 2009). Hence, the enteric zoonotic pathogens *Campylobacter*, *Cryptosporidium* and *Giardia* in ruminants were chosen as the focus of this thesis.

Recent advances in molecular biology have allowed the development of more sensitive testing methods, revolutionising source tracking of pathogens. For example, the use of next-generation sequencing technology in the investigation of a large foodborne *Escherichia coli* O104:H4 outbreak that occurred in 2011 allowed tracking of the infection to contaminated fenugreek seeds imported from Egypt (Bielaszewska et al., 2011; Scheutz et al., 2011). Such integration of molecular techniques in epidemiological studies of diseases has provided a better understanding of the transmission dynamics of pathogens. Multiple molecular epidemiological studies have also been conducted using techniques such as species-specific PCR and multi-locus

sequence typing to track the source of *Campylobacter*, *Cryptosporidium* and *Giardia* contamination in water (Kwan et al., 2008). However, due to complex transmission pathways of these pathogens, molecular epidemiological studies are warranted for unravelling these pathogens' transmission dynamics.

Therefore, this review focuses on the major aspects of the waterborne zoonotic pathogens *Campylobacter*, *Cryptosporidium* and *Giardia* species; the diffusion of these pathogens into water; their environmental survival and transport into waterways; the current methods used for identification of pathogens in water and/or faeces, and approaches to prevent waterborne diseases. Pathogens refer to "*Campylobacter*, *Cryptosporidium* and *Giardia"* spp. where not specified.

1.2 *Campylobacter* spp. infection in humans and ruminants

Theodor Escherich, a German-Austrian paediatrician and bacteriologist, first discovered the bacterium *Campylobacter* in 1886 when investigating the infant stool specimen. Subsequently, Smith and Taylor grouped this organism into the genus *Vibrio* in 1919 (cited in Nachamkin et al., 2008). However, Sebald and Veron (1963) identified serological and biochemical differences of this organism from *Vibrio*, and therefore established a new genus "*Campylobacter"*. The genus *Campylobacter* comprises small, spiral-shaped, non-spore-forming, motile, microaerophillic, Gram-negative bacteria. *Campylobacter* is a taxonomically complex genus that currently comprises 22 species and eight subspecies (Debruyne et al., 2010), of which *Campylobacter jejuni*, *Campylobacter coli*, *Campylobacter lari*, *Campylobacter upsaliensis,* and *Campylobacter helveticus* are commonly isolated from humans and animals worldwide. Of these, *C. jejuni* and *C. coli* are two zoonotic species of *Campylobacter* that are of major public health concern worldwide.

1.2.1 Campylobacteriosis in humans

Ruminants have been considered an important source of human campylobacteriosis only since 1980, although *Campylobacter* bacteria were known to cause disease (termed campylobacteriosis) in animals for more than seven decades before then. Contaminated food is considered the main transmission vehicle in the majority of human campylobacteriosis cases, but waterborne campylobacteriosis cases have been frequently reported in many countries, such as in the USA, UK, Norway and New

Zealand (CDC, 2013a; 2013b; EFSA, 2014; ESR, 2014). Waterborne campylobacteriosis cases are often related to drinking improperly treated or untreated water. It is estimated that the consumption of 500 *Campylobacter* organisms can cause gastroenteritis with the onset of prodromal symptoms: fever, headache, myalgia and malaise followed by diarrhoea within 48 hours of infection (Black et al., 1988; Kirkpatrick & Tribble, 2011). *Campylobacter* infections are self-limiting with symptoms lasting between three to six days, but some individuals (1-2%) may develop sequelae of reactive arthritis, irritable bowel syndrome and Guillain-Barré syndrome (Smith, 1995; Keithlin et al., 2014). Antibiotic treatment is controversial and most *Campylobacter* strains have been reported to be resistant to cephalosporins and fluoroquinolones, two groups of antibiotics that are generally used to treat animal and human illness (Koenraad et al., 1995; Alfredson and Korolik, 2007; Silva and Teixeira, 2015).

In many countries, including New Zealand, Australia, the UK and Nordic countries, seasonal peaks of campylobacteriosis have been reported in summer (Hudson et al., 1999; Nylen et al., 2002; Meldrum et al., 2005). Campylobacteriosis cases have also been reported across all ages, although higher incidence is reported in immunocompromised people and young children. Young children (<4 years of age) living in close proximity to high densities of livestock and people living in rural areas have been shown to be at greater risk than those living in urban dwellings (Green et al., 2006; Strachan et al., 2009; Fitzenberger et al., 2010; Spencer et al., 2011; Gilpin et al., 2013). Strachan *et al.* (2009) found that the source of infection in rural areas is predominantly ruminant-associated *C. jejuni* whilst poultry meat seems to be a more common source in urban areas. Therefore, it is likely that environmental sources such as water are implicated in *Campylobacter* infection in people living in rural areas.

The mechanisms by which *Campylobacter* spp. cause disease are not yet fully understood despite the presence of genomic information on *Campylobacter* spp. Studies have shown that *C. jejuni* can attach to and invade intestinal epithelial cells, respond to changing environments, colonise intestines, secrete toxins and evade host immune responses (Sasakawa, 2009; Dasti et al., 2010; Man, 2011). The flagellum and chemoreceptor genes in *C. jejuni* are responsible for colonisation and infection by facilitating the invasion and adhesion of bacteria into the epithelial cells, resulting in inflammation of the gut. *C. jejuni* has also been reported to produce cytotoxins and a cholera-like enterotoxin, provoking severe symptoms (Sasakawa, 2009; Dasti et al., 2010; Man, 2011). Moreover, it is believed that superoxide dismutase (sodB), which

converts superoxide into hydrogen peroxide, and a heat shock protein virulence factor in *Campylobacter* spp. enable the intracellular survival of *Campylobacter* in the host.

1.2.2 Campylobacteriosis in farmed ruminants

Campylobacter spp. were isolated for the first time from infected animals in 1909, and have long been suspected as a cause of infectious abortion in cattle, buffalo and sheep (as cited in Nachamkin et al., 2008). *Campylobacter* spp. appear to be commensal in the gastrointestinal tract, and clinical disease is not commonly observed in ruminants (Altekruse et al., 1994; Silva et al., 2011). Some *Campylobacter* spp. can cause venereal disease characterised by infertility, embryo mortality, abortion, irregular oestrous cycles and long calving intervals (Irons et al., 2004). In cattle, genital campylobacteriosis may cause a reduction in pregnancy rates to as low as 20%, abortion rate as high as 10%, and sterility in up to 11% of infected heifers (Hum, 1987; McCool et al., 1988). *Campylobacter fetus venerealis* is the main cause of cattle genital campylobacteriosis (Mshelia et al., 2010). *Campylobacter fetus* and *C. jejuni* found in the genital and intestinal tracts of cattle, sheep and goats may also cause infectious infertility and a wide variety of invasive diseases in humans (Mannering, 2006; Sahin et al., 2007; Wagenaar et al., 2014).

1.2.2.1 Occurrence of *Campylobacter* spp. in farmed ruminants

Campylobacter species often inhabit the intestinal tract of ruminants, but their epidemiology is complex and not yet fully elucidated. Although *Campylobacter* spp. have been isolated from many ruminant hosts such as sheep, goats and buffalo, cattle are considered the main reservoir of infection (Nachamkin et al., 2008). *Campylobacter jejuni* and *C. coli* shed by ruminants are considered the most important source of human infections through water (Nachamkin et al., 2008). However, little data are available on the role of other *Campylobacter* species in waterborne illnesses related to ruminants. For example, *Campylobacter fetus* subsp *fetus/veneralis* that causes fertility disorders in cattle and sheep has also been isolated from humans (Wagenaar et al., 2014). This finding suggests that, though rare, zoonotic transmission of *C. fetus* is possible. It has been hypothesised that direct contact with cattle and sheep is the main cause of *C. fetus* transmission to humans (Wagenaar et al., 2014). The environmental transmission route has not yet been established. Therefore, the following subsections briefly summarise *Campylobacter* spp., particularly *C. jejuni* and/or *C. coli*, in cattle, sheep, goats and other ruminants.

1.2.2.1.1 Cattle as a reservoir of Campylobacter spp.

The prevalence of *Campylobacter* spp. in cattle has been widely studied worldwide. These studies reported *Campylobacter* prevalence ranging from 0.8 to 93% in dairy cattle (Rosef et al., 1983; Nielsen, 2002; Adhikari et al., 2004; Grinberg et al., 2005; Oporto et al., 2007; Chatre et al., 2010; Ramonaite et al., 2013) and 6 to 68% in beef cattle (Oporto et al., 2007; Chatre et al., 2010; Sproston et al., 2011; Khalifa et al., 2013). The variations in reported prevalences might be due to differences in sampling techniques such as sampling of rectal swabs or from freshly voided faecal samples, sampling time, the age of the animal, isolation and analytical methods, management practices and environmental factors. Intermittent shedding of *Campylobacter* spp. by cattle might also produce differences in *Campylobacter* prevalence (Devane et al., 2005). In one Norwegian study, the presence of *Campylobacter* spp. was reported in only 2/254 (0.8%) rectal swabs collected from cows (Rosef et al., 1983). The lower prevalence reported could be due to the storage of samples at 4 ^{o}C for 2-3 days before analysis (Rosef et al., 1983; Giacoboni et al., 1993). Another study that collected 686 samples directly from the rectum and faecal pats on the ground of 15 cattle farms in Washington State, USA, showed prevalences of 43.2% and 34.1% of *Campylobacter* spp. and *C. jejuni* respectively (Bae et al., 2005). A high faecal prevalence of *Campylobacter* spp. (87%; 2415/2776) was reported in freshly voided faecal samples collected from feedlot cattle in Alberta, Canada (Hannon et al., 2009). It is likely that seasonal and spatial factors may affect the prevalence of *Campylobacter* spp. in cattle. For example, a longitudinal study was conducted on five cattle farms in Cheshire, UK over an 11-month period; this identified high *C. jejuni* prevalence in faecal samples taken in summer and spring (62.2%; 751/1208) as compared to autumn and winter (37.7%; 455/1208) (Kwan et al., 2008). In contrast, Hannon *et al.* (2009), reported non-significant differences in the prevalence of *C. jejuni* in summer (70% [95% CI=67-72%]; 963/1376) and winter (64% [95% CI=58-70%]; 896/14000) in feedlot cattle in Canada.

Another factor to consider is the health status of cattle. The frequency of *C. jejuni* isolation was not significantly different between healthy cattle and cattle with diarrhoea (Chatre et al., 2010; Ramonaite et al., 2013). An Austrian study that evaluated faeces from diarrhoeic and healthy calves also reported non-significant differences, although this study found more *Campylobacter* spp. shedding in diarrhoeic calves (Klein et al., 2013).

Many studies have reported that farm management factors such as feeding regime, cattle kept indoors or outdoors, and cattle density can influence *Campylobacter* prevalence. For example, it was reported that dairy cattle fed with whole cotton, cottonseed and alfalfa were more likely to shed *C. jejuni* than cattle not fed these types of feed (Wesley et al., 2000). Feeding grain to feedlot cattle was also reported to increase *Campylobacter* shedding (Garcia et al., 1985). Gregory *et al.* (2000) observed that pasture-fed animals shed more liquid faeces than hay-fed animals, and the authors postulated that soiling of fresh faecal matter on the hide possibly spreads pathogens between cattle or to its environment. In one study, *C. jejuni* were detected at greater frequency in faeces of cattle that were raised indoors than outdoors i.e. on pasture (Ellis-Iversen et al., 2009), and postulated that close proximity of cattle in indoor environments facilitates transfer of the bacteria among them (Weijtens et al., 1999; Minihan et al., 2004a). In contrast, Grove-White *et al.* (2010) reported that cattle kept outdoors are more likely to excrete *C. jejuni* than those kept indoors, probably due to a greater exposure of outdoor cows to environmental sources of *C. jejuni*. A study conducted in New Zealand did not find any differences in the *C. jejuni* prevalence between dairy herds managed under a housing system (herd home or stand-off pads) combined with outdoor pasture grazing and herds managed without a housing system (outdoor pasture only) (Rapp et al., 2014). Animals were kept indoor for 2-18 hours per day in the latter study, which may have influenced their findings. Increases in cattle density were also found to be associated with increased prevalence of *Campylobacter* spp. in dairy herds (Minihan et al., 2004; Grove-White et al., 2010).

Campylobacter spp. have been reported in all age groups of cattle, but higher prevalences have been reported in young calves compared to adult cattle. One Danish study found *Campylobacter* spp. in 42.1% of 107 samples from calves <4 months (45/107) of age and in 9.2% of 120 samples from adult cows (Nielsen et al., 2002), while a Japanese study reported higher *Campylobacter* prevalence in calves that were <1-year-old (97%; 32/34) compared to those which were >1-year-old (46.7%; 28/60) (Giacoboni et al., 1993). Calves can become colonised with *Campylobacter* spp. soon after the birth, thus, their prevalences differ widely between different age groups of cattle (Chatre et al., 2010; Ramonaite et al., 2013). Nonetheless, *C. jejuni* were detected at a significantly lower rate in adult cattle than in young calves, implying that calves near waterways pose a higher risk for water contamination (Chatre et al., 2010; Ramonaite et al., 2013).

Campylobacter jejuni prevalence in cattle is reported to be greater than *C. coli.* In a Canadian study, *C. jejuni* and *C. coli* were isolated in 50% and 1% of 100 slaughtered beef cattle sampled, respectively (Garcia et al., 1985). Similarly, a US study reported *C. jejuni* and *C. coli* in 38% and 1.8% of faecal samples from 2085 dairy cattle farm, respectively (Wesley et al., 2000). A French study that collected 2255 samples of faeces between 2002 and 2006 from healthy cattle for slaughter also reported lower *C. jejuni* (12.8%) and higher *C. coli* (3.7%) prevalence compared to previous studies (Chatre et al., 2010). Another cross-sectional study conducted in cattle in England and Wales found prevalences of *C. jejuni* (25.5%; 60/235) higher than *C. coli* (2.12%; 5/235) (Ellis-Iversen et al., 2009). These studies likely suggest that cattle could not be substantial contributors in *C. coli* contamination into waterways as compared to *C. jejuni*.

1.2.2.1.2 Sheep, goats and other ruminants as reservoirs of Campylobacter spp.

Campylobacter is less frequently observed in sheep faeces than in cattle, even though retail sheep carcases are considered as potential vehicles for *Campylobacter* infection sources (Stanley *et al.,* 1998; Wong et al., 2006; Lazou et al., 2014). *Campylobacter* prevalences in sheep faeces are reported to be between 0-58% and are more commonly detected in lambs than adults. (Stanley et al., 1998; Jones et al., 1999; Kassa et al., 2007; Oporto et al., 2007; Moriarty et al., 2011a; Sproston et al., 2011; Chanyalew et al., 2013; Pao et al., 2014). The true seasonal variation of *Campylobacter* shedding and prevalence was reported to exist in sheep, with the highest prevalence reported during the summer months (Stanley et al., 1998; Sproston et al., 2011). John *et al.* (1999) reported the lowest *Campylobacter* shedding (0%) when sheep were fed hay and silage compared with feeding on grazing pasture. It was postulated that increased stress could cause ewes to shed *Campylobacter* spp. three days after lambing (Mawdsley et al., 1995), and lambs may acquire infection within a few days (~ 5days) (Jones et al., 1999). Therefore, peaks of campylobacteriosis outbreaks in sheep can be observed around lambing the highest rates of shedding (100%) were reported in lambs stressed by weaning and movement onto new pasture. *C. coli* prevalence in sheep was often reported to be greater than in cattle. For example, Grove-White *et al.* (2010) reported 47.4 and 4.4% *C. coli* from 1720 and 7779 isolates originated from sheep and cattle faeces, respectively. Sproston *et al.* (2011) also identified 59.4 and 2.3% *C. coli* in 30 and 74 isolates originated from sheep and cattle faeces, respectively. Hence, although many studies reported lower *Campylobacter* spp. prevalence in sheep

than in cattle, the role of sheep in water contamination should not be underestimated because of a substantial presence of *C. jejuni/coli* in sheep faeces.

Although *Campylobacter* spp. have been reported in the meat of goats and buffalo, there is little information available on *Campylobacter* spp. prevalences in faeces of goats and other farmed ruminants such as buffalo and deer. Adesiyun *et al.* (1992) reported *Campylobacter* spp. in 4/10 diarrhoeic and in 3/8 non-diarrhoeic freshly voided faecal samples from kids below 6 months of age. Stone *et al.* (2013) reported that 3.7% of 252 rectal swabs from goats of Grenada were *Campylobacter* positive, with 3.1% being *C. jejuni*. Conversely, a high proportion of *C. jejuni* was reported in rectal swabs of goats in rural Ghana (33.3%; 24/72) and in freshly voided faeces of goats in Venda region of South Africa (17.5%; 35/200) (Abrahams et al., 1990; Uaboi-Egbenni et al., 2011). These high proportions in goats could be attributable to contact with other animal species such as pigs and poultry that are high *Campylobacter* shedders (Corte et al., 2006; Rosef et al., 1983). *Campylobacter* spp. were not detected in faecal samples from red deer, wild boar and other ungulates in a Spanish study, nor in samples from mule deer in a Canadian study (Díaz-Sánchez et al., 2013; Van Donkersgoed et al., 1990). *C. hyointestinalis* (6%; 24/399), however, was reported in faecal contents from reindeer in six slaughterhouses of Finland (Hänninen et al., 2001). One Canadian study detected *C. jejuni* in 6 of 7 faecal samples of buffalo collected from Oldman River watershed at Alberta (Jokinen et al., 2011) A recent a recent study conducted in Cambodia, however, detected no *Campylobacter* spp. in 25 faecal swabs from water buffalo, when cultured and PCR tested (Osbjer et al., 2016). Therefore, as there is a scarcity of studies on *Campylobacter* spp. shedding in faeces of goats, deer and buffaloes, understanding the role of these ruminants in *Campylobacter* spp. transmission needs further investigation.

1.3 *Cryptosporidium* spp. in humans and ruminants

Earnest Edward Tyzzer (1875-1965), an American parasitologist, was the first to identify and describe the genus *Cryptosporidium* in 1910, after the frequent detection of the parasite in the gastric glands of domesticated mice. Tyzzer named this parasite *Cryptosporidium muris*. He also found another species, *Cryptosporidium parvum,* in the small intestine of the mice in 1912. However, *Cryptosporidium* was not considered as an important pathogen until Slavin (1955) associated *Cryptosporidium meleagridis* with illness and death in turkeys, and Panciera (1971) reported that *Cryptosporidium* spp. were associated with diarrhoea in cattle. Subsequently, two groups reported the

first human cryptosporidiosis cases in 1976 (Meisel et al., 1976; Nime et al., 1976). Currently, *Cryptosporidium*, protozoan parasites, have been isolated from a wide range of hosts including mammals, amphibians, reptiles, fish, and birds. Although there is a lack of consensus in *Cryptosporidium* taxonomy, 30 valid *Cryptosporidium* species have been recognised to date (Slapeta, 2013). Among these species, *C. parvum* and *C. hominis* are the most important species implicated in human and animal cases of cryptosporidiosis (Fayer and Xiao, 2008). The list of major *Cryptosporidium* species, their hosts and their public health importance is shown in Appendix A.

1.3.1 Cryptosporidiosis in humans

Historically, the first cases of reported cryptosporidiosis were from a 3-year-old child and a 39-year-old immunosuppressed patient who were living on a farm with cattle and a dog, and from 9-year-old boy and a 52-year-old man with immunosuppressive conditions who were not in contact with animals (Nime et al., 1976; Meisel et al., 1976; Lasser et al., 1979; Weisburger et al., 1979). However, cryptosporidiosis in humans was not studied in greater detail until the Centers for Disease Control and Prevention (CDC) in the United States reported severe diarrhoea, due to *Cryptosporidium* infection, in 21 male patients who had Acquired Immune Deficiency Syndrome AIDS (Anonymous, 1982). Thereafter, many studies have confirmed *Cryptosporidium* spp. as the cause of diarrhoea in immunocompromised people (Baxby et al., 1983; Current et al., 1983; Ma and Soave, 1983; Tzipori, 1983a; Soave et al., 1984). At present, cryptosporidiosis is identified worldwide and is ranked as one of the most important zoonotic pathogens by public health agencies (ESR, 2014; OECD, 2014; CDC, 2014).

Cryptosporidium oocysts are small, oval or round-shaped, contain four parallel sporozoites, and are mostly morphologically indistinguishable from each other when observed by light microscopy (Fayer and Xiao, 2008). Infected hosts excrete environmentally resistant oocysts that can contaminate the, including water. Thus, *Cryptosporidium* infections can occur through the ingestion of contaminated drinking and recreational water, through food, or via direct contact with infected animals or persons (Tzipori and Ward, 2002; Fayer and Xiao, 2008; Caccio & Putignani, 2014). The infective dose of *C. parvum* is relatively low, with a median infective dose (ID_{50}: the extrapolated dose required to infect 50% of the test subjects) ranging from 10 to 132 oocysts (Du Pont et al., 1995; Fayer and Xiao, 2008). Once infected, sporozoites in the oocyst release and invade the microvillus border of epithelial cells in the intestine, following which villous atrophy and clinical symptoms may appear within a week on

average (Tzipori and Ward, 2002; Fayer and Xiao, 2008; Caccio & Putignani, 2014; Shikani and Weiss, 2014). The reported incubation period ranges from 3 to 22 days. The most common clinical signs of cryptosporidiosis in humans are self-limiting watery diarrhoea, abdominal cramps, vomiting and weight loss, but immunocompromised people can become chronically affected (Shikani and Weiss, 2014). Duration of the symptoms can range from 1 to 10 days (Jokipii et al., 1983; Insulander et al., 2005). Occasionally, watery diarrhoea can persist for a month, and intermittent diarrhoea was reported for up to 36 months in a Swedish study (Chalmers and Davies, 2010; Insulander et al., 2013). Infected people can shed oocysts for up to two months after symptoms cease (Jokipii and Jokipii, 1986; Chalmers and Davies, 2010). It has been reported that persons infected with *Cryptosporidium hominis* can shed higher numbers of oocysts for longer periods (mean days: 13.9) compared with *C. parvum*-infected persons (mean days: 6.4) (McLauchlin et al., 1999; Xiao et al., 2001; Cama et al., 2008).

In Nordic countries, a 1% *Cryptosporidium* prevalence was reported in asymptomatic and 3% in diarrhoeic individuals (Horman et al., 2004). Another coprological survey showed 0-2% prevalence in asymptomatic individuals and 0.1-27.1% prevalence in diarrhoeic patients (O'Donoghue, 1995). On the contrary, *Cryptosporidium* prevalences in developing countries vary from 0-31.6% in asymptomatic individuals and 0.1-31.5% in diarrhoeic patients (O'Donoghue, 1995; Esteban et al., 1998). These reports imply that individuals in developing countries could be carriers of *Cryptosporidium* without showing signs of overt disease. *Cryptosporidium* infection, particularly in children <5 years old, is considered one of the major causes of persistent diarrhoea in developing countries (Amadi et al., 2002; White, 2010) and is strongly associated with malnutrition, stunted growth and even infant death (Molbak et al., 1997; Amadi et al., 2001; Hunter and Nichols, 2002). Many AIDS patients have chronic and fulminant cryptosporidiosis and are associated with high morbidity and mortality in the lack of combination antiretroviral therapy (cART) (Farthing, 2000). Therefore, decreased immunity, malnourishment and poor hygiene could explain the differences in *Cryptosporidium* infection rates between developed and developing countries.

Cryptosporidium infection is a debilitating disease in immunosuppressed people, and *Cryptosporidium* is found spread throughout the gastrointestinal tract, in the biliary system epithelium and pancreatic duct, and as far as the respiratory tract within the host (Berk et al., 1984; Clavel et al., 1996; Farthing, 2000; Shikani and Weiss, 2014). Other cryptosporidiosis complications reported in immunocompromised persons,

particularly in cases of AIDS, were pneumatosis cystoides intestinales, oesophageal damage, appendicitis, and gastric cryptosporidiosis (Oberhuber et al., 1991; Ventura et al., 1997; Hunter and Nichols, 2002). Ideal treatment regimens for cryptosporidiosis are still unavailable, although some therapeutic agents such as salicylamide, derivatives of ntirothiazole and nitazoxanide and synthetic quinazolinone halofuginone lactate, have been used to reduce the duration of diarrhoea (Amadi et al., 2002; Trotz-Williams et al., 2011; Al Mawly et al., 2013; Shirley et al., 2012). Studies found heavy rainfall correlated with high *Cryptosporidium* infection, as rainfall might contaminate the source of drinking water with sewage and animal waste (Moodley et al., 1991; Adegbola et al., 1994; Atherholt et al., 1998; Curriero et al., 2001; Lal et al., 2013). Therefore, prevention of contamination of water and food is crucial for reducing the *Cryptosporidium* infection burden in both immunocompetent and immunosuppressed populations.

1.3.2 Cryptosporidiosis in farmed ruminants

Cryptosporidium was first considered as a cause of neonatal calf diarrhoea in 1970 (Panciera et al., 1971), and was recognised as a primary pathogen in neonatal diarrhoea only in 1980 (Tzipori et al., 1980). Baker and Carbonell (1974) and Mason *et al.* (1981) first described cryptosporidiosis in diarrhoeic lambs and 2-week-old kids in Australia. Currently, cryptosporidiosis has been reported in ruminants worldwide (Santin et al., 2008; Maurya et al., 2013). Ruminants can shed between 10^3 and 10^9 *Cryptosporidium parvum* oocysts per gram of faeces, even without showing any clinical signs of infection (O'handley et al., 2002; Fayer and Xiao, 2008). Therefore, healthy adults and young carriers can serve as sources of infection for other animals. Clinically, *C. parvum* infections in calves aged 1-3 weeks are characterised by acute onset of profuse watery diarrhoea, that is sometimes accompanied by depression, weakness, anorexia and dehydration (Tzipori et al., 1983b; Schnyder et al., 2009). The infection usually lasts 1-2 weeks, but in some cases, the severity and duration of diarrhoea may lead to death (Sanford and Josephson, 1982; Tzipori et al., 1983b; Fayer et al., 1998). Clinical signs related to *Cryptosporidium bovis*, *Cryptosporidium ryanae* and *Cryptosporidium andersoni* have not been reported in cattle (Fayer et al., 2005; 2008). Maldigestion and reduced in weight gain in young stock, and reduced milk production in cows have, however, been reported in *C. andersoni* infections (Anderson 1998; Lindsay et al., 2000; Ralston et al., 2010).

As in cattle, *C. parvum* infections in lambs cause mild to severe neonatal diarrhoea and have been associated with high morbidity and mortality (Angus et al., 1982; Sari et al., 2009). Retarded growth and low dressing percentage are other observed effects of cryptosporidiosis in lambs (Sweeny et al., 2011). In sheep, *C. parvum, Cryptosporidium xiao,* and *Cryptosporidium ubiquitum* are believed to cause diarrhoea, although clinical signs were not observed in lambs experimentally infected with *C. xiao* and *C. ubiquitum* (Navarro-i-Martinez et al., 2007; Fayer and Santin 2009; Diaz et al., 2010). Similarly, cryptosporidiosis in goatherds can reach 100% morbidity and 50% mortality, implying a high economic cost (Munoz et al., 1996; Johnson et al., 1999; Sevinç et al., 2005; Santin, 2013). *Cryptosporidium*-infected kids have shown anorexia, prostration and diarrhoea (Munoz et al., 1996; Paraud et al., 2010; Santin, 2013). Recently, calf diarrhoea in water buffalo was reported in association with *Cryptosporidium* infection (Bhat et al., 2012; Diaz de Ramrez et al., 2012; Maurya et al., 2013). Although *C. parvum* was found to be associated with diarrhoea in goats and buffalo, significant molecular studies are warranted on species-specific cryptosporidiosis in those species, as they represent major sources of food in developing countries where cryptosporidiosis-related diarrhoea is much more common (WHO, 2004; 2012).

1.3.2.1 Occurrences of *Cryptosporidium* in farmed ruminants

Ruminant farming, particularly of cattle and sheep, is the economic backbone of many countries including New Zealand. Infections with *Cryptosporidium* parasites are prevalent in ruminants worldwide. Acute diarrhoea in ruminants has negative impacts on the economy. In addition, infected ruminants are important reservoirs and amplifiers of *Cryptosporidium*, in particular, the zoonotic species *C. parvum*. Therefore, a better understanding of the epidemiology of *Cryptosporidium* spp. in farmed ruminants and their zoonotic potential is required to control cryptosporidiosis in animals and humans.

1.3.2.1.1 Cattle as reservoir of Cryptosporidium species

Cattle, particularly calves, are recognised reservoirs of *Cryptosporidium.* Numerous studies have, however, reported disparate prevalences at the animal level, ranging between 0% and 100% in cattle populations (Santin et al., 2004; Fayer et al., 2006; 2010; Plutzer and Karanis, 2007a; Thompson et al., 2007; Santin and Trout, 2008; Karanis et al., 2010; Khan et al., 2010; Smith et al., 2010; Castro-Hermida et al., 2011; Kvack et al., 2011; Meireles et al., 2011; Muhid et al., 2011; Wang et al., 2011; Chen and Huang et al., 2012; Ng et al., 2011; Budu-Amaoako et al., 2012a; 2012b; Helmy et al.,

2013; Maurya et al., 2013), perhaps due to the different locations, study design and diagnostic methods employed (Feng and Xiao, 2011). Often, studies have reported point prevalence (single sample per animal collected) of *Cryptosporidium*, and only a few have been studied for *Cryptosporidium* prevalence over time (multiple samples from the same animal collected), which showed that at some point all the cattle within a herd can be infected. Cattle are intermittent shedders of *Cryptosporidium* and have a short patent period (Santin et al., 2008). Therefore, point prevalence studies may underestimate the true prevalence in a population. There may also be distortions in the estimates because of the sensitivity of the identification methods used to determine the differences in the observed *Cryptosporidium* prevalences. These may occur because *Cryptosporidium* oocysts are small and may easily be missed during microscopy. Direct molecular testing of samples, on the other hand, may overestimate the prevalences, due to amplification of naked DNA (Lorenzo et al., 1993; Scott et al., 1995; Fayer et al., 2000). Despite these factors, numerous studies have reported *Cryptosporidium* prevalences between 0 and 100% at the herd level (Chang'a et al., 2011; Grinberg et al., 2005; Olson et al., 1997; Santin et al., 2004). Muhid *et al.* (2011) showed higher *Cryptosporidium* prevalence in calves on intensive farms (31.7%; 38/120) compared to those on semi-intensive farms (22.5%; 27/120). Another study in central Spain reported *Cryptosporidium* prevalences of 45.9% (387/844) and 50.4% (425/844) in intensive and semi-intensive systems, respectively (Castro-Hermida et al., 2002). High *Cryprosporidium* prevalence is not necessarily able to be attributed to whether a farming system is intensive or semi-intensive. Instead, poor hygiene could have increased *Cryptosporidium* prevalence, because studies have reported that calves in pens that were not disinfected were at a higher risk of infection than calves in periodically disinfected pens (Castro-Hermida et al., 2002; Hamnes et al., 2006; Muhid et al., 2011). In addition, *Cryptosporidium* spp. infection in calves increased if calves were kept on a sand floor, on slatted floors, on a bedding of 0-5 cm depth, and if calves were fed with saleable milk (Mohammed et al., 1999; Brook et al., 2008; Muhid et al., 2011).

Another important factor in the understanding of *Cryptosporidium* infection epidemiology is the age of cattle. Brook *et al.* (2008) collected 215 faecal samples between April and May 2004 from unweaned calves on 41 farms within a 10 km x 10 km area of Cheshire, UK. They found *Cryptosporidium* spp. in 28% of samples from 66% herd, with calves between 8 and 21 days of age being 5.24 times more likely to be infected than calves aged between 0 and 7 days.. Satin *et al.* (2008) investigated 30 calves in a purebred Holstein dairy farm in Maryland, the USA from birth to 24 months

of age, and collected 33 faecal samples from each calf. They reported the highest prevalence of *Cryptosporidium* in pre-weaned calves (<8-weeks-old; 45.8%) followed by post-weaned calves (3-12 months; 18.5%). The majority of these *Cryptosporidium* detections (96.6%) were found in 2-week-old preweaned calves. A Danish study also showed a *Cryptosporidium* prevalence of 96, 84, and 14% in young calves, older calves, and cows, respectively (Maddox-Hyttel et al., 2006). Many other studies also provided data on the age-related *Cryptosporidium* prevalence patterns (Xiao and Herd, 1994; Sischo et al., 2000; Sturdee et al., 2003; Olson et al., 2004; Santin et al., 2004; Kvac et al., 2006), but there are many hypotheses behind the correlation of age and *Cryptosporidium* prevalence. However, these studies indicated that the calving to weaning period is an important period in which to target the reduction of the *Cryptosporidium* burden in a farm environment, and thereby reduce the *Cryptosporidium* infection risk to people.

In addition, age-related patterns have also been observed in the distribution of *Cryptosporidium* species. Primarily, *C. parvum* is more prevalent in pre-weaned calves and *Cryptosporidium bovis* and *Cryptosporidium rynae* in post-weaned calves, whereas in adult cattle *Cryptosporidium andersoni* is more prevalent (Santin et al., 2004; Fayer et al., 2007; Plutzer and Karanis 2007; Santin et al., 2008; Brook et al., 2009). However, a number of studies also reported *C. bovis* dominating in –pre-weaned and *C. andersoni* in pre-weaned and pos-tweaned calves (Thompson et al., 2007; Silverlas et al., 2010; Muhid et al., 2011; Wang et al., 2011; Budu-Amoako et al., 2012a; 2012b), and a few studies have reported no significant correlation between the calf age and the species of *Cryptosporidium* detected (Geurden et al., 2008; Winkworth et al., 2008). Some studies have even reported mixed infections of two or more *Cryptosporidium* spp., between 1-17% of cattle samples analysed (Thompson et al., 2007; Muhid et al., 2011; Murakoshi et al., 2012; Ng et al., 2012; Helmy et al., 2013; Shrestha et al., 2014). All these *Cryptosporidium* spp. are morphologically similar but differ genetically, and appear to be linked to different clinical manifestations in cattle. For example, *C. parvum* has been identified as one of the causes of diarrhoea in calves, whereas *C. andersoni* has been implicated as a cause of reduced milk production. Therefore, although the clinical effects of many *Cryptosporidium* spp. on humans have not been defined, findings of different species of *Cryptosporidium* in different age groups of cattle added valuable information to understand the transmission of cryptosporidiosis in cattle.

Many studies have reported subtypes of *C. parvum* in cattle samples based on sequence analysis of the 60-kDa glycoprotein (gp60) gene, and have identified two families of

subtypes, IIa and IId, in cattle (Xiao, 2010; Feng et al., 2007; Ng et al., 2011; Helmy et al., 2013; Shrestha et al., 2014; Al Mawly et al., 2015). The most common subtype found in calves worldwide is IIaA15G2R1 (Alves et al., 2003; 2006; Feng et al., 2007; Xiao et al., 2007). However, some countries such as Australia and New Zealand have reported a higher prevalence of subtype IIaA18G3R1 (Thompson et al., 2007; Xiao, 2010; Ng et al., 2011). Both IIaA15G2R1 and IIaA18G3R1 subtypes and families IIa and IId have been reported in humans, suggesting that subtyping of *Cryptosporidium* isolates is important to determine potential zoonotic transmission.

1.3.2.1.2 Sheep, goats and other ruminants as reservoir of Cryptosporidium

Studies of *Cryptosporidium* infections in sheep, goats and deer are lower in number and have not been well described in comparison to cattle. A review that summarises the published survey data since 1989 to 2009 showed that *Cryptosporidium* prevalence ranged from <5% to >70% (mean~30%, n=20) in sheep and from<5% to >35% (mean ~15%, *n*=11) in goats (Roberston, 2009). A few recent *Cryptosporidium* studies also showed prevalences between 2.2 and 78% in sheep (Koinari et al., 2014; Yang et al., 2014; 2015; Wells et al., 2015; Mirhashemi et al., 2016) and between 3.8 and 16.5% in goats (Koinari et al., 2014; Wang et al., 2014; Mi et al., 2014; Peng et al., 2016). The wide variation in prevalences observed could be due to differences in sample collection variables such as age of animal, selection of diarrhoeic animals, individual or pooled samples etc., and in sample analysis methods used such as microscopy of wet mounts, or microscopy after concentration of samples or after immunofluorescent antibody staining (IFA), or PCR.

As in cattle, *Cryptosporidium* spp. are more prevalent in lambs and kids compared with adult animals. In a Belgian study, 13.1% (18/137) lambs of ≤10 weeks of age from 4/10 herds and 9.5% (14/148) kids of ≤10 weeks of age from 6/10 herds were found infected with *Cryptosporidium* (Geurden et al., 2008). In an Australian study of *Cryptosporidium* infection in 2-week- old to8-month-old lambs (n=235) from two herds, an 18.5-42.6% infection prevalence was reported (Sweeny et al., 2011). Similarly, Wang *et al.* (2010) showed 10.8% and 4.3% prevalences of *Cryptosporidium* spp. in pre-weaned (n=378) and post-weaned (n=585) lambs respectively. Delafosse *et al.* (2006) reported 16.2% (142/879) *Cryptosporidium* prevalence in kids aged between 5 and 30 days from France. Conversely, Maurya *et al.* (2013) found low *Cryptosporidium* spp. prevalence (3.5%, 4/116) in kids of ≤3-months-old from India. Although it is postulated that with increasing age the sheep gains immunity to *Cryptosporidium* infections, the variation in prevalence could be due to sample analysis

methods used for detection of *Cryptosporidium*, sample size and study designs of the studies (Robertson et al., 2014), and geographical differences in prevalences of *Cryptosporidium*. Therefore, more studies with better epidemiologic designs are needed to understand the transmission of *Cryptosporidium* spp. in both sheep and goats.

In sheep, several species and genotypes have been recognised. However, the most prevalent species are *C. parvum, C. ubiquitum* and *C. xiao* (Fayer and Xiao, 2008). Other species identified are *C. scrofarum, C.* sheep genotype I, *C. andersoni, C. hominis, C. suis* and *C. fayeri* (Fayer and Xiao, 2008). Geurden *et al.* (2008) and Robertson *et al.* (2010) reported a high prevalence of *C. ubiquitum* in young lambs, whereas *C. parvum* was found more frequently in lambs in Australia, Italy, and Romania (Yang et al., 2009; Paoletti et al., 2009; Imre et al., 2013). The possible reasons for this variation are not known and warrant more studies. Studies have also reported *C. bovis* in sheep (Soltane et al., 2007; Wang et al., 2010). Only *C. parvum* and *C. xiao* have been reported to be prevalent in goats (Fayer and Xiao, 2008). However, only a limited number of studies have been conducted in goats, making the species distribution inconclusive. As in cattle, potential zoonotic *C. parvum* gp60 subtype families, IIa and IId were also reported in sheep and goats. Mallon *et al.* (2003a) found sheep subtypes clustered with human and cattle isolates, indicating possible zoonotic transmission. As *C. ubiquitum* has been reported in sporadic human cases, storm water, raw water and drinking water, sheep could be a contributor for potentially zoonotic *Cryptosporidium* contamination in water and related waterborne outbreaks.

Studies on *Cryptosporidium* infection in buffalo are relatively recent. Studies suggested 3-38% *Cryptosporidium* prevalences in buffalo with the presence of *C. parvum, C. rynae, C. bovis,* and *C. ubiquitum* (Fayer and Xiao, 2008; Robertson et al., 2014). Because those studies used calves, *Cryptosporidium* species distribution across different age groups of buffalo is unknown. Like in cattle, *C. parvum* is prevalent in buffalo calves, and these belong to subtypes of families IIa and IId (Fayer and Xiao, 2008; Robertson et al., 2014). *Cryptosporidium* species studies conducted on wild deer indicated prevalences ranging between 1.3-12.5% (Castro-Hermida et al., 2011a; 2011b; Robinson et al., 2011; Santin and Fayer, 2015), whereas diarrhoea associated with *Cryptosporidium* was reported in farmed deer (Tzipori et al., 1981; Orr et al., 1985). Recently, studies have reported *C. parvum, C. bovis*, *C. rynae*, and *C. ubiquitum* (Garcia-Presedo et al., 2013; Santin and Fayer, 2014) in wild deer, indicating that deer could be a reservoir of

zoonotic *Cryptosporidium* spp. However, there is scarce data on *Cryptosporidium* in farmed deer and this warrants detailed investigations.

1.4 *Giardia* spp. in humans and ruminants

Krunstler established the generic name *Giardia* in 1882, to describe the flagellate found in the tadpole, although Antony van Leeuwenhoek, a biologist and clinician, discovered *Giardia* in 1681 and Lambl first described it in 1859 (as cited in Thompson and Monis, 2012). *Giardia* species are morphologically indistinguishable from each other; therefore, Filice (1952) proposed a trophozoite morphological characters- based system to re-evaluate *Giardia* species (as cited in Thompson and Monis, 2012). On a morphological basis, currently, *Giardia* consists of six valid species. Of the six species, *G. duodenalis* (syn. *Giardia intestinalis*, *Giardia lamblia*) is known to infect both humans and animals (Feng and Xiao, 2011). Furthermore, *G. duodenalis* consists of eight distinct genetic groups or assemblages (A through H) based on DNA polymorphism (Monis et al., 2003; Caccio and Ryan, 2008; Ryan and Caccio, 2013). Recently, a species name has been proposed for *G. duodenalis* assemblages A, B and E after whole genome sequence analysis of them, but it has not yet been validated (Franzen et al., 2009; Jerlstrom-Hultqvist et al., 2010). There is also genetic variation within the assemblages, and a number of sub-assemblages have been identified (identified by Roman numerals as suffixes) (Monis et al., 2003; Read et al., 2004; Traub et al., 2005). A list of currently recognised *Giardia* species, assemblages and sub-assemblages are listed in Appendix A. So far, assemblages A and B have been found to infect humans (Homan et al., 1998; Read et al., 2004).

1.4.1 Giardiasis in humans

Approximately 200 million human giardiasis cases have been reported annually worldwide (Lane and Lloyd, 2002; Yason and Rivera, 2007). The infected hosts excrete egg-like cells called 'cysts', which are immediately infectious upon excretion (Fayer and Xiao, 2011). However, the clinical effect of *G. duodenalis* infection depends on the immune status of the host and the degree of exposure (number of cysts consumed) of the host to cysts. Adult volunteer feeding trials showed an infective dose (ID_{50}) of 50 *G. duodenalis* cysts (Hibler et al., 1987), whereas Rose and Gerba (1991) determined an ID_{50} of 35 cysts based on a dose-response curve. In addition, the ID_{50} of *G. duodenalis* could be <35 if the cyst source is from humans (Rendtorff, 1978). The faecal-oral route is the commonest path for *Giardia* transmission. Person-to-person *Giardia*

transmission has been well documented, but there is also evidence of environmental transmission, particularly through contaminated water (Fraser et al., 2000; Hoque et al., 2002; Craun et al., 2010). Recent evidence of the occurrence of assemblages A and B in humans, canines, felines and cattle suggested the possibility of zoonotic transmission of *Giardia duodenalis*. The mechanism by which *Giardia* causes illness is not well understood, and no specific virulence factors have been identified. However, it is believed that trophozoites (the asexual phase) colonise the small intestine, reduce the absorptive surface area and cause electrolyte transport abnormalities, leading to increased intestinal motility (Feng and Xiao, 2011). Giardiasis is characterised by a spectrum of symptoms ranging from asymptomatic carriage to manifestations of acute or chronic diarrhoea, abdominal pain, nausea, vomiting and anorexia (Fraser et al., 2000; Eckman, 2002; Caccio and Ryan, 2008; Craun et al., 2010). Some people may develop lactose intolerance and malabsorptive syndrome as a consequence of giardiasis (Ryan and Caccio, 2013). Studies showed that a *Giardia*-infected person can show symptoms within 3-20 days of infection (average: one week), and can excrete up to 10^7 cysts/g of faeces (Danciger et al., 1975; Jokipii et al., 1985; Nash et al., 1987). In a Netherlands study, *Giardia duodenalis* assemblage A infections were characterised by mild intermittent diarrhoea, whereas assemblage B caused profuse diarrhoea, weight loss and fatigue (Homan et al., 2001). Read *et al.* (2002) reported more frequent diarrhoea in children infected with assemblage A than with assemblage B. In contrast to cryptosporidiosis, human giardiasis is pharmacologically treatable with antibiotics such as metronidazole, tinidazole, and paromomycin (Gardner and Hills, 2001; Pasupuleti et al., 2014; Watkin and Eckman, 2014).

Giardiasis has been reported worldwide and the reported *Giardia* prevalence is generally lower in developed countries (0.4-7.5%) than in developing countries (8-30%) (Feng and Xiao, 2011). Poor sanitation and contaminated water supplies are believed to be responsible for the higher prevalence in developing countries. Like *Cryptosporidium,* immunocompromised people, children and the elderly are more vulnerable to *Giardia* infection (Dwivedi et al., 2007). Stark *et al.* (2009) reviewed and reported a 1.5-17.7% prevalence of *G. duodenalis* in HIV-infected individuals, although giardiasis is not considered a major cause of enteritis in HIV patients. However, there are frequent reports of *G. duodenalis* cysts in the stools of hypogammaglobulinemic patients with symptoms of chronic diarrhoea.

Case-control studies have reported that children <5- years- old are more vulnerable to infection with *Giardia* spp. than those >5- years- old (Espelage et al., 2010; Julio et al.,

2012). Seasonal variations are also evident, and in many countries, there is a higher incidence of human giardiasis in the warmer months (spring/summer) than in winter (Yoder and Beach, 2007; Yoder et al., 2012). In addition, people living in rural areas are at higher risk of *G. duodenalis* infection compared to the people in urban areas (Lujan and Svard, 2011). A retrospective case-control study in rural New England reported the household use of shallow water sources as the main risk factor for giardiasis, followed by foreign travel and attending day-care centres (Chute et al., 1987). Another case-control study in Italy also identified travelling abroad and exposure to surface water as the main risk factors for giardiasis (Faustini et al., 2006). A recent *G. duodenalis* assemblages-specific case-control study showed assemblages A and B being more prevalent in >15 and 15-44 -year -old people, respectively (Minetti et al., 2015). In addition, they also found a strong association of assemblage A with dog ownership whereas assemblage B were positively associated with contact with other people and severity of clinical symptoms. Therefore, it is expected that specific risk factor analysis of the assemblages could provide a better understanding of *Giardia* epidemiology in future.

1.4.2 Giardiasis in farmed ruminants

G. duodenalis infections in ruminants are not as severe as reported in *Cryptosporidium* infections. Yet, giardiasis has been described as an important enteric disease of ruminants due to its high prevalence, the clinical effect on young animals and production losses (O'Handley et al., 1999; Olson et al., 2004). Subsequent clinical signs vary widely between ruminant species and within species because the concentration of *G. duodenalis* cysts and host immunity play important roles in giardiasis. *Giardia duodenalis* has been frequently reported in both asymptomatic (23-34%) and diarrhoeic (7-29%) veal calves. Bjorkman *et al.* (2003) found *G. duodenalis* cysts in 21% (26/124) of healthy calves in Sweden and indicated that *G. duodenalis* can persist on the farm without clinical signs being apparent in animals. Consequently, direct contact with an infected host could be the potential route of transmission for *G. duodenalis* in ruminants. Clinically, *G. duodenalis* infected calves showed pasty to fluid faeces with mucus in diarrhoea (Constable, 2014). Experimentally infected calves showed acute to chronic diarrhoea (St. Jean, 1987, Geurden et al., 2006a; 2006b), and infected lambs excreted unformed faeces and showed reduced weight gain, impaired feed efficiency and decreased carcass weight (Olson et al., 1995; Ralston et al., 2003). In addition, diarrhoea in giardiasis does not respond to antibiotic or coccidiostatic

treatment. Therefore, giardiasis in ruminants could potentially influence production and the economy.

1.4.2.1 Occurrences of *G. duodenalis* in farmed ruminants

Giardia duodenalis infections have been detected in a wide range of ruminants, but most prevalence data are primarily available for cattle (O'Handley and Olson, 2006). Similar to *Cryptosporidium*, reported *Giardia duodenalis* prevalences in ruminants vary considerably, according to management style, location, the design of the study and diagnostic techniques. Therefore, *Giardia* occurrence in farmed ruminants should be investigated in detail to reduce the potential of zoonotic transmission. With the advancement in molecular techniques, *Giardia duodenalis* assemblages A and B (which are found in humans) have also been reported in farmed ruminants. Usually, cattle, sheep and goats are found infected with *Giardia duodenalis* assemblages E (Caccio and Ryan, 2008; Minetti et al., 2014).

1.4.2.1.1 Cattle as reservoir of Giardia duodenalis

A wide range of *Giardia duodenalis* prevalences in cattle has been reported worldwide. *Giardia* prevalences reported were between 2.12% (n=1366) in China (Huang et al., 2014) and up to 57.8% in Canada and Australia (O'Handley et al., 2000) at animal level, whereas at the farm level, reported prevalences were as high as 96-100% in Canada (Dixon et al., 2011; Budu-Amoako et al., 2012a), UK, and Germany (Geurden et al., 2012). In addition, at the farm level, cumulative prevalences of *Giardia* were reported as high as 100% in North American cattle (Xiao and Herd 1994; Ralston et al., 2003). These studies indicate both that the occurrence of *G. duodenalis* is widespread and that, at some point in time, all cattle within the farm could be infected.

Apparently, *G. duodenalis* prevalences also vary between ages in cattle (Xiao and Fayer 2008). In one USA study, *Giardia duodenalis* prevalence ranged from 9 to 93% in 1-7-week-old pre-weaned calves on 14 farms (Trout et al., 2004). *G. duodenalis* prevalence in Canada was also reported to be lower in calves (<6 months; 22.6% of 605) than adult cattle (>24 months; 17% of 605) (Gow and Waldner, 2006). In addition, analysis of risk factors associated with *Giardia* infection showed that young calves (<2 months) have a likelihood of higher prevalences than older cattle (Geurden et al., 2012). *Giardia* spp., therefore, have been reported with a greater frequency in calves <6 months- old than in calves >6 months of age (Geurden et al., 2010), implying that young calves could be potential reservoirs for *Giardia* contamination of farm environments.

In more recent years, molecular studies of *G. duodenalis* indicated that assemblage-specific assays[1] would better estimate the *G. duodenalis* prevalence than immunofluorescence assay (Caccio and Ryan, 2008; Geurden et al., 2012). In cattle, the host-specific assemblage E of *G. duodenalis* appears to be the most prevalent genotype (Thompson et al., 2010; Feng and Xiao, 2011; Fayer et al., 2012; Ryan and Caccio, 2013; Helmy et al., 2014a). However, in New Zealand, only assemblage B was found in cattle (Learmonth et al., 2003; Winkworth et al., 2008). Recently, assemblage B was also reported in 7/16 isolates from cattle in China, as determined by typing at the *tpi* locus (Liu et al., 2012). There are increasing reports of assemblage A in cattle, with reported prevalences of 28% in Italy, 29% in the UK, 41% in Germany and 61% in France (Geurden et al., 2012). Some studies have found a higher frequency of assemblage A (15%; 7/14 farms) in pre-weaned calves (<8 weeks) and of assemblage E (33% on 13/14 farms) in older cattle (Trout et al. 2004; 2005; 2006; 2007). Assemblage E is typically not zoonotic whereas assemblages A and B are considered zoonotic. Therefore, although public health risk of cattle-associated giardiasis is considered minimal, the findings of assemblages A and B should not be ignored. Further, frequent reports of mixed infections with assemblages A and E in cattle suggested that zoonotic *Giardia* assemblages might be transient in cattle (Fang and Xiao, 2011; Khan et al., 2011). Although uncommon, assemblages C, D and F have also been reported in cattle in the UK, as determined by typing at 18S locus (Minetti et al., 2013).

Genetic variability within assemblages has been reported, and therefore assemblages A and B have been divided into four sub-assemblages (AI, AII, AIII and AIV; and BI, BII, BIII and BIV) based on protein polymorphisms of 23 loci (Monis et al., 2003). In animals, subassemblages AI, AIII and AIV, and BI and BII have been reported. (Monis et al., 2003). Subtyping of assemblage A isolates showed subtype-AI (62%) as the most common subassemblage in cattle (Sprong et al., 2009). In India, genotyping at the bg locus of *G. duodenalis* isolates showed the presence of a similar subassemblage AI in both cattle and farm workers, and suggested zoonotic potential (Khan et al., 2011), because usually humans are reported to be infected with subassemblages AII (Xiao and Fayer, 2008). Therefore, cattle are possibly zoonotic reservoirs for *Giardia duodenalis*, and transmission could occur through contaminated water (Feng and Xiao, 2011; Budu-Amoako et al., 2012a).

[1] In vitro amplification of nucleic acid using PCR or related methodologies

1.4.2.1.2 Sheep, goats and other ruminants as reservoir of Giardia duodenalis

The reported *G. duodenalis* prevalences in sheep and goats in different studies ranged from 1.5 to 38% and 4 to 53%, respectively (Castro-Hermida et al., 2005; Bomfim et al., 2005; Santin et al., 2007; Geurden et al., 2008; Ruiz et al., 2008; Yang et al., 2009; Gomez-Munoz et al., 2009; 2012; Robertson et al., 2010; Zhang et al., 2012; Tzanidakis et al., 2014). In these studies, farm prevalences were between 45 and 100% on both sheep and goat farms. Only a handful of studies investigated the *Giardia duodenalis* occurrences in buffalo. Rinaldi *et al.* (2007) conducted a cross-sectional survey of *Giardia duodenalis* in 90 water buffalo farms from central Italy. They tested 347 faecal samples for copro-antigens using enzyme-linked immunosorbent assays (cELISA) and reported 18.1% animal prevalences and 30% farm prevalences. Few studies have reported *Giardia* prevalences in wild deer, but until now, no *Giardia* surveys in farmed deer have been reported in the literature.

Similar to cattle, *Giardia duodenalis* assemblage E was commonly reported in sheep and goats, although assemblage A has also been found in some countries (Feng and Xiao, 2011; Jafari et al., 2012; Paz e Silva et al., 2014; Robertson et al., 2009). Sprong *et al.* (2009) showed 28/36 isolates belonging to the subassemblage AI and the remaining to AII subassemblage in goats, whereas Lebbad *et al.* (2010) found that all isolates from sheep belonged to the AI subassemblage. Mixed infection with assemblages A and E was also reported in sheep. The genotypes of *Giardia duodenalis* varied widely according to the locus typed (Gómez-Muñoz et al., 2012). At the 18S locus, assemblages A and E differed by a single nucleotide substitution, therefore, there is a need to genotype *Giardia duodenalis* isolates at loci other than 18S (Gómez-Muñoz et al., 2012). Caccio *et al.,* (2007) reported 6/8 assemblage E isolates from buffalo, and the remaining two were assemblage A when a β-giardin locus was used. An Australian study detected assemblage A in 56/476 (11.8%) and assemblage E in 6/476 (1.3%) samples (Abeywardena et al., 2013). In addition, assemblage B has also been reported in sheep and goats (Robertson et al., 2010; Zhang et al., 2012), suggesting that sheep, goats and buffalo could be potential reservoirs of zoonotic *Giardia duodenalis*, and may contaminate the farm environment including water.

1.5 Pathogens contamination of waterways

Fayer and Trout (2005) estimated that 10-100 billion tonnes of agricultural animal manure are generated annually on a global scale, and these animal wastes may contain

high concentrations of microorganisms including bacterial and protozoal pathogens (Mawdsley et al., 1995; Thurston-Enriquez et al., 2005). During heavy rainfall events, enteric bacteria and protozoa may be released from the faeces deposited on land, to flow through the water and contaminate surface waters used for drinking purposes, recreational swimming and bathing. However, the impact of the presence of pathogens such as *Campylobacter*, *Cryptosporidium* and *Giardia* in these biomasses on human and animal health may depend on how much faecal load is present in the environment, how these pathogens are transported into the water and how they survive there. These questions are discussed in brief in the following subsections.

1.5.1 Faecal loading rate of pathogens in the environment

Shedding of faeces on pasture contributes to the loading of pathogens in environments and to agricultural run-off. Therefore, the farm environment may act as a continuous reservoir of infection for humans. *Campylobacter*-positive ruminants can shed 10^2 CFU/g of organisms with a minority (<10%) potentially shedding from 10^6 to 10^8 CFU/g of faeces. *Cryptosporidium*-positive dairy calves aged between one and four weeks can shed 1×10^5 to 6×10^7 oocysts/g of faeces (Xiao & Herd, 1994; Atwill et al., 1998; Uga et al., 2000; Moore et al., 2003; Starkey et al. 2005; Santin et al., 2008). Assuming a mean body mass of a 1-to-4-week-old calf to be 40-60 kg and a mean daily faecal output of 3.3% of the body mass, a calf could shed a few billion to hundreds of billions of oocysts/day. Similarly, sheep and goats can shed between ~6800 and ~232,000 oocysts/g of faeces, respectively (Geurden et al., 2008). An adult sheep could generate 4×10^6 oocysts/day, assuming 0.7 kg faecal excretion/day. Ortega-Mora and Wright (1994) showed that the shedding intensity of *C. parvum* oocysts was higher in six-day-old lambs (2.2×10^9) compared with two –month- old lambs (2.5×10^7). Geurden *et al.* (2008) reported *G. duodenalis* infected sheep and goats shedding a mean of ~4600 and ~18,000 cysts/g faeces respectively. As a result, the environmental loading with *G. duodenalis* cysts would be 3.2×10^6 cysts/adult sheep/day, considering an adult sheep produces 0.7 kg faeces/day. Feedlot steers in the US were found to be shedding ~2030 cysts/g faeces, which could generate an environmental loading rate between 2×10^7 and 8×10^7 cysts/animal/day (assuming 10-40 kg faeces excreted). These reports show that high prevalence rates of pathogens in the farm may contribute to high intensities of pathogen loading into the environment, including water. However, animal husbandry practices such as calving pen hygiene, and manure management practices, pathogen viability and survival and transport of the pathogens

in the farm and water environment will influence water contamination and the potential for waterborne diseases.

1.5.2 Transport of pathogens to the waterways

The degree of pathogen contamination from land to water depends on the type of livestock production system and the hydrological system, topography and soil types. In extensive grazing systems like in New Zealand, cattle and sheep frequently have direct access to streams and rivers, thereby directly depositing pathogens and polluting the water environment over long periods of time (Wilcock, 2006; McKergow and Hudson, 2007). In intensive production systems as in the United States and the Netherlands, rainfall event transport the pathogens to surface water sources through surface runoff from land-applied manure, or after manure ponds leak (Davies-Colley et al., 2004). Therefore, there are possibilities of contamination of surface and ground water in both production systems.

1.5.2.1 Transport of pathogens to surface water

Runoff is the flow of excess water from rain, ice-melt water, or other sources, over the land. Runoff is a key mechanism for microbial contamination of surface water sources and is influenced by many factors, such as the occurrence of storms, their intensity and duration, land topography and soil types. A study conducted in the Toenepi stream of New Zealand estimated that 95% of the total faecal pollution is due to storm flows (Davies-Colley et al., 2008). There is also evidence of increased *Campylobacter* concentration during storm flows (Stott et al., 2011). Faecal contamination can drop in magnitude at low flows, but on the other hand, at high flows, some bacteria may be retained in wetlands, resulting in slow bacterial release to streams (Collins, 2002). High-velocity runoffs are often seen in the hill country, whereas rolling country generates moderate runoff, and flat country is less likely to generate runoff (Donninson and Ross, 2009; NIWA, 2006). One study showed 46 to 388 mm of subsurface runoff in flat country, implying that soil types irrespective of topography also influence runoff and pathogen transportation (Donninson and Ross, 2009; NIWA, 2006). Clay-rich soils have a low permeability and can promote surface runoff, whereas bare soil can contribute a substantial pathogen load to runoff (Rosen, 2000). *Campylobacter* spp. and other bacteria can run off from gley soil at least 28 days after deposition (Donninson and Ross, 2009). Artificial drainage such as mole and tile drains can rapidly transport microbes to waterways (Donninson and Ross, 2009). However, although relationships between subsequent rainfall events and microbial loads are inconsistent, it is desirable

to remove stock from paddocks near waterways during prolonged heavy rainfall events in order to reduce waterway contamination with faeces.

1.5.2.2 Transport of pathogens to ground water

Ruminant-associated pathogens may also contaminate groundwater, resulting in waterborne outbreaks. For example, in 2004, a massive groundwater-associated outbreak due to bacteria occurred in Ohio, USA, and affected 1450 people (Fong et al., 2007). Microbe transportation into groundwater is characterised by complex interactions between the organisms and physical (filtration/straining) and chemical (adsorption/desorption) processes in soil columns. For example, smaller microbes like viruses are more able to filter to the bottom of a soil column through fine-grained soils than larger protozoa and bacteria (Jamieson et al., 2004). Heavy rainfall likely increases the speed of pathogen transportation into groundwater compared with lighter rainfall. In addition, recent manure application or faecal deposition on coarse or sandy soil followed by heavy rainfall increases groundwater contamination risks (Bowman, 2009; USEPA, 2005). Shallow unconfined aquifers are more vulnerable to contamination than are deep confined aquifers, although the latter may be contaminated if cracks are present in the confining layer (Borchardt et al., 2007). During infiltration, viruses may be adsorbed within the first few inches of soil but rainfall cause desorption of viruses leading to more downward migration (Borchardt et al., 2007). Microbial cells with a negative surface charge and mobile microbes exhibit a greater downward migration, while high clay and iron oxyhydroxide and soil pH <7 increase the adsorption of the cells. In addition, unsealed or unprotected well-heads also lead to the groundwater contamination, like *Campylobacter* outbreaks in Walkerton, Canada (Hrudey et al., 2003).

1.5.3 Survival of pathogens in the environment

Ruminants may shed numerous pathogens in faeces, but these pathogens must also be able to survive in the soil, water and food in order to infect a new host. In this section, how *Campylobacter, Cryptosporidium,* and *Giardia* survive in the environment are discussed.

1.5.3.1 Survival of *Campylobacter* spp.

Campylobacter spp. growth outside the host has not been reported because these organisms are thermophillic and microaerophillic, and the presence of environmental stresses such as ambient temperature and UV radiation inhibit their replication (Jones,

2001; Rollins and Colwall, 1986; Sinton et al., 2007). In fact, thermotolerant *C. jejuni* can survive above ambient temperature but becomes inactive at or above 48 °C in acidic conditions (pH<5.0). Sinton *et al.* (2007) showed that 90% of *C. jejuni* cells in cattle faeces pats are inactivated within 6 days of deposition. Moriarty *et al.* (2011b) also reported 90% of *Campylobacter* spp. inactivation in sheep faeces on pasture, particularly in warmer weather, within 6 hours during summer.

Campylobacter spp. was found surviving much better at cooler water temperatures (<15 °C) (Cools et al., 2003; Gonzalez and Hanninen, 2012). *Campylobacter* was also found thriving in water in harsh environmental conditions (e.g. low nutrient- and osmotic-stress) by forming biofilms and entering a viable but non-culturable (VBNC) state (Rollins and Colwall, 1986; Murphy et al., 2006). The molecular mechanisms of *Campylobacter* biofilm formation are still not well understood, though it has been reported that the ability to form biofilms varies between *C. jejuni* strains (Bronowski et al., 2014; Buswell et al., 1998; Joshua et al., 2006). VBNC is the state at which bacteria are unable to grow in normal growth media and reduce their metabolic activity, but they can retain viability for recovery and infection under favourable conditions (Barrer and Harwood, 1999). Some studies suggested that VBNC state *Campylobacter* organisms, if consumed, could potentially infect and cause disease, whereas other studies disagreed on the VBNC infectivity (Baffone et al., 2006; Bronowski et al., 2014; Koenaard et al., 1997; Thomas et al., 1999; Ziprin and Harvey, 2004). Some studies reported that VBNC state varies between *C. jejuni* strains and could be the reason why certain *C. jejuni* strains are often reported from environmental sources (Cools et al., 2003; Lazaro et al., 1999; Medema et al., 1992; Tholozan et al., 1999). *C. jejuni* was also reported as persisting in the environment when co-occurring with other bacteria such as *Pseudomonas* that enhance biofilm formation, and when swallowed up by protozoa such as *Acanthoamoeba polyphagia* (Snelling et al., 2006; Sanders et al., 2007; Teh et al., 2010).

1.5.3.2 Survival of *Cryptosporidium* and *Giardia* spp.

Cryptosporidium and *Giardia* (oo)cysts are resilient and can persist in faeces for many years (King and Monis, 2007; Fayer and Xiao, 2008). At low temperatures and in the absence of freeze-thaw cycles, *Giardia* cysts may remain infective for up to a month at 1.1-7 °C, and *Cryptosporidium* oocysts up to 24 months at <15 °C (Fayer et al. 1998b; Jenkins et al. 2002; Robertson and Gjerde, 2006). The (oo)cysts are extremely susceptible to temperatures above 37 °C and lose their infectivity when incubated at 54 °C for 10 minutes (*Giardia* cysts) and at 64 °C for 5 minutes (*Cryptosporidium*

oocysts). In an experimental assay, *Giardia* cysts were found to survive up to 7 weeks in soil and up to 7 days in solid cattle faeces (Olson et al., 1999). In cattle faeces, *Cryptosporidium* oocysts were found to survive nearly six months at 4 °C (Robertson et al., 1992). These studies clearly indicate that (oo)cysts are highly susceptible to the higher temperature (>15 °C). Besides temperature, pH and ammonia generated in manure can also adversely affect (oo)cyst survival. Jenkins *et al.* (1998) reported that 0.06 M ammonia would inactivate 99.9% of freshly purified oocysts in 8.2 days at a temperature of 24 °C. Another study that exposed (oo)cysts to low ammonia concentration (5-50 mg NH3 l^{-1}) for 4 days found a significant reduction of oocyst viability (41·5-14·8%, respectively) (Reinoso et al., 2007). The latter study also reported 80-51% reduction in oocyst viability when they were exposed at pH 10 for 6 days at 25°C. Therefore, it is important to consider pH, ammonia and temperature of the environment to determine the survival of the (oo)cysts.

However, *Giardia* cysts and *Cryptosporidium* oocysts can survive much longer in a wet environment (Ziemer et al., 2010; Cotruvo et al., 2004). It was reported that *Giardia* cysts survived for up to 77 days at 4-8 °C water and *Cryptosporidium* oocysts up to one year at 5 °C water (Ziemer et al., 2010; WHO, 2004b). At low water temperatures (-10 °C), oocysts are able to maintain viability for a few days (Ziemer et al., 2010; WHO, 2004). *Cryptosporidium* oocysts are more environmentally resistant than *Giardia* cysts. Therefore, *Cryptosporidium* oocysts require higher doses and longer exposure (up to 28000 mg/L for 24 h) to chlorine than *Giardia* cysts (5-15 mg/L for 30 min to 2 h) for inactivation (Bukhari et al., 1999; Craik et al., 2001). However, *Cryptosporidium* oocysts are susceptible to low doses of UV radiation (9 mJ/ml), and to ozonation at lower temperatures (Bukhari et al., 1999; Craik et al., 2001). It implies that natural sunlight in mid-summer could kill (oo)cysts in the environment, but further study is required to determine how other factors such as pH and moisture of faeces may affect oocyst survivability.

1.5.4 Occurrence of pathogens in waterways

Campylobacter, Cryptosporidium and *Giardia* species have been isolated from various waterways including surface, ground and recreational waters (Koenraad et al., 1997; Rose, 1997; Thomas et al., 1999; Jones, 2001; Robertson et al., 2001; Savill et al., 2001; Fayer and Xiao, 2008; Vereen et al., 2007; Lujan and Svard, 2011; Jokinen et al., 2012; Hokajärvi et al., 2013). *Campylobacter* species were detected in 70% of water samples from rivers and lakes in the Warsaw region of Poland (Popowski et al., 1997). A study

in the UK found 40.5% water samples positive for *Campylobacter* (Kemp et al., 2005). Van Dyke *et al.* (2010) found *Campylobacter* spp. in 0–23% of samples using a culture-based method and 57-79% using quantitative PCR assay from different surface water sources collected over three years in Ontario, Canada. In England, *Campylobacter* spp. were not detected in the samples from streams in upland moors, but were reported in the samples from the same streams running through lowland or grazed pasture (Jones et al., 1990; Jones and Hobbs, 1996). Similar findings were also reported for the river water samples in the same study. Many studies reported a higher frequency of *C. jejuni* than *C. coli* in streams/rivers than (Stanley et al., 1998; Thomas et al., 1999; Jones et al., 1999; Brown et al., 2004; Kemp et al., 2005). *C. jejuni* were commonly isolated from trough-water and running-water sources and *C. coli* were isolated from standing water (Kemp et al., 2005). Subtyping of *C. jejuni isolates* showed that most of the *C. jejuni* sequence types identified in river water could be attributed to wild bird faecal contamination followed by ruminants (Carter et al., 2009; Jokinen et al., 2011). There was also evidence of *C. jejuni* occurrence in surface water in all seasons, with the lowest frequency found in the summer, when elevated UV levels and temperatures exist (Jones, 2001), and in ground water (Stanley et al., 1998).

In different countries, *Cryptosporidium* spp. oocysts were detected in between 4 to 100% of surface water samples, with oocyst concentrations ranging between 0.01 and 5800 oocysts/L (Fayer and Xiao, 2008; Rose et al., 2002; Smith et al., 1995). *Giardia* cyst occurrences were between 30 and 100% of samples of surface water, with concentrations ranging from <0.01 up to <32,400 cysts/L (Anceno et al., 2007; Lim et al., 2008; Schets et al., 2008). In Sweden, *Cryptosporidium* oocysts were detected in 32% (16/50) and *Giardia* cysts in 26% (13/26) of water samples (Hansen et al., 1998). The (oo)cyst occurrence in water varies according to the type of land use in the watershed (Smith et al., 1995; Rose et al., 2002;Fayer and Xiao, 2008;). (Oo)cysts, though less frequent, were also reported in groundwater samples, which is alarming as groundwater has a relatively low risk of contamination from pathogens (Rose, 1997; Karanis et al., 2006; Fayer and Xiao, 2008; Lim et al., 2008). Yet, 5.8% of 258 well-water samples from the UK were positive for *Cryptosporidium* oocysts, and 22% of 18 well-water samples from Bulgaria and 18% of 28 well-water samples from Malaysia were found contaminated with *Giardia* cysts (Rose, 1997; Karanis et al., 2006; Lim et al., 2008). Similarly, *Cryptosporidium* and *Giardia* (oo)cyst occurrence in recreational water is well documented, although the level of contamination is generally low (Fayer and Xiao, 2008; Robertson and Lim, 2011). Lim *et al. (*2009) and Wicki *et al.* (2009) reported *Giardia* cysts in 77.8% (7/9) and 97.5% (39/40) of recreational water

samples collected from Malaysia and Switzerland, respectively, implying that *Giardia* spp. are ubiquitous in recreational water in those countries. However, *Cryptosporidium* oocysts were not detected in that Malaysian study. It is estimated that if (oo)cysts are absent in >90% of water samples analysed, the ambient concentration of pathogens in surface water could be 1 (oo)cyst per 100 L of water (Ongerth and Saaed, 2013).

From the above studies, it is clear that both surface and ground water are often contaminated with *Campylobacter*, *Cryptosporidium* and *Giardia* spp., indicating that faecal contamination in water is a serious public health risk. Therefore, identifying and understanding the characteristics of different pathogen detection methods is important to track contamination sources, and subsequently, develop effective protection strategies for better water quality.

1.6 Detection of pathogens in faeces and/or water

With the advancement of molecular biology, microbial source tracking has improved rapidly, providing a much better understanding of the role of ruminants as water contaminants. Yet, the causes of waterborne outbreaks often remain unknown because water parameters of turbidity, temperature and pH influence the analytical sensitivity of the detection methods, making pollution source tracking difficult. In addition, laboratory detection of pathogens is challenging due to their relatively low test-sensitivity, and the long time required to perform the tests (Hunter et al., 2003; WHO, 2004). Here, the main methods available for the detection of *Campylobacter* spp., and *Cryptosporidium* and *Giardia* (oo)cysts in faecal/water samples for tracking the sources of water pollution as applied are described in brief.

1.6.1 Pathogen isolation

Several isolation techniques for *Campylobacter* spp. and concentration techniques for *Cryptosporidium* and *Giardia* (oo)cysts have been described. The choice of isolation or concentration technique may, however, influence the result of the testing. In this PhD study, *Campylobacter* spp. were isolated using a culture technique, and *Cryptosporidium* and *Giardia* spp. were detected using immunofluorescence assay (IFA). The culture of samples in selective media is a popular method of *Campylobacter* spp. isolation in both faeces and water. Usually, faecal samples are used directly for culturing *Campylobacter* spp., whilst rectal swabs in a transport medium such as Amies or Stuart are also used for culture. For water samples, prefiltration of at least 100 ml of water and culture of the filter is recommended to isolate *Campylobacter* spp. It has

been reported that 0.22 µm size filters retain a higher amount of *Campylobacter* spp. than 0.45 µm size filters (Donnison, 2003). Modified charcoal, cefoperazone, desoxycholate agar (mCCDA) is the recommended selective medium for *Campylobacter* spp. culture, although other blood -containing media are also used. Studies have reported that *Campylobacter* spp. are better isolated if samples are enriched in Preston or Bolton broths media, with antibiotics added to suppress other bacterial growth (Bolton et al., 1983; Mason et al., 1999). To recover *Campylobacter* spp*.,* the samples are cultured at microaerophillic (O_2 <10%), capnophillic (CO_2 >5%) and thermophillic (42 ^{o}C) atmospheres on the selective media for 48 hours. Some *Campylobacter* species (e.g. *C. fetus)* may require hydrogen and lower incubation temperature (>37 ^{o}C).

Techniques for the concentration of (oo)cysts from faecal samples have improved the sensitivity of *Cryptosporidium* and *Giardia* spp. detection compared with analysis of direct smears, although the process entails a considerable loss of (oo)cysts (Casemore et al., 1985; Weber et al., 1992; Garcia et al., 1983). For example, the concentration of faecal samples by formalin-ether sedimentation detected 72% of *Giardia*-positive samples, compared with a 31% detection rate in direct smears (Baughn and Morales, 1971). Formalin-ethyl acetate sedimentation and sugar floatation performed much better than zinc sulphate flotation in detecting *Cryptosporidium* oocysts (McNabb et al., 1985; Mtambo et al., 1992; Bukhari et al., 1995). Several studies indicated that the use of immunomagnetic separation (IMS) provides higher sensitivity for the concentration of (oo)cysts from faeces than traditional methods based on centrifugation without IMS (Webster et al., 1996; Pereira et al., 1999; Power et al., 2003).

A sufficient volume of water should be collected for better (oo)cyst recovery. During the 1980s and 1990s, 100 to 1 000 litres of water were concentrated with yarn-wound cartridge filtration, and found (oo)cyst recovery efficiencies ranging between 8-9% for *Cryptosporidium* and 12-28% for *Giardia* at different (oo)cysts inoculation levels (Nieminski et al., 1995; Shepherd et al., 1996; Hsu et al., 2001a). Collection of a large volume of water was sometimes counterproductive for detection of *Cryptosporidium/Giardia* spp., however, as multiple factors including high water turbidity and quickly clogging filters affected the isolation of (oo)cysts. Subsequently, many filtration techniques were developed, modified and validated for (oo)cyst recovery. Currently, the most common procedure used for detection of *Cryptosporidium/Giardia* (oo)cysts is the United States Environmental Protection Agency approved method 1623 (USEPA 1623; 1996, 2012). The USEPA 1623 protocol includes the manual elution of protozoa from filters or the use of an automatic wash

station, concentration of (oo)cysts using immunomagnetic separation, and detection by direct immunofluorescence assay (IFA). The collection of at least 10 litres of water is recommended, using filters such as Filta-Max® or Envirocheck filters for retention of protozoa from raw water (Smith et al., 2010). The mean recovery efficiency for (oo)cysts is as high as 90.2% (Sartory et al., 1998) for the USEPA 1623, but this depends on inoculation levels, water turbidity, and interlaboratory variations. For example, Feng *et al.* (2003) spiked 2010 oocysts in 10 litres of water, and obtained the highest mean recovery of *Cryptosporidium* oocysts (85%) at 5 nephelometric turbidity units (NTU) using pleated membrane filtration. DiGiorgio *et al.* (2002) recovered 51% *Cryptosporidium* oocysts and 53% *Giardia* cysts at a turbidity level of 11 NTUs when 100 (oo)cysts were spiked into 10 litres of water. A moderate level of turbidity may enhance the recovery of (oo)cysts, but higher levels affect the filtration capacity by clogging the filter and filtering lower volumes of water (DiGiorgio et al., 2002). Paramagnetic beads coated with antibodies against *Cryptosporidium* and *Giardia* used in the IMS step have successfully recovered (oo)cysts from environmental samples, although recovery efficiencies varied widely depending on the commercial IMS kit used. Variable recovery rates ranging between 62% and 100% at high turbidity were reported using Dynal® (Oslo, Norway) IMS kit (Rochelle et al., 1999; Hsu et al., 2001b; McCuin et al., 2001; Feng et al., 2003; Nichols et al., 2006).

1.6.2 Identification of pathogens

Discriminating and sensitive methods are required to identify the isolated pathogens. Therefore, phenotyping or genotyping methods are used either alone or in combination for tracking the sources of water pollution.

1.6.2.1 Identification by phenotyping

Phenotyping methods are still widely used for microbial characterisation at the genus level, including morphology characterisation, serotyping, biotyping and phage typing. The morphology characterisation is quite common for protozoan identification but is not always the preferred method for bacterial strain identification due to the possibility of misdiagnosis.

1.6.2.1.1 Phenotypic identification of Campylobacter spp.

The colonies found on culture media are observed for the presence of motile, spiral-shaped bacteria under a phase-contrast microscope, for identification of presumptive *Campylobacter*. *Campylobacter* colonies are slightly pink, round, convex, smooth and

shiny, with a regular edge on Skirrow or other blood-containing agarised media (Nachamkin et al., 2008). On charcoal-based media such as mCCDA, they are observed as greyish, flat and moistened, with a tendency to spread colonies, and they may have metal sheen colonies. Suspected colonies are usually also subjected to an oxidase test to identify *Campylobacter* species, which are oxidase- positive. Other phenotyping techniques such as biotyping, phage typing and fatty acid profile typing are also used for *Campylobacter* species identification, but they suffer from poor typeability and low discriminatory power (Fox et al., 1996; Steele et al., 1998; Klena, 2001).

1.6.2.2 Phenotypic identification of *Cryptosporidium* and *Giardia spp.*

The immunofluorescence assay (IFA), developed in 1985, is commonly used for morphological identification of *Cryptosporidium*/*Giardia* (oo) cysts under a fluorescence microscope (Sauch, 1985; Johnston et al., 2003; Fayer and Xiao, 2008). In IFA, a stain containing fluorophore-labelled antibody is used. Both polyclonal (pAbs) and monoclonal (mAbs) antibodies have been used to identify (oo)cysts in faeces/water (Ongerth et al., 1987; Rose et al., 1989). The mAbs are believed to improve the sensitivity of the technique, but anti-*Cryptosporidium* mAbs may still cross-react with debris such as algae in water or with faecal yeasts, giving rise to non-specific fluorescence (Sterling et al., 1986; Rodgers et al., 1995). The antibody attaches to cell wall antigens of the (oo)cysts illuminating with bright green colour under the dark field fluorescence microscopy. The criteria proposed by Rose *et al.* (1989) are commonly used for *Cryptosporidium* and *Giardia* spp. identification in both faeces and water. The criteria include (i) the degree of fluorescence should be at least 50% of that seen in the control slide; (ii) a distinct fluorescence around the wall of (oo)cysts; and (iii) a spherical shape of 4-6 µm oocyst and an oval shape of 8-18 x 5-15 µm cysts. Studies have shown that IFA has 93-100% sensitivity and 99.8-100% specificity for *Giardia* compared to other staining techniques (Arrowood and Sterling, 1989; Quilez et al., 1996; Lemos et al., 2005; Baig et al., 2012). Another antibody-based test that is commonly used is enzyme-linked immunosorbent assay (ELISA). ELISA is used to identify *Cryptosporidium*/*Giardia* antigens in faeces by assessing colour shift using visual examination or spectrophotometry. Compared with IFA, ELISA has a lower sensitivity for *Cryptosporidium* (68-99%) and *Giardia* (89-100%), with 99-100% specificities for both protozoa in faeces, when ELISA kits designed for humans are used (Garcia et al., 1997; Johnston et al., 2003; Fayer and Xiao, 2008). Although many methods have been developed, IFA is considered the ‘‘gold standard’’ test to identify *Cryptosporidium*/*Giardia* spp. because of its ability to identify entire oocysts and its

usefulness in quantifying the oocysts, rather than just detecting soluble antigens (Langhout et al., 2010; Chalmers et al., 2011). However, neither IFA nor ELISA can differentiate *Cryptosporidium*/*Giardia* species (Monis and Thompson, 2003) and genotypes/subtypes. The current phenotyping methods can, therefore, provide adequate discrimination for microbial source identification at genus level in faeces/water, but genetic level determination is required to provide insight into the pathogens' diversity, their sources and role in diseases and transmission dynamics.

1.6.2.3 Identification by genotyping

Genotyping involves the detection of single or multiple fragments of certain genes, or the whole genome, particularly through amplification of the nucleic acids *in-vitro*. DNA molecular tools have been widely used to determine the species, genotypes and subgenotypes of pathogens because they enhance discriminatory power and sensitivity (Hadfield et al., 2011). In addition, improvement in DNA technology, its ease of use, high throughput of processing large numbers of samples, and the lowering of costs have revolutionised the epidemiological study of pathogens (Duim et al., 2000; Caccio and Ryan, 2008).

Standard techniques for genotyping *Campylobacter*, *Cryptosporidium* and *Giardia* isolates from faecal and environmental samples include extraction of DNA from the sample, PCR amplification and visualisation of the DNA, and/or sequencing of the DNA. The amplification of DNA depends on an efficient DNA extraction. Therefore, in the following sections, I have provided a brief overview of DNA extraction methods and different genotyping and/or subtyping tools for *Campylobacter, Cryptosporidium* and *Giardia* with an emphasis on the methods used in this thesis.

1.6.2.3.1 Campylobacter spp. genotyping and/or subtyping methods

1.6.2.3.1.1 DNA extraction of Campylobacter spp.

The initial step, crucial in genotyping, is the extraction of DNA from presumptive *Campylobacter*, as a sufficient amount of DNA is required for genotyping and/or subtyping methods. A variety of DNA extraction protocols has been developed for the isolation of *Campylobacter* DNA from faecal/water samples. The protocols include steps of Proteinase K digestion, boiling, use of Chelex®100, phenol-chloroform, immunomagnetic separation, and use of commercial kits (Kirk and Rowe, 1994; Mohran, et al., 1998; Shah et al., 2002). However, none of them has been defined as the best standard methodology because of relatively low sensitivity. Therefore, different laboratories use different methods according to their preference.

1.6.2.3.1.2 Methods of genotyping and/or subtyping of Campylobacter spp.

Many methods have been developed to genotype and/or subtype *Campylobacter* spp. Here, some of the methods, such as pulsed field gel electrophoresis, restriction fragment length polymorphism, polymerase chain reaction, and multilocus sequence typing, that are commonly used in various countries, are described in brief.

Pulsed-field gel electrophoresis (PFGE) is a restriction enzyme analysis technique that separates long strands of DNA by length during alternate electric fields of gel electrophoresis. PFGE has been used successfully for many years in subtyping bacterial strains. In PFGE, 4-20 distinct DNA bands are generated per sample. PFGE is still used to examine polymorphism within the *Campylobacter* genome due to its excellent discriminatory power, high resolution, and available rapid protocols. PFGE is unsuitable, however, for typing large numbers of samples, as it lacks stability and reproducibility for genotyping and/or subtyping *Campylobacter*, and shows inter-laboratory variation (Duim et al., 2000; Klen, 2001).

Restriction fragment length polymorphism (RFLP) is the method of selective amplification of restricted fragments generated from DNA. This method can distinguish between genetically related and unrelated strains, is relatively simple and reproducible, and is equally discriminatory as PFGE (Cardarelli-leite et al., 1996). Therefore, RFLP has been used for *Campylobacter* epidemiological surveillance. However, like PFGE, it is also band-based and cannot discriminate fingerprints between two distinct sequences with the same peak, which may lead to misclassification.

Polymerase chain reaction (PCR) is a popular molecular technique used to amplify a specific region of a gene/genome (often between 0.1 and 10 kbp), and the product is visualised using agarose gel electrophoresis. Different types of PCR exist, of which quantitative PCR (qPCR), nested PCR, reverse-transcription PCR (RT-PCR) and multiplex-PCR are often used in the laboratory. qPCR has a high degree of precision and is used in microbial source tracking as it can quantify various bacteria, protozoa and viruses, even unculturable organisms, in the laboratory (Fong and Lipp, 2005; Stelzer et al., 2012). Multiplex PCR is used to identify multiple strains and multiple traits of same strains simultaneously, and is employed in differentiating *Campylobacter* species (Yamazaki-Matsune et al., 2007). Conventional PCR is also a choice of genotyping tool for the initial screening of many organisms, including *Campylobacter* spp. The sensitivity of PCR, however, depends on various factors such as the quantity

and quality of DNA, enzymatic reaction and the quality of reagents used. PCR performed directly from water without previous culture cannot differentiate between viable or non-viable pathogens in the aquatic environment, and may overestimate the risk, especially when water treatment kills the pathogen without disrupting nucleic acids. Therefore, sequencing of the PCR-products (amplicons) is increasingly used to avoid false representation of organisms. DNA sequencing has a more precise measure of genetic variability and leads to the development of molecular subtyping methods such as direct PCR-sequencing, PCR-RFLP sequencing and multilocus sequence typing.

Multilocus sequence typing (MLST) was developed in 1991, and used first in determining virulent lineages of the bacterial pathogen *Neisseria meningitidis* (Maiden et al., 1998; 2006; Enright and Spratt, 1999; Dingle, et al., 2001). This method characterises strains by measuring and comparing the variations in DNA sequence in a set of seven conserved housekeeping genes of *Campylobacter* spp. (Dingle, et al., 2001). The seven housekeeping genes used for MLST are: aspatase (aspA), glutamine synthetase (glnA), citrate synthase (gltA), serine hydroxy methyl transferase (glyA), phospho glucomutase (pgm), transketolase (tkt), and ATP synthase alpha subunit (atpA/uncA). This method is analogous to multi-locus enzyme electrophoresis (MLEE) but differs in that it assigns alleles directly by DNA sequencing, rather than indirectly by electrophoretic mobility of their gene products. MLST is able to distinguish isolates that were indistinguishable with PFGE (Sails et al., 2003). Moreover, in contrast to most typing procedures, the MLST output (the allelic profiles of the isolates) can be compared to those in a virtual library database via the internet (e.g. PubMLST[2]). MLST has been commonly used in *Campylobacter* species subtyping, and the data have enabled a study to be carried out concerning source attribution for human campylobacteriosis (Wilson et al., 2005; Mullner et al., 2009). MLST has also elucidated *Campylobacter* diversity in animal hosts or environmental niches, thereby increasing the understanding of transmission dynamics of *Campylobacter*. Thus, MLST is considered a powerful tool for global epidemiology and population biology studies of *Campylobacter* (Maiden, 2006), although it appears to be less discriminative than PFGE in short-term epidemiological studies (Sails et al., 2003).

1.6.2.3.2 Cryptosporidium and Giardia genotyping and/or subtyping

1.6.2.3.2.1 DNA extraction of Cryptosporidium and Giardia spp.

[2] http://pubmlst.org/campylobacter/

Several research groups have grown *Giardia* cysts in culture media with or without a mammalian serum (Bifulco and Schaefer 1993; Clark and Diamond, 2002; Mata-Cárdenas et al., 2012). Despite some recent success (Aldeyarbi and Karanis, 2014; Yang et al., 2015), the cell-free culture of *Cryptosporidium oocysts* has not been very successful. *Cryptosporidium* and *Giardia* (oo)cysts are highly resistant to many chemicals and heat compared with *Campylobacter* cells. Therefore, a reliable (oo)cyst disruption process is required to extract DNA from faeces and water. There are no gold standard methods for *Cryptosporidium/Giardia* DNA extraction methods. Methods that are used for *Cryptosporidium* can also be used for *Giardia,* as *Giardia* cysts are less robust than *Cryptosporidium* oocysts (Nichols et al., 2006; Smith and Nichols, 2008). The common methods applied for the extraction of DNA from partially purified (oo)cysts are boiling, freeze-thaw cycles, chemical, enzymatic or mechanical treatments, and bead-beating (Balatbat et al., 1996; McLauchlin et al., 1999; Elwin et al., 2001; Amar et al., 2002, 2007; Lindergard et al., 2003; Nichols et al., 2006). Thereafter, DNA is extracted using commercial spin columns, glass milk, or chelex resin (Xiao et al., 2001; Read et al., 2004). The choice of DNA extraction protocol used varies between laboratories, thus validation of the protocols used in the laboratory is recommended(Smith and Nichols, 2009).

Old or frozen *Cryptosporidium* and *Giardia* (oo)cysts containing samples have been used for DNA extraction (Elwin et al., 2012), including the (oo)cysts on stained microscopy slides (if the original sample is unavailable) (Amar et al., 2001). Such slide-genotyping technique is commonly used for water samples, where the number of recovered (oo)cysts is extremely low, although there is a need to improve the recovery of DNA for amplification. In slide-genotyping, the smeared material on the slide is scraped off into lysis buffer, and the lysate is used for DNA extraction using suitable methods (Amar et al., 2001; Di Giovanni et al., 2010). It is reported that mounting media containing formalin used during slide preparation interferes with DNA extraction. In addition, closed cell foam used for scrapping slides was found to have a better recovery than cotton swabs (Di Giovanni et al., 2010). This modification technique has been able to recover *Cryptosporidium* and *Giardia* DNA allowing the genotyping of (oo)cysts present in water samples, but further studies are required to improve the recovery efficiency.

1.6.2.3.2.2 Methods of Cryptosporidium spp. genotyping and/or subtyping

Molecular methods targeting various genetic loci have been used for identifying *Cryptosporidium* species/genotypes in faecal and water samples (Xiao and Ryan, 2008).

Analysis of at least two genomic loci from an isolate is recommended for reliable confirmatory identification (Caccio et al., 2005; Fayer and Xiao, 2008; Xiao and Ryan, 2008). Preferably, one of these should target the highly or moderately conserved coding region, and the other should be one that can identify species and subtypes (Caccio et al., 2005). Coding loci, such as small subunit ribosomal RNA (18S rRNA) genes for eukaryotes, structural and housekeeping genes should be analysed. These coding regions include the *Cryptosporidium* (oo)cyst wall protein (COWP), the heat shock protein 70 (Hsp70), actin genes and 60kDa glycoprotein (gp60) for *Cryptosporidium* (McLauchlin et al., 1999; Xiao et al., 2000, 2001; Chalmers et al., 2005a, 2011; Xiao and Ryan, 2008). The 18S rRNA is a widely used locus for *Cryptosporidium* genotyping. It is particularly useful for taxonomic typing because the 18S rRNA locus comprises a species-specific region (Caccio et al., 2005). In addition, its analysis offers improved sensitivity due to the presence of multiple copies (five) of the gene and the presence of both conserved and polymorphic regions. Other target-gene primers may not amplify DNA from all species, resulting in an underestimation of some species. For example, gp60 analysis does not identify *C. bovis* whereas 18SrRNA does, and this may lead to a report of an absence of *C. bovis* in the analysed samples. Furthermore, the primers of the 18S rRNA are not universal.

Nested and semi-nested conventional PCR of the locus is often used for amplification and amplicons are visualised by gel electrophoresis to confirm the taxon. However, PCR followed by restriction fragment length polymorphism (PCR-RFLP) analysis of the 18S rRNA locus has also been described in many studies for differentiation of *Cryptosporidium* species and genotypes (McLauchlin et al., 1999; Xiao et al., 2000; Heitman et al., 2002; Coupe et al., 2005). The PCR-RFLP analysis can distinguish between genetically related and unrelated strains, but its interpretation is complex due to the presence of indistinguishable banding patterns. Yet, because of its relative simplicity and reproducibility, RFLP has been widely used for *Cryptosporidium* epidemiological surveillance. Other analytical methods such as real-time PCR and digital PCR that detect and quantify the *Cryptosporidium* species may be used widely in food and water microbiology, but these need further validation (Elwin et al., 2012; Yang et al., 2014). These PCR-based genotyping tools can differentiate between anthroponotic and zoonotic types of *Cryptosporidium* species, but their resolution is too low to be used in epidemiological studies. Therefore, genotyping/subtyping of *Cryptosporidium* species, particularly *C. parvum* and *C. hominis,* have been used increasingly, especially in targeting the gp60 gene, minisatellites and microsatellite markers, double-stranded RNA elemenst, and/or the internal transcribed spacer-2 of

the rRNA gene sequencing (Aiello et al., 1999; Caccio et al., 2000, 2001; Peng et al., 2001; Sulaiman et al., 2001; Xiao et al., 2001; Alves et al., 2003; Gasser et al., 2004; Widmer et al., 2004, 2006; Chalmers et al., 2005a; Leoni et al., 2006; Tanriverdi et al., 2006; Gatei et al., 2007). Of these, the gp60 gene as well as mini- and micro-satellite typing offer sufficient subspecies discrimination in population structure studies for understanding aspects of transmission dynamics (Xiao et al., 2004; Caccio et al., 2005; Smith et al., 2006). These markers have sequence motifs repeated in tandem on 1-4 base pairs (microsatellites and gp60) or more (minisatellites) that are hypervariable. In gp60 sequencing, in addition to the number of trinucleotide (TCA/TCG/TCT) repeats, non-repeat regions also vary extensively. Thus, sequencing of the gp60 gene allows a definition of subtype families Ia, Ib, and Id to Ii of *C. hominis* and IIa to IIo of *C. parvum* (Fayer and Xiao, 2008; Nichols et al., 2014). Therefore, sequencing of the gp60 gene is widely used in *Cryptosporidium* subtyping, particularly for *C. hominis* and *C. parvum*, because of being a very polymorphic marker (Fayer and Xiao, 2008). Both single locus and multi locus (such as multilocus typing) schemes have been used for the subtyping of *Cryptosporidium*, though single locus is preferred for source tracking studies because of its ease of use and its lower cost.

1.6.2.3.2.3 Methods of Giardia spp. genotyping and/or subtyping

Highly or moderately conserved gene regions such as the 18S SSU rRNA, β-giardin (bg), glutamate dehydrogenase (gdh), and triosephosphate isomerase (tpi) (Feng & Xiao, 2011) are used for *Giardia* genotyping. PCR protocols have been developed for various markers used for the identification of *Giardia* to the species, assemblage and subassemblage levels (Feng and Xiao, 2011). These markers differ widely, however, because of the existence of considerable genetic variations. The markers such as 18S rRNA and elongation factor1-α are strongly conserved and only used for the identification of *Giardia duodenalis* assemblages (Wielinga and Thompson, 2007). Additionally, the 18S SSU rRNA locus has GC richness requiring special PCR buffers for amplification (Ryan and Caccio, 2013). Therefore, these markers are less useful in determining genetic variability within *Giardia duodenalis*. Consequently, the highly intraspecific variable loci tpi, bg and gdh have been used along with 18S rRNA for the current classification of *Giardia* (Amar et al., 2002; Wieling and Thompson, 2007; Caccio and Ryan, 2008; Geurden et al., 2009; Feng and Xiao, 2011; Thompson and Monis, 2012). It has been suggested that a single marker with high genetic heterogeneity can provide a resolution as high as the multilocus approach, although

the latter provides accuracy on genotyping *Giardia* (Caccio et al., 2008; Lebbad et al., 2008).

Conventional PCR is also the commonest method used to amplify target markers of *Giardia* for genotyping. PCR-sequencing is recommended for determining assemblages or subassemblages. Other PCR- based methods such as PCR-fingerprinting, PCR-RFLP, real-time PCR and multiplexed tandem-PCR have also been developed for *Giardia* genotyping (Amar et al., 2003; Read et al., 2004; Haque et al., 2007; Jex et al., 2012; Zhang et al., 2012; Koehler et al., 2014). Real-time PCR can also be used for examining *Cryptosporidium/Giardia* isolates from water (Fontaine and Guillot, 2003; Koehler et al., 2014). However, these methods do not detect all sequences and may not always delineate multiple distinct types of sequences within a sample (Koehler et al., 2014). Therefore, more research is required to determine their specificity and sensitivity.

1.6.2.3.2.4 Identification of mixed infections

Molecular typing and subtyping have also increasingly reported mixed infections of *Cryptosporidium* species or *Giardia* assemblages (McLauchlin et al., 2000; Cama et al., 2006; Morse et al., 2007; Helmy et al., 2013; Ng-Hublin et al., 2013; Ryan and Caccio, 2013; Shrestha et al., 2014). Mixed infections complicate the understanding of the genetics of these protozoa, as the true occurrence of mixed species genotypes in samples or recombination of these taxa is poorly understood (Grinberg et al., 2013; Ryan and Caccio, 2013). Many studies have used single locus or genus-specific molecular typing tools. As a result, mixed infections may not be readily detected because of preferential PCR amplification of the predominant genotypes present in the samples. Therefore, multilocus typing is recommended for confirmation of the species present. However, amplification at two different loci may also detect the different *Cryptosporidium* species or *Giardia* assemblages because of preferential PCR amplification of those loci (Sprong et al., 2009; Ryan and Caccio, 2013). In spite of this, conventional PCR, PCR-RFLP or real-time PCR have shown that mixed infections are not uncommon in animals and humans (Limor et al., 2002; Mallon et al., 2003a; Fayer and Xiao, 2008; Geurden et al., 2009; Cooper et al., 2010; Rayan and Caccio, 2013). Nevertheless, PCR assays with broad specificity coupled with species- or assemblage-specific tools provide better options for mixed infection analysis, and this is used in chapter 4 of this thesis.

1.6.2.4 Sequencing approaches to identify pathogens

Many molecular techniques such as PCR, RFLP and MLST are often coupled with sequencing for genotyping. Therefore, sequencing has become an integral part of genotyping techniques for understanding the pathogens' diversity and their impact on human/animal diseases. Each DNA molecule contains up to four nucleotides: adenine, guanine, cytosine and thymine arranged in a precise order. During sequencing, the order of these nucleotides is determined to define a unique sequence. Although the conventional sequencing (Sanger sequencing) approach has been widely used, application of the latest sequencing technology (next-generation sequencing) is about to provide a deeper understanding of the ecology of pathogens.

1.6.2.4.1 Sanger sequencing

DNA sequencing technology was first introduced in the 1970s (Sanger et al., 1973). In 1977, Sanger and Gilbert developed a sequencing technology that became known as "Sanger sequencing". Sanger sequencing was rapidly adopted due to its reliability and ease of use (Sanger et al., 1977). Over the next two decades, the Sanger method was improved by the advent of fluorescent labelling, capillary electrophoresis (CE) and automation that resulted in the production of the first human genome database (~3Gigabases) at a cost of three billion USD (Human Genome Sequencing, 2004; Schmutz et al., 2004). Therefore, this automated Sanger sequencing method is also called the "first generation sequencing". In Sanger sequencing, a single strand DNA molecule reacts with dideoxy-nucleotides triphosphates and primer in the presence of DNA polymerase. Then, a chain terminator is tagged in the sequences and fractionated via electrophoresis, or with a chromatography capillary column to identify colour tags. These colour tags are then read to produce the consensus sequence (Perez-Losada et al., 2013). This sequencing approach can sequence single molecules and decrease artificial recombination impacts with simple post-processing of the sequences and a low error rate (0.0001-1%) (Ewing et al., 1998; Perez-Losada et al., 2013). Therefore, Sanger sequencing is still the gold standard for DNA-sequencing and has been applied widely in both basic and applied biological research including research involving *Campylobacter*, *Cryptosporidium* and *Giardia,* to define the species (Harismendy et al., 2009; Perez-Losada et al., 2013). However, Sanger sequencing not only has low throughput but it is also laborious and time-consuming to process large numbers of samples (Perez-Losada et al., 2013). In the last eight years, there have been rapid changes in sequencing technologies, with the development of so-called "next-generation sequencing" (syn: high-throughput sequencing) technologies.

1.6.2.4.2 Next-generation sequencing (NGS)

Millions of fragments of DNA/RNA can be sequenced in parallel in a single stroke, and processed much more quickly at lower cost using next-generation sequencing (NGS) (Mardis, 2008; Shokralla et al., 2012; Perez-Losada et al., 2013). Therefore, NGS has gained popularity in the sequencing world and has greatly improved our understanding of the microbes, their ecology, intra-host diversity and transmission dynamics (Hazen et al., 2013; Perez-Losada et al., 2013). The underlying principle of next- generation sequencing involves the sequencing of DNA molecules in a stepwise repetitive process, or in a continuous real-time manner, where each individual template fragment is independently sequenced and counted among the total sequences generated (Pareek et al., 2011).

1.6.2.4.2.1 Types of Next-generation sequencing platforms

Next-generation sequencing platforms include technologies like Illumina (Solexa) sequencing, Roche 454 sequencing and Ion torrent: Proton / PGM sequencing (Mardis, 2008; Shokralla et al., 2012). In Illumina and Roche 454 sequencing, longer DNA fragments are ligated to generic adaptors and are annealed to a glass slide using the adaptors (Mardis, 2008; Shokralla et al., 2012). A PCR is carried out to amplify each DNA molecule, which is separated into single strands and sequenced at once by reading optical signals as bases are added (Mardis, 2008; Shokralla et al., 2012). Roche 454 sequencing provides greater flexibility by accurately annotating the reads (nucleotides sequences) in ecological applications, but it lacks terminating moiety to stop the extension run, generates high sequence errors, and is relatively costly for reagents per megabase sequencing outputs (Sogin et al., 2006; Mardis, 2008; Claesson et al., 2010; Shokralla et al., 2012). Illumina can sequence homopolymer regions more accurately, as nucleotide detection is performed one at a time, and can give high output per run at a low cost compared to Roche 454 sequencing (Shokralla et al., 2012). However, this technology generates a relatively shorter read-length than Roche 454 sequencing, thus limiting the application of Illumina where there is no reference sequence available (Shokralla et al., 2012). Ion torrent sequencing includes fragmentation of DNA or RNA, amplified using emulsion PCR, and reads the pH changes due to H^+ ion release to determine the bases (Mardis, 2008; Shokralla et al., 2012). Therefore, the choice of NGS platform is often difficult, because of large variability among and within the platforms in terms of template size needed and construction, read-length, throughput and coverage, each with its own pattern of bias (Harismendy et al., 2009; Metzker, 2010).

1.6.2.4.2.2 Library preparation

Each technology has its own library preparation methods. Briefly, the DNA samples need to be converted into a special library (Bentley et al., 2008). The input DNA is fragmented into short lengths (400-1000bp), and the ends of each fragment ligated using two end-specific adaptors. The adaptors act as primers for the amplification that hybridised the double- stranded section of the template, and the amplification is repeated several times to produce ~1,000 copies of the original sequence. The double-stranded DNA (dsDNA) is denatured and a single- stranded DNA (ssDNA) library is prepared (Adessi et al., 2000; Fedurco et al., 2006; Kircher and Kelso, 2010).

1.6.2.4.2.3 Application of next-generation sequencing platforms

Next-generation sequencing has been used to study groups of genomes (metagenomics) or the whole genome of a single organism or a targeted gene of the organisms (amplicons sequencing). For example, the whole genome shotgun approach of metagenomics was used to investigate the microbial ecology of the freshwater. The whole genome sequencing of *Campylobacter jejuni* sequence type 474 (ST-474) was used to investigate the diversity of this ST-474, and sequencing of the 16S rRNA or 18S rRNA amplicons were used to determine the bacterial and protist diversity in fresh water samples. respectively. Metagenomics study has become a popular tool to monitor environmental organisms, whilst whole genome and amplicon sequencing of organisms have broadened the NGS scope into human or animal health disease investigation, genomic epidemiology and ecology of organisms.

Metagenomics

The term "metagenomics" was first used in 1998 to describe the analysis of a collection of similar but not identical genetic items (Handelsman et al., 1998). Now, metagenomics is known as the study of genomes that are directly obtained from the environment, bypassing conventional culture-based methods (Su et al., 2012). Metagenomics has been used successfully to elucidate complex microbial communities present in the ecosystem, and to understand their structure, genetic and functional potential over time or under different environmental pressures (Cardenas and Tiedje, 2008).

Metagenomic analysis using NGS platforms has overcome the bias of cultivation bottleneck which exists when applying conventional (Sanger) sequencing (Handelsman, 2004), and has identified many microbes in the environment that were not reported before (Cardenas and Tiedje, 2008; Hugenholtz and Tyson, 2008; Su et al.,

2012). There are two main categories of DNA-based metagenomic studies. The first is the study that targets one or a few marker gene amplicons to reveal the composition and diversity of the microbiota. For example, the 16S rRNA gene has been investigated extensively for analysis of bacterial diversity using metagenomic studies. This method is now called metabarcoding. Another type of metagenomic study use is an entire genomic sequencing approach, or shotgun metagenomics, generated in a random approach (Kuczynski et al., 2012). The first NGS based metagenomic studies determined the presence of microbial communities from deep-sea sediments and groundwater in an iron mine by targeting the 16S rRNA region (Edwards et al., 2006; Sogin et al., 2006). Recently, NGS technologies have facilitated metagenomics studies from a variety of ecosystems, including marine water, fresh water, drinking water, soil, terrestrial and gut microbiota (Shokralla et al., 2012). For example, Staley *et al.* (2013) used Illumina sequencing technology to characterise the bacterial community at ten sites along the Upper Mississippi River and to evaluate shifts in community composition determined by land use changes, by targeting the 16S rRNA region. However, there is little information on microbial communities of surface water destined for drinking purposes. Therefore, chapter 6 of this thesis explores the metagenomic approach, using both 16S rRNA-specific and shotgun sequencing, in two surface water areas of New Zealand.

A number of caveats regarding metagenomic sequencing need to be addressed during the interpretation of the data. For example, there might be amplification biases in some data due to the use of PCR-bases that may alter the conclusion (Abulencia et al., 2006). In addition, many of the low abundance organisms are often missed, suggesting insufficient coverage of the microbial community (Hazen et al., 2013). There is also a need to use string annotation pipelines to obtain accurate metagenomic sequence data (Hazen et al., 2013). Other factors that need to be addressed for better estimation of the microbial diversity are inefficient DNA extraction, data storage and sharing, and limited data analysis tools (Wang et al., 2013).

Amplicon sequencing using NGS

Next-generation sequencing (NGS) technologies have been used for PCR-amplicon sequencing of genes of interest. In amplicon sequencing, a two-step PCR process is used to amplify the targeted gene (Bybee et al., 2011). At first, primers for the interested gene are designed with overhang adapters, the gene is amplified using conventional PCR and an amplicon-based template library is generated. Thereafter, a second PCR amplification is carried out to attach the indices and sequencing adaptors.

The libraries are normalised and pooled, and loaded into the NGS machine to sequence. This method is fully scaleable, relatively inexpensive, and applicable to phylogenetic and population genetic analysis (Bybee et al., 2011). One of the applications of amplicon sequencing is in multilocus sequence typing (MLST) of various organisms such as *Salmonella* and *Pseudomonas* (Boers et al., 2012; Singh et al., 2012). In chapter 5 of this thesis, NGS-based MLST amplicon sequencing of *Campylobacter* species found in ruminant faeces and water samples is applied.

1.7 Mitigation of pathogens' presence in water

Land use change is one of the more effective methods of pathogen mitigation because cattle or sheep not grazing or being away from waterways will reduce the faecal loads. On the other hand, it will have a huge impact on the economy of the country. Therefore, a possible way to reduce the pathogen contamination of water is to reduce the pathogen loads in ruminants themselves. This could be achieved by vaccinating ruminants against the pathogens. Vaccine development to limit the occurrence of pathogens is, however, difficult, costly and challenging. Alternatively, efforts have been devoted to: 1) the development of microbial probiotics to reduce or exclude the pathogens within the intestinal tract of livestock including cattle and sheep; 2) the inclusion of antimicrobials in the diet of ruminants at subtherapeutic levels so that pathogens can be reduced; and 3) the use of pathogen-specific bacteriophages that only target the specified pathogens without harming the commensal bacteria. However, the effect of probiotics on pathogens that are excreted in faeces still need further investigation whereas antimicrobial use may cause the development of resistance in pathogens as well as commensal flora.

Besides animal-level measures, some biosecurity measures in ruminant production systems can be employed as a farm-level approach. For intensive ruminant production, confining animals to some areas with no direct access to surface water sources is a good approach, but for extensive production, pathogens should be monitored closely by keeping the animals in a separate barn or paddock at least for 30 days before introducing them into the herd/flock during animal import, and by keeping clinically ill animals in separate paddocks, away from young calves and calving pens (Wells et al., 2002). In addition, cattle should be prevented from grazing on paddocks treated with fresh manure. Their feeding or water trough should also be routinely tested for pathogen contamination, and fencing should be deployed to reduce vectors such as rodents that may introduce the pathogens into animals. All these measures are very

difficult to implement, so a more practical solution to reduce water contamination is the exclusion of stock from waterways (Davies-Colley et al., 2004).

Cattle should be prevented from direct access to streams or waterways through fencing, and/or establishing riparian areas. Riparian areas prevent faecal deposition on the banks and act as a filter promoting entrapment of the organisms. A riparian area with 3 m grass strips removes 43-74% of microbes during rainfall. However, the entrapment of pathogens in riparian areas depends on flow characteristics and types of vegetation used. Heavy rainfall has a minimal impact on the settling out of faecal bacteria (38-84%) in riparian areas, but in lower magnitudes of rain, riparian areas can be effective in controlling 95% of faecal bacteria (Collins et al., 2004). Stockpiling or composting of faeces can be another effective method because the interior temperature of stockpiles may reach up to 70 ^{o}C, which could kill many pathogens (Topp et al., 2008). Nevertheless, none of these approaches is the absolute solution for elimination of pathogens, and such elimination may require a multiple-barrier approach, including animal-level, farm-level and catchment-level measures, along with a risk assessment for the control of the diseases and proper treatment and distribution of source water.

1.8 Waterborne zoonoses in New Zealand context: current knowledge

Campylobacter, *Cryptosporidium* and *Giardia* infections are three of the top five notifiable diseases of people in New Zealand. These three diseases were responsible for 56% of 17,711 notified cases and 91.9% of 62 waterborne outbreaks reported in 2013 (ESR, 2014). Therefore, the following sections briefly describe the current knowledge about these three pathogens in humans, animals, and water in New Zealand.

1.8.1 Epidemiology *of Campylobacter*, *Cryptosporidium and Giardia* infections in humans

1.8.1.1 *Campylobacter* spp.

The rate of campylobacteriosis in New Zealand is among the highest reported in developed countries (ESR, 2013). This disease was included in the notifiable disease

list[3] in New Zealand in 1988, and the highest rate of infection (315 per 100 000) was reported in 2006 (Figure 1.1). The increasing trend in campylobacteriosis cases (as reported in ESR[4]) that incurred a cost of >61.7 million New Zealand dollars led researchers and the New Zealand government to take action to identify the sources of campylobacteriosis for implementation of control strategies. Several university researchers provided evidence that the consumption of fresh rather than frozen poultry is the major cause of human campylobacteriosis in New Zealand (Eberhart-Phillips et al., 1997; Baker et al., 2006; 2007; Mullner et al., 2010). Consequently, NZFSA, in collaboration with large poultry industries, implemented poultry risk management strategies that ultimately helped to reduce the rate of campylobacteriosis to ~59% in 2008 (Sears et al., 2011). Thereafter, the campylobacteriosis notification rate has been fairly stable (157 cases per 100,000 population in 2013) (ESR, 2014). Like in other developed countries, the highest rates of campylobacteriosis have been reported in children aged between one and four years, and in the male population (281.7 and 174.1 per 100 000 population in 2013, respectively) compared to other age groups and females, respectively (ESR, 2014). Between 2008 and 2013, annually, there was an average of ~515 hospitalised campylobacteriosis cases reported ,with one death reported in 2013. It is estimated that ~25% of campylobacteriosis cases contribute to Guillain-Barré Syndrome (GBS) in New Zealand (Baker et al., 2012). Given that it has the highest rate of campylobacteriosis among developed countries, and that this disease accounts for 38.3%of enteric diseases in the country, this disease is a social and economic burden for New Zealand.

Potential routes of *Campylobacter* transmission identified in New Zealand were consumption of food from retail premises, contact with farm animals, and consumption of untreated water, which have been consistently reported as three major risk factors since 2008 (Gilpin et al., 2008; ESR, 2014). Source attribution models developed using multi-locus sequence typing (MLST) of *Campylobacter* isolates from humans, cattle, sheep, wild birds and water in New Zealand determined that poultry (50-76%) and ruminants (11-40%) were major causes of human infections (Mullner et al., 2009; Anonymous, 2014). These studies also showed that 8-15% of human cases are attributable to other sources, including water. Regional and spatial variations in

[3] Diseases included in the notifiable diseases list under the Health Act 1956 and the Tuberculosis Act 1948 of New Zealand to prevent and control the diseases. Medical officers and national laboratories are responsible for the notification of these diseases. In general, demography information, health outcome, basis of diagnosis, risk factors and some clinical management data are collected.

[4] The Institute of Environmental Science and Research (ESR) is a Government-owned Crown research institute that delivers world class knowledge, research and laboratory services to help New Zealand.

campylobacteriosis cases have also been reported in New Zealand, with high summer and autumn incidences and the greatest seasonal variations in urban North and South Islands compared to the rural North Island (Hearnden et al., 2003). The inconsistent temporal and regional patterns determined could be due to multifactorial risk factors that cause campylobacteriosis. For example, it is possible that contact with farm animals could be the major factor that differentiates campylobacteriosis cases between rural and urban settings. In fact, Mullner *et al.* (2010) found that *Campylobacter* genotypes identified in cases from urban and rural areas were associated with poultry and ruminant STs, respectively. The temporal, spatial and regional variations in campylobacteriosis cases, and the increase in ruminant and other source attribution trends of human campylobacteriosis (as shown by Mullner et al., 2009 and French and Marshall, 2014) suggest that there is a need for in-depth molecular and epidemiological combined studies on campylobacteriosis sources in New Zealand to better understand their potential transmission routes.

1.8.1.2 *Cryptosporidium* and *Giardia* spp.

Cryptosporidiosis and giardiasis are also important protozoal diseases in New Zealand, ranking among the top five most reported infectious diseases (ESR, 2013; 2014). These two protozoan diseases were included in the notifiable diseases[1] list in New Zealand in 1996. Since then, annual notifications of cryptosporidiosis cases have ranged between 584 to 977, and peaked at 1208 cases in 2001 and at 1348 cases in 2013 (ESR, 2014) (Figure 1.1). Annual notifications of giardiasis cases decreased steadily from 1998 to 2006 (2183 to 1214 cases), then the number of cases reported increased, peaking at 1985 cases in 2010, and then decreased steadily again (1729 cases in 2013) (ESR, 2014)(Figure 1.1). Cryptosporidiosis and giardiasis cases are, therefore, endemic in New Zealand with an occasional epidemic and those cases are related mostly to recreational activities. Like campylobacteriosis, cryptosporidiosis and giardiasis were reported at their highest rate in children aged one to four years (168.7 and 151.0 per 100 000 population in 2013, respectively). In addition, cryptosporidiosis rates were consistently reported to be higher in females than males, whereas there were no significant differences reported between males and females for giardiasis cases (ESR, 2014). Although hospitalisations and death rates due to these two protozoal diseases are low, these diseases have the highest notification rate among developed countries, have potential for large outbreaks, and a high economic burden (~1.5 million NZD as estimated by Snel et al., 2009) in New Zealand, which cannot be ignored for the betterment of public health.

Among the different infection routes, direct contact with farm animals and the drinking of untreated water have been implicated in human cryptosporidiosis and giardiasis in New Zealand (Snel et al., 2009; ESR, 2014). In addition, people in rural areas exposed to cattle or sheep were found to be at a higher risk of acquiring these infections, particularly during the winter and spring months (Snel et al., 2009; Spencer et al., 2012). It is possible that the higher incidence of cryptosporidiosis and giardiasis in the winter and spring months in New Zealand could be related to the combination of a higher animal density, heavy rainfall leading to runoff and contamination of the farm environment including water, and an increased number of calves (Britton et al., 2010).

In New Zealand, only a few studies have genetically characterised *Cryptosporidium* or *Giardia* isolates to understand the potential transmission routes for cryptosporidiosis and giardiasis. *C. hominis* were predominantly reported in urban areas and *C. parvum* in rural areas, suggesting that person-person transmission is common in cities and zoonotic transmission in rural places (Learmonth et al., 2004). In addition, *C. hominis* and *C. parvum* were predominantly reported in autumn and spring cryptosporidiosis cases, suggesting the attribution of calving/lambing and heavy rainfall towards the peak of cryptosporidiosis cases in spring (Learmonth et al., 2003; 2004). Occupational outbreaks associated with *Cryptosporidium* were also reported in veterinary students of Massey University in 2006 (Grinberg et al., 2011). The authors characterised *Cryptosporidium* isolates from seven stool samples at 18S rRNA and gp60 loci and determined the rare IIaA21G4R1 *C. parvum* subtypes in two isolates, indicating a point-source infection. In New Zealand, IIaA18G3R1 and IIaA19G4R1 alleles of *C. parvum* have been frequently reported in humans and in cattle samples (Anonymous, 2013), and the IIaA18G3R1 allele has been identified as waterborne cryptosporidiosis in other countries (Glaberman et al., 2002; Chalmers et al., 2005a). However, two studies reported *Giardia assemblages* A and B in human isolates from intensive cattle farming (Learmonth et al., 2003; Winkworth et al., 2008). As both A and Bassemblages could be found in cattles and humans, it is difficult to confirm whether zoonotic or anthropogenic transmission are predominant in New Zealand. Therefore, further molecular studies are anticipated to unravel the epidemiology of *Giardia* as well as *Cryptosporidium* infections in New Zealand.

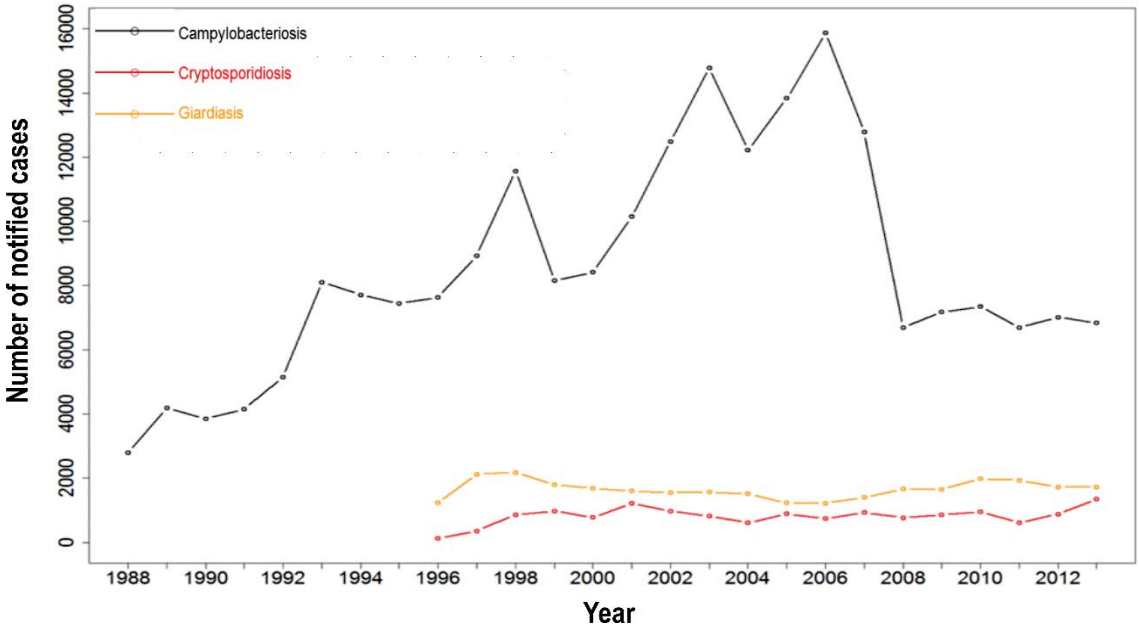


Figure 1.1: Campylobacteriosis (black), cryptosporidiosis (red) and giardiasis (orange) cases notified in people of New Zealand between the year 1988 and 2013 (ESR, 2014). Cryptosporidiosis and giardiasis were included in the notification list from 1996 only.

1.8.2 Epidemiology of *Campylobacter*, *Cryptosporidium* and *Giardia* in farmed-ruminants in New Zealand

Only very limited data are available on the epidemiology of *Campylobacter*, *Cryptosporidium* and *Giardia* in farmed ruminants, the majority being dairy cattle in New Zealand (Al Mawly et al., 2015b).

1.8.2.1 *Campylobacter* spp.

In 1989, Meanger and Marshall investigated the seasonal prevalence of *Campylobacter* in dairy cows at Massey University farms in the North Island, and found 24% (17/72), 31% (33/106), and 12% (11/95) prevalences in summer, autumn and winter respectively. Adhikari *et al.* (2004) reported *Campylobacter* in 54% (28/52; CI: 40-67%) of dairy cattle faeces collected from No. 4 dairy farm. Another survey of 185 new born calves of 24 dairy farms in the Manawatu region of the North Island showed *Campylobacter* presence in 36% (58/161) of calves from 18/24 farms (Grinberg et al., 2005). In 2008, Gilpin *et al.* reported the prevalence of *Campylobacter jejuni* in 410 dairy cattle from 36 dairy farms in the Matamata-Puako district of the North Island, and compared the isolates with 58 human *C. jejuni* isolates from the same area. They estimated *C. jejuni* prevalences of 59% (CI: 52-65%) in cattle (n=225) and 75% (CI: 68-

81%) in dairy calves (n=185) and showed the presence of indistinguishable genotypes among *C. jejuni* isolates from cattle and human sources. Another study by the same authors (Gilpin et al., 2008b) investigated the transmission route for human cases linked to seven dairy farms in the North Island and reported that 4/7 cases had most likely had contact with dairy cow faeces. Similarly, the potential reservoirs and transmission routes of human pathogenic *C. jejuni* in Ashburton (South Island) were investigated, and reported *C. jejuni* prevalences of 97.8% (89/91) and 83.9% (52/87) in dairy and beef cattle, respectively (Devane et al., 2005). Recently, the occurrence of *C. jejuni* was investigated over two years in 21 commercial dairy farms, representing three management systems, located in the Waikato region of the North Island (Rapp et al., 2014). This study reported prevalences of 55%, 49% and 54% for the herd home system, the standoff pad system and the pasture system, respectively. All these studies imply that dairy cows and calves may act as a potential source of human campylobacteriosis in New Zealand.

Only a handful of studies has determined the prevalence of zoonotic *Campylobacter* spp. in sheep and goats in New Zealand, although many studies targeted for *C. fetus,* a cause of ovine abortions (Collins et al., 1985; Dempster et al., 2011; Mannering et al., 2003). Devane *et al.* (2005) reported a *C. jejuni* prevalence of 59.8% (52/87) in sheep faeces, whereas Moriarty *et al.* (2011a) found *Campylobacter* spp. in 80.9% and 30.4% of faecal specimens from lambs (n=105) at slaughter and sheep (n=220) on pasture, respectively. A recent study reported a 30% (74/249) *Campylobacter* spp. prevalence in healthy adult dairy goats in the Waikato region of the North Island, predominantly *C. jejuni* (Rapp and Ross, 2012). These studies also indicated that both sheep and goats could represent a possible source for human infections in New Zealand.

1.8.2.2 *Cryptosporidium* and *Giardia* spp.

A limited number of studies reported the *Cryptosporidium* and *Giardia* species in farmed ruminants in New Zealand, and the majority of the studies were performed in dairy cattle. In 1987, Townsend and Lance reported *Cryptosporidium* spp. in 37% (206/550) of calves' faecal diagnostic samples submitted to the Ruakura Animal Health Laboratory between 1984 and 1986. They found the highest prevalence in calves aged between four and fourteen days. Hunt *et al.* (2000) reported *Giardia duodenalis* in 40% (286/715) of calves <8 weeks of age from farms in the Waikato and Manawatu regions. However, another study reported low *Cryptosporidium* spp. (0.6% and 8%) and *G. duodenalis* (4.5% and 10.5%) prevalences in specimens from 354 cows and 304 calves collected from 36 herds in the Waikato region, respectively (Learmonth et al., 2003).

As these authors used different detection techniques, the prevalences reported could have been varied. From the above reported prevalences of *Giardia,* it could be postulated that *G. duodenalis* are more prevalent than *Cryptosporidium* spp*.* in New Zealand cattle farms, and this indicates a need for further attention to reduce the public health risk.

Geographical differences in *Cryptosporidium* and *Giardia* spp. prevalences were also noted in New Zealand. For example, Grinberg *et al.* (2005) conducted a cross-sectional study during the winter calving season of 2002 and reported *C. parvum* in 21.1% (33/154) of calves from 10/24 farms in the Manawatu region of New Zealand. A longitudinal study conducted during the spring calving season in 2005 and 2006 in the Otago region of the South Island, conversely, reported very low (2.6%; 31/1190) *Cryptosporidium* spp. prevalences in freshly voided faecal samples on the ground of 10 dairy farms containing calves aged between 1 and 7 weeks (Winkworth et al., 2008). The variations could be due to sampling design but geographical variations could not be ignored, and further studies are warranted in New Zealand.

Recently, a countrywide survey on enteropathogens in 1283 calves from 97 dairy farms reported 18% (144/797) and 52% (223/429) *C. parvum* prevalences in 1-to-5-day-old and 9-to-21-day-old calves, respectively (Al Mawly et al., 2015a), suggesting <21 days calves as a potential source of environmental contamination in New Zealand. Mixed *Cryptosporidium* species, *C. parvum, C. bovis* and *C. andersoni*, infections in calves were also recently reported in New Zealand cattle, after employing iterative molecular characterisation at three different loci (Shrestha et al., 2014). However, only three studies characterised *Giardia* isolates in New Zealand. Learmonth *et al.* (2003) and Winkworth *et al.* (2008) analysed *Giardia* isolates from cattle faecal samples at the tpi and β-giardin loci, respectively, and determined zoonotic potential assemblages A and B in cattle. These studies did not find any cattle-specific *G. duodenalis* assemblage E, whereas assemblage E were reported in cattle from the South Island when the gdh locus was analysed (Abeywardena et al., 2012). Although studies on *Cryptosporidium* spp. and *G. duodenalis* in cattle and sheep are limited, these host species may potentially contaminate water sources and pose a public health risk in New Zealand.

Only one study has reported *Cryptosporidium* spp. and *G. duodenalis* in 28.6% and 37.1% of faecal specimens from lambs (n=105) and 3.6% and 0.0% of sheep (n=220), respectively (Moriarty et al., 2011b). Cryptosporidiosis cases in red deer calves have also been reported in New Zealand (Orr et al., 1985). No study has characterised

Cryptosporidium and *Giardia* in sheep, goats or other farmed ruminants at a molecular level. Therefore, further molecular analysis coupled with epidemiological studies of *Cryptosporidium* and *Giardia* spp. in samples from farm animals including cattle and sheep should be conducted to better understand the zoonotic potential and transmission pathways of these pathogens in a New Zealand context.

1.8.3 Survival and abundance studies of *Campylobacter, Cryptosporidium* and *Giardia* spp. in New Zealand environment

In New Zealand, few studies have been conducted to determine the survival and abundance of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. in faeces on pasture and soil. Moriarty *et al.* (2008) reported cattle shedding a median of $3{\cdot}9 \times 10^5$ C.F.U. of *C. jejuni*. Gilpin *et al.* (2009) studied the survival of *Campylobacter* spp. in 10 freshly collected cow pats deposited on pasture during summer and reported *Campylobacter* counts below the detection limits in three pats. In the other seven pats, *Campylobacter* counts dropped below detection limits within 14 days, with 90% of *Campylobacter* spp. inactivated in 2·2 days. Similarly, *Campylobacter* spp. were rapidly inactivated at higher temperatures (> 20 ^{o}C), with complete inactivation within 4 days (Moriarty et al., 2011a). Recently, 35 cows shedding up to 3.6 $\log_{10}$ median concentration of *Campylobacter* spp. were observed in two New Zealand dairy farms (Rapp et al., 2012). Likewise, Grinberg *et al.* (2005) found up to 10^6 oocysts per gram of faeces in calves aged<21days, whereas low numbers of oocysts (1–25) and cysts (1-17) per gram of faeces were reported in a survey of fresh cattle faeces on pasture (Moriarty et al., 2008). An experiment was conducted to study the survival of *C. jejuni* in four different soils: (1) Hamilton (granular); (2) Taupo (pumice); (3) Horotiu (silt loam), and (4) Waihou (allophonic) on farms (Ross and Donninson, 2006). In this study, dairy farm effluent (FDE) was applied to intact soil cores at a rate of two litres per square metre and was incubated at 10°C for up to 32 days. This study demonstrated that at least 99% of *C. jejuni* were retained in the top 5 cm of the four soils, and *C. jejuni* survived for 25 days in Hamilton and Taupo soils, 32 days in Waihou soil and for more than 32 days in Horotiu soil.

These findings clearly indicate that farmed ruminants, particularly cattle and sheep, are significant reservoirs of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. in New Zealand. These hosts might represent a significant source of water contamination and a public health risk for humans residing in farming areas. However, more

epidemiological studies are required to fill the gaps of knowledge on the epidemiology of these agents in different age groups of farmed ruminants within watershed areas.

1.8.4 Occurrences of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. in New Zealand surface water

Many rivers and streams in New Zealand are freely accessible for recreational purposes and used for drinking purposes. A few studies in New Zealand have shown that consumption of water contaminated with *Campylobacter*, *Cryptosporidium* and *Giardia* spp. occurs during recreational activities (Fraser and Cooke 1991; Duncanson et al., 2000; Hoque et al., 2002; Hearndon et al., 2003; Till et al., 2008). Little is known, however, about occurrences of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. in different types of water in New Zealand (Lechaevallier et al., 1991; Savill et al., 2001; Anonymous 2013). Lechaevallier *et al.* (1991) reported *Cryptosporidium* oocysts in 87% and *Giardia* cysts in 81% of 66 surface drinking water samples. A cross-sectional study was conducted on roof-collected rainwater supplies (n=125) in four rural Auckland districts of the North Island (Simmons et al., 2001). This study did not report *Campylobacter* and *Giardia* spp. in roof water but found *Cryptosporidium* spp. in 4% (2/50) of water supplies. *Campylobacter* spp. were isolated from 60% of river (n=30), 75% of ground (n=18), 37% of roof (n=24) and 29% of drinking (n= 24) water samples analysed (Savill et al., 2001). This study also highlighted that infiltration galleries[5] should not be used as the sole means of treating drinking water. Eyles *et al.* (2003) investigated the occurrence of *Campylobacter* spp. in the Taieri River, and reported a greater median level of *Campylobacter* numbers during the summer months, particularly in areas where intensive agricultural farming is present. The *Campylobacter* prevalence in water samples from the Ashburton River of the South Island was 55.2% (162/293) (Devane et al., 2005). Another study that investigated the microbial quality of shallow ground water in a border-strip irrigated dairy farm catchment in Canterbury of the South Island, reported *Campylobacter* spp. in 12% (16/135) of water samples (Close et al., 2008).

More recently, since 21 September 2009, a survey was conducted every trimester of the year for identifying the presence of *Cryptosporidium* and *Giardia* spp. in 16 high-risk surface water sources and four shallow bore ground water sources (Anonymous,

[5] Infiltration gallery is a horizontal drain made from open jointed or perforated pipes or block drain that is laid below the water table and collects groundwater. It can be used to collect sub-surface flow from rivers, and either withdrawn directly or pumped to a storage tank. (WHO fact sheet 2.5)

2013). From 2009 to 2013, the two Waikato Rivers were found to be contaminated with *Cryptosporidium* and *Giardia* spp. on about half of the occasions, with oocyst or cyst concentration ranging from 1-12 or 1-18 per 100L, respectively. *Cryptosporidium* and *Giardia* spp. were found in 6/20 sites analysed. *Campylobacter* spp. were also analysed in water samples from the same 20 sites, and found in 19.7% (59/300) samples from 15/20 sites analysed. This study also showed that *Cryptosporidium* spp. were most likely to be detected in September, *Giardia* spp. in May and July, and *Campylobacter* spp. in January.

Few molecular epidemiological studies of *Campylobacter* have been conducted in New Zealand using various molecular tools. Carter *et al.* (2009) characterised 244 *C. jejuni* isolates from the Taieri and Ashburton Rivers in the South Island and the Manawatu River in the North Island using MLST. They reported 88 different sequence types (STs) belonging to 13 clonal complexes (CC) and seven unique STs in New Zealand rivers. Although the majority of STs were wild bird-associated, the study also identified CC ST-61 and CC ST-42 that are associated with cattle and sheep (Carter et al., 2009). Previous studies used Penner typing and the pulse-field gel electrophoresis method to investigate *C. jejuni* isolated from water (Devane et al., 2005; Eyles et al., 2005). These studies identified indistinguishable macrorestriction profiles in river, human, and/or cattle isolates. Molecular analysis of *Cryptosporidium* and *Giardia* isolates from water has not yet been reported. However, the abovementioned epidemiological and molecular studies suggest that river water in New Zealand is a potential source of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. infection in humans, and therefore further attention is required to elucidate the ecology of *Campylobacter*, *Cryptosporidium* and *Giardia* in environmental niches in New Zealand.

1.9 Conclusion

Campylobacter, *Cryptosporidium* and *Giardia* spp. are zoonotic pathogens of significant importance to public health and the water industry worldwide, They can cause life-threatening illnesses such as Guillain-Barré syndrome and irritable bowel disease in humans, and death in extreme cases. Farmed ruminants, particularly cattle and to some extent sheep, are considered a reservoir of *Campylobacter*, *Cryptosporidium* and *Giardia* spp., and are recognised as important sources for human infections with these three pathogens. Humans acquire infections of these three pathogens through the consumption of contaminated food and water, and by direct and indirect contact with these pathogen-shedding animals, particularly on farms. Although a substantial

amount of research has been conducted to gain knowledge on the epidemiology of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. related waterborne zoonosis worldwide; there is very little information available on their prevalence and distribution in cattle and sheep in New Zealand, and in surface water destined for drinking or recreational purposes. In addition, the use of molecular technology to characterise isolates of these three pathogens will allow the zoonotic potential of these pathogens to be determined. Hence, epidemiological data coupled with molecular analysis are required to identify the sources and transmission pathways of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. infection in humans. Subsequently, these studies help in the development of appropriate and cost-effective control strategies, and may contribute to a reduction in the incidence rate of *Campylobacter*, *Cryptosporidium* and *Giardia* species in New Zealand, which is currently among the highest in developed countries.

Chapter 2

Dynamics, Source and Population Structure of *Campylobacter jejuni* Isolated from Six High-Use Recreational Rivers in New Zealand

2.1 Preamble

This chapter includes the analysis of the retrospective data on *C. jejuni* isolated from six high-use recreational rivers in the Manawatu-Wanganui region of New Zealand. Here, dynamics, source and population structure of isolated *C. jejuni* were investigated using the generalised additive model, generalised linear/logistic regression model, and the minimum spanning tree. The presence of potential zoonotic *C. jejuni* was also assessed by analysing the *C. jejuni* subtypes isolated.

2.2 Abstract

Campylobacter jejuni, a leading cause of gastroenteritis worldwide, has been frequently isolated from recreational rivers and streams in New Zealand, yet the public health significance of this is unknown. This study uses molecular tools to improve our understanding of the epidemiology and sources of *Campylobacter* in recreational rivers with the hope of preventing human infection. Epidemiological and microbiological data were collected between 2005 and 2009 from six high-use recreational rivers in the Manawatu region of the North Island. *Campylobacter* spp. and *C. jejuni* were isolated from 38% and 21% of 507 samples, respectively. High recovery of *Campylobacter* was observed at low and high river flows. After adjusting for the confounding effects of river flow, there was a significantly higher likelihood of isolating *Campylobacter* in the winter month of June compared to January (OR=57, $p<0.0001$). A high diversity of *C.*

jejuni sequence types was seen with the most commonly isolated being the ST-2381 (19 isolates, 21%), ST-1225 (8 isolates, 9%), and ST-45 (6 isolates, 7%). The water rail-associated ST-2381 was found in all rivers, while the most commonly isolated STs from human cases in New Zealand, the poultry-associated strain ST-474, was isolated only in one river. Although the majority of *Campylobacter* sequence types identified in river water were strains associated with wild birds that are rarely associated with human disease, poultry and ruminant-associated *Campylobacter* strains that are found in human infection were also identified, and these strains could present a public health risk.

2.3 Introduction

Campylobacter is a Gram-negative, thermophilic and microaerophilic bacterium that causes approximately 5-14% of diarrheal illness worldwide (Nachamkin et al., 2008; WHO, 2004). Most of the identified *Campylobacter* species can cause human campylobacteriosis but, of these, *Campylobacter jejuni* and *C. coli* are the major causes of gastroenteritis, accounting for 95% of all reported human campylobacteriosis cases (WHO, 2004). This organism naturally inhabits the intestines of warm-blooded animals including poultry, wild birds and ruminants (Stanley and Jones, 2003). Several epidemiological studies have shown these animals to be potential reservoirs for human infections (Stanley and Jones, 2003; Wong et al., 2006; French et al., 2009). These studies also indicate that animal hosts are the main sources of food and water contamination, whilst the food chain route has been shown to be the predominant infection pathway for human campylobacteriosis, particularly via poultry meat (Corry and Atabay, 2001; Woodward et al., 2005). Sporadic cases of *Campylobacter* infection have been most commonly linked to food (ESR, 2011b; Samuel et al., 2004), whereas outbreaks of campylobacteriosis have been associated with drinking contaminated water, the accidental ingestion of water during recreational activities, and the consumption of poultry meat and raw milk (Evans et al., 1996; Frost et al., 2002; Schönberg-Norio et al., 2004). Hence, each potential exposure pathway needs to be studied in detail in order to increase our understanding of *Campylobacter* transmission dynamics and thus inform the design of effective, country-specific, prevention programmes.

Campylobacteriosis is the most common zoonotic bacterial enteric disease in New Zealand, and over the last decade, this country has persistently reported one of the highest campylobacteriosis rates among developed countries (Wilson, 2005; Baker et

al., 2007b; Muellner et al., 2011). From 1980 (the first year of mandatory notification of cases) to 2006, the annual number of reported cases of campylobacteriosis increased steadily from a few hundred to 15,873 (Baker et al., 2007a; ESR, 2007; ESR, 2011b). These cases accounted for 70% of all the notified enteric diseases in New Zealand and resulted in an estimated cost of NZD 75 million per year (Scott et al., 2000; NZFSA, 2010). Around 25% of campylobacteriosis cases incurred a further disease burden due to sequelae such as Guillain-Barré syndrome with an average rate of 2.32 hospitalisations per 100 000 population per year (Baker et al., 2012; Muellner et al., 2009). An investigation into the sources of human campylobacteriosis cases identified poultry as a major source of infection (Wilson, 2005). Subsequently, in early 2007, the three major poultry suppliers implemented voluntary and regulatory interventions to reduce poultry carcasss contamination. In the two years subsequent to the interventions, there was a 59% reduction in reported human campylobacteriosis cases from 383 cases in 2006 to 157 cases per 100,000 population in 2008) (Sears et al., 2011). Nevertheless, the campylobacteriosis rate in New Zealand remains among the highest of the industrialised countries, indicating the need to understand the role of other sources of infection, including water and ruminants.

In New Zealand, thermophilic *Campylobacter* spp. have been isolated from river water, streams, lakes, ponds, runoff water, and drinking water (Koenraad et al., 1997; Till et al., 2008). Savill and co-workers (Savill et al., 2001) found that 60% of 30 recreational water sites were positive, with counts of *Campylobacter* reaching 11 most probable number (MPN[1]) 100 mL^{-1}. Between December 1998 and February 2000, a large-scale survey was conducted in 25 freshwater recreational and drinking water supply sites distributed throughout New Zealand (Till et al., 2008). This study showed that 60% of the 725 samples were positive for *Campylobacter* spp. with 48% of the positive samples identified as *C. jejuni*. Till *et al.* (2008) also estimated that 5% of campylobacteriosis cases could be attributed to recreational water. This freshwater survey and a quantitative risk assessment led to develop a new national recreational freshwater water quality guideline in 2003. However, applying these guidelines to recreational rivers is complicated because many factors, including the rate of river flow, land use, animal access to the waterways and surface runoff, could influence the bacterial risks to human health.

[1] Most probable number (MPN) is the method to estimate the concentration of viable microorganisms in a sample by means of replicate liquid broth growth in ten-fold dilutions and is particularly useful with samples that contain low concentrations of organisms (<100/g) (Blodgette, 2010)

Campylobacter spp. only replicate in animals, and although they are frequently recovered from the environmental water, they have not been demonstrated to multiply outside the host (WHO, 2004). The growth of *C. jejuni* and *C. coli* requires both thermophilic (>30 ^{o}C) and microaerophilic (<15% of O2 and CO2) atmospheric conditions (WHO, 2004). However, it has been suggested that *Campylobacter* spp. may persist in the environment by entering a viable but non-culturable (VBNC) state, and/or by forming monospecies biofilm or colonising pre-existing biofilm (Chaisowwong et al., 2011; Reuter et al., 2010). Even so, ambient temperatures, high oxygen concentrations, UV radiation, and desiccation may decrease the survival of *Campylobacter* in the environment (Rollins et al., 1986; WHO, 2004; Inglis et al., 2010). Therefore, the presence of *Campylobacter* in water most likely indicates recent faecal contamination from either a point source (e.g. meat plant effluent) or nonpoint sources (e.g. agricultural runoff).

The subtyping of *C. jejuni* has been used successfully to identify the sources of human infections. Several methodologies, including serotyping and PFGE have been developed for subtyping of *C. jejuni* (Lorenz et al., 1998; Nielsen et al., 2000). However, multilocus sequence typing has good discriminatory power and better reproducibility than other typing methods (Dingle et al., 2001). Further, the use of MLST for typing *C. jejuni* has also provided important insights into the population genetics of this organism (Dingle et al., 2001; 2005) and has helped to increase the understanding of transmission pathways of human campylobacteriosis (Wilson et al., 2008). Source attribution modelling has also been used in human campylobacteriosis cases in the Manawatu, New Zealand. Prior to intervention in the poultry industry, an estimated 70% of human cases were attributable to poultry sources. Subsequently, there has been a decrease in the proportion of human cases attributed to poultry and an increase in the proportion of cases attributable to ruminant sources, which suggest that livestock are an important source of infection, particularly in rural areas (Dingle et al., 2005; Muellner et al., 2011).

Previous investigations into *Campylobacter* in rivers showed evidence of seasonal differences in retrieving *Campylobacter* from water samples (Jones, 2001; Obiri-Danso et al., 2001). In addition, diverse sequence types of *C. jejuni* have been isolated from rivers, and the majority of sequence types identified as those associated with wild birds (Obiri-Danso et al., 2001). Here, we report the findings of a longitudinal study of *Campylobacter* in six high-use recreational rivers in the Manawatu region from 2005 to 2009. The primary objectives of this study were:

1) to assess the proportions of *Campylobacter* spp. and *C. jejuni* positive water samples from each of the study sites, and

2) to determine the potential associations between the presence of both *Campylobacter* spp. and ruminant-associated *C. jejuni* and explanatory variables such as month and site of sample collection, and river flow rates.

In addition, the population genetic structure of *C. jejuni* was assessed to explore possible animal sources of river-borne isolates, and the potential of the strains present to cause human infection.

2.4 Materials and methods

2.4.1 Sources of *Campylobacter* data

Data used for this study were obtained from the Manawatu Sentinel Surveillance Program (MSSP) New Zealand. During the MSSP, water samples were collected from six waterways identified as high-use recreational swimming sites by the Regional Council (Horizons). These waterways were the Mangapapa Stream, Woodville; the Manawatu River, Hopelands Picnic Reserve, Hopelands; Oroua River, Timona Park, Fielding; the Manawatu River, Albert Street, Palmerston North; Tokomaru River, Horseshoe Bend, Tokomaru; and Kaikokopu Stream, Himatangi Beach (Figure 2.1). The rate of river flow for each sampling occasion was obtained from the Regional Council's automatic flow recording system. There were no river flow data available for Tokomaru River and Kaikokopu Stream. In total, 507 water samples were collected over 40 months between December 2005 and April 2009.

Water samples were collected in 100 mL sterile bottles fortnightly from each site. The samples were transported in a cool box within one hour to the Molecular Epidemiology and Public Health Laboratory (mEpiLab) at the Hopkirk Research Institute, Massey University. *Campylobacter* spp. were isolated from the water samples using microaerobic culture techniques as described by Mullner *et al.,* 2009. The presence of greyish, flat and moistened colonies on blood agar was considered confirmation of a presumptive *Campylobacter*-positive sample. From each sample, up to five colonies were selected for genotyping. The presumptive *Campylobacter* isolates were speciated for *C. jejuni* using a polymerase chain reaction procedure adapted from Stucki *et al.* (1995) and described by Mullner *et al.* (2009). Those isolates confirmed as *C. jejuni* were typed by multilocus sequence typing based on the method published by Dingle *et*

al. (2001. The multilocus sequence typing alleles, sequence types, and clonal complexes (CCs) were assigned using the *Campylobacter* PubMLST database (http://pubmlst.org/campylobacter/) developed by Keith Jolley and hosted by the University of Oxford. Ruminant-associated *C. jejuni* data were obtained after running the island model as described by Wilson *et al.,* 2008.

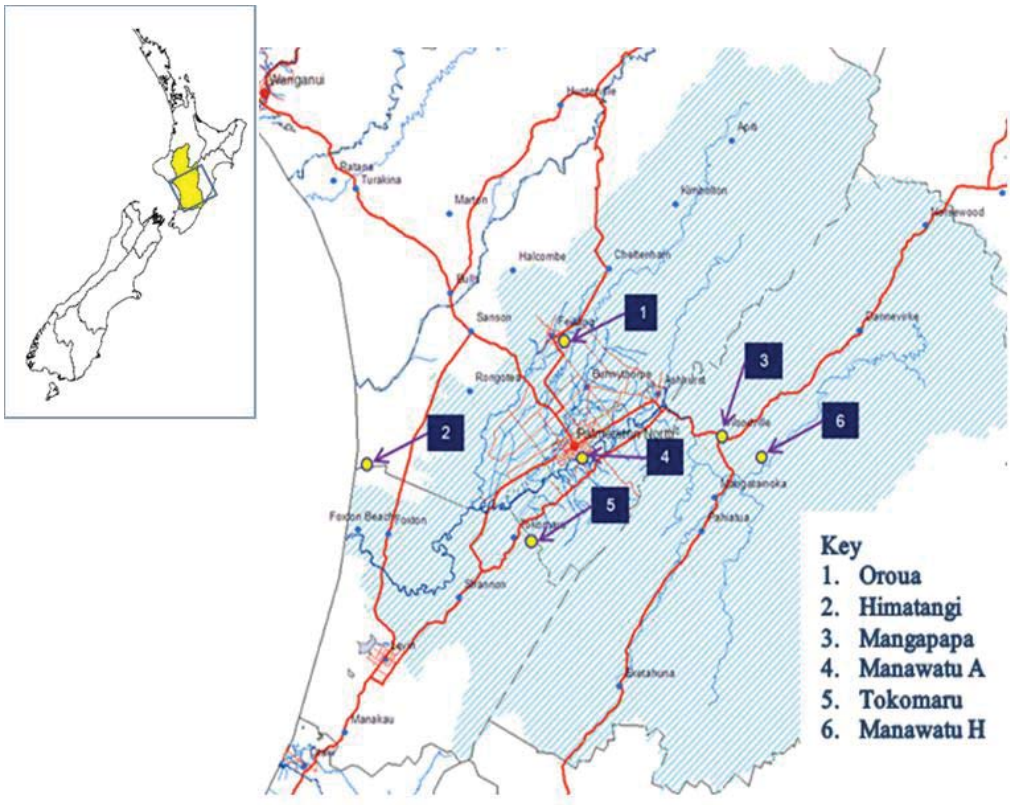


Figure 2.1: The recreational swimming water study sampling sites within the Manawatu, New Zealand. Manawatu A refers to the Albert Street section of the Manawatu River, and Manawatu H refers to the Hopelands Picnic Reserve section of the Manawatu River.

Table 2.1: Information available in the data sets used for building logistic regression models to identify the relationship between the isolation of *Campylobacter* spp. from freshwater samples and various covariates.

Variables	Type of variables	Data set A[2]	Data set B[3]
Standardized river flow rates (m^3sec^{-1})	Continuous	No	Yes
Source Site	Categorical		
Kaikokopu Stream, Himatangi		Yes	No[4]
Mangapapa Stream		Yes	Yes
Manawatu River at Hopelands		Yes	Yes
Manawatu River at Albert St		Yes	Yes
Oroua River		Yes	Yes
Tokomaru River		Yes	No[5]
Months	Ordinal		
January to December		Yes	Yes
Presumptive *Campylobacter* Presence or absence	Dichotomous	Yes	Yes

[2] Data set A contained 507 water samples.
[3] Data set B contained 344 water samples.
[4] River flow data were not available for the Himatangi sampling sites
[5] River flow data were not available for the Tokomaru sampling sites

2.4.2 Statistical analysis

Two response variables were assessed during the analyses: the presence or absence within a water sample of presumptive *Campylobacter* spp. in general, and the presence or absence of ruminant-associated *C. jejuni* in particular. The variables available within the MSSP database are shown in Table 2.1. Due to the different data available for the sampling sites, two different datasets were analysed. Dataset A comprised available information on all six sites and was used for descriptive analyses, whereas dataset B contained only the four sites with river flow data and was used for regression analyses. The descriptive, logistic and linear regression analyses were conducted in R version 2.13.1 (Hosmer and Stanley, 2000).

The mean base flow rate of four rivers at the time of sampling was not normally distributed, and it varied considerably between sites and within a site over time. Therefore, the flow rates were log-transformed to normalise the data, and the normalised flow data were then standardised to reduce the effect of between-site variations.

The Manawatu *C. jejuni* surveillance database contains the information about the isolates from food, water, and environmental sources. This information was used to estimate the ruminant-associated *C. jejuni* in our data. Firstly, the food and environmental isolates in the database were classified into 7 sources: Cattle, Sheep, Ducks, Swans or Geese, Wild birds, Dogs or Cats, and Poultry. Then, asymmetric island model (Wilson et. al 2008) was run in software R version 2.13.1 to attribute water isolates to their most likely source, and the model provided the estimates of the probability that each STs found in water came from one of the 7 sources. Finally, the probability that each STs was ruminant-associated was determined by the sum of the probabilities that the STs came from either Cattle or Sheep.

2.4.2.1 Descriptive analysis

The percentage of water samples that tested positive for presumptive *Campylobacter* and *C. jejuni* were calculated for each site along with their confidence intervals. Confidence intervals were calculated using the following formula as described by Fleiss (1981) (Section: 3.4.9.1). Pearson's Chi-square tests were also performed to determine if there is any variation in the distribution of presumptive *Campylobacter* positive samples across the sampling sites and months. A bar plot and an error bar plot were produced to show variation in the proportion of water samples positive for presumptive *Campylobacter* across the months and

standardised river flow rates, respectively. Notched box-and-whiskers plot was produced to investigate temporal variation of standardised river flow rates for four river sites.

The percentage of different clonal complexes (CCs), sequence types (STs), and the number of STs that were attributed to various likely sources were also calculated in this study. A histogram was produced to estimate the number of STs that are likely associated with ruminant sources.

2.4.2.2 Generalised linear models

We used a generalised additive model (GAM) and logistic and linear regression models to determine the relationship between covariates and response variables.

2.4.2.2.1 Generalised additive model (GAM)

The generalised additive model (GAM) was originally developed by Trevor Hastie and Robert Tibshirani, and is an extension of the generalised linear model (GLM) by combining the additive model with GLM. In GAM, the linear predictor depends linearly on unknown smooth functions of some predictor variables to maximise the quality of prediction of a dependent variable Y from various distributions, by estimating unspecific (non-parametric) functions of the predictor variables, which are "connected" to the dependent variable via a link function (Hosmer and Stanley, 2000; Menard, 2009).

GAM employs a link function to relate each linear predictor to the mean response:

$$Y = \alpha + \sum_{j=1}^{p} fj(xj) + \varepsilon \quad \text{(i)}$$

where, α is the intercept, $f_j(x_j)$ are linear terms, j is the parameters and ε is an error term. Here, GAM was used to evaluate the functional form of the relationship between:

a) a continuous covariate, standardised river flow, and the detection of presumptive *Campylobacter* spp. in water samples
b) ruminant-associated strains of *C. jejuni* and month of year adjusted for river sites.

2.4.2.2.2 Logistic regression model

The relationships of the presence of presumptive *Campylobacter* (*Y*) with the site, month and standardised river flow data (*x)* were explored using a logistic regression model as described by Hosmer and Lemeshow (2000). Initially, univariable models (Equation ii) were used to assess each covariate independently, and then those covariates with p-values of less than 0.2 were included in a multivariable model (Equation iii). The multivariable regression model was also used to control for the confounding effects of river flow rate data.

$$Y = f(x) = \beta_0 + \beta_{(\text{i}=month\ or\ site\ \text{or}\ river\ flow\ \ rate)} * x_i + \varepsilon \ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(ii)}$$

$$Y = f(x) = \beta_0 + \beta_{(\text{i}=month)} * x_i + \beta_{(\text{j}=site)} * x_j + \beta_{(\text{k}=river\ flow\ rate)} * x_k\ \varepsilon \ldots\ldots \text{(iii)}$$

where, Y = logit(p) is the logit of the response variable, β_o is the intercept, β_i is the parameter of $x_{i/j/k}$ variable in the model, and ε are residuals.

Two multivariable logistic regression models were built which differed in the form of the month used. In model A, a second-degree polynomial term of the month was fitted to identify temporal relationships present with presumptive *Campylobacter* species. In model B, the month was used as a categorical variable to detect the likelihood of isolating presumptive *Campylobacter* spp. in each month. The output of model A was used to plot a probability function curve as described by Hosmer and Lemeshow (42). For producing the curve, the following formula was used:

$$P = \frac{1}{1+e^{-(\beta o + \sum \beta_j X_j)}} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(iv)}$$

where, P = probability of isolating *Campylobacter*, e is the base of the natural logarithm, β_o is an intercept and β_j is the parameter for the variables X_j in the model A.

The Akaike information criterion (AIC) was used to compare relative fit of the two multivariate logistic regression models. Overdispersion in the data used for the models was also assessed by calculating ĉ to avoid underestimation of the variance of parameter estimates. The ĉ is calculated by dividing deviance statistics (D) per covariate pattern by degrees of freedom (df). If ĉ is reasonably close to one that is >0.95 to <1.10, the data are not underdispersed or overdispersed (Hosmer and Stanley, 2000; Menard, 2009).

2.4.2.2.3 Linear regression model

Similarly, linear regression models were fitted consecutively for each covariate (month, river site and river flow rates) followed by multivariable analysis to identify associations of these variables with 'ruminant-associated' *C. jejuni*. Rather than classifying STs as being ruminant-associated using an arbitrary cut-off, instead, the probability that each STs was ruminant-associated directly as the outcome variable of our linear regression, after having transformed it using logit, were used. The logit transformations of these proportions of *C. jejuni* attributable to ruminant origin were used as a response variable in the model.

2.4.3 Population structure analysis

Phylogenetic relationships between the *C. jejuni* isolated were described using a Kruskal's algorithms based Minimum Spanning Tree in BioNumerics version 6.1 software (Applied Maths NV Sint-Martens-Latem, Belgium).

2.5 Results

2.5.1 Descriptive analysis

Thirty-eight per cent (192 of 507; 95% CI = 33% to 42%) of water samples tested positive for presumptive *Campylobacter* spp. and ~21% of samples (103 of 507; 95% CI = 17% to 24%) were confirmed by PCR as *C. jejuni*. The breakdown of positive samples across each of the six study sites are shown in Table 2.2. The proportion of presumptive *Campylobacter*-positive samples were significantly different between months (X^2 = 58.40, df = 11, p <0.0001) and between sites (X^2 = 29.80, df = 5, p<0.0001).

Figure 2.2 shows that a higher proportion of presumptive *Campylobacter*-positive samples were found in the winter months (June, July and August) compared to the summer months (December, January and February). Similarly, Figure 2.3 illustrates that river flow rates were higher in winter months than summer months, with the greatest monthly range of flow rates being seen in March, i.e. in early autumn. Our data showed higher proportions of presumptive *Campylobacter*-positive samples when standardised river flows were between the third and fourth deciles, and above the fifth decile compared to first decile (Figure 2.4).

A 91 of 116 (77%) of *C. jejuni* PCR isolates from 103 samples were fully typed using MLST. In total, 14 different clonal complexes (CC) and 51 different STs were

identified. Approximately 49.5% (45/91) of the total *C. jejuni* isolates could not be assigned to a clonal complex (U/A). Of the recognised clonal complexes and sequence types, the most dominant genotypes are shown in Table 2.3. In this study, 70/91 STs were attributed to various sources such as wild birds, cattle, sheep and poultry (Wilson et al., 2008). Only 11/70 STs were found likely associated source of ruminant when arbitrary cut-off of 50% was used to define the ruminant-associated STs (Figure 2.5).

Table 2.2: Total number of water samples collected from six recreational river sites in the Manawatu region of New Zealand, and the number and percentage of presumptive *Campylobacter* spp. and *C. jejuni* in those samples.

Source sites	**No. of water samples**	**No. of presumptive *Campylobacter* (%)[6]**	**No. of *C. jejuni* (%)[7]**
Himatangi[a]	85	26 (30.6)	12 (14.1)
Mangapapa	86	47 (54.7)	30 (34.9)
Manawatu A[b]	87	46 (52.9)	26 (32.2)
Manawatu H[c]	85	25 (29.4)	15 (17.6)
Oroua	86	29 (33.7)	14 (16.3)
Tokomaru	78	19 (24.4)	6 (7.7)
Total	**507**	**192 (37.9)**	**103 (20.71)**

[a] Kaikokopu Stream, Himatangi [b] Manawatu River at Albert Street [c] Manawatu River at Hopelands

[6] A sample was considered presumptive *Campylobacter* spp. positive when samples cultured on mCCDA and BA produced colonies with typical *Campylobacter* morphology.

[7] A sample was considered *C. jejuni* positive when at least one isolate from that sample was confirmed as *C. jejuni* by PCR.

Table 2.3: Clonal complexes (CC), sequence types (ST), seven housekeeping genes, total number and relative frequency of *C. jejuni* isolated from six recreational rivers. Blue, green, red colour respectively denotes the first, second and third highest frequency of sequence types obtained in this study.

CC	ST	House-keeping genes							Total no.	Relative frequency
		aspA[6]	glnA[6]	gltA[6]	glyA[6]	pgm[6]	tkt[6]	uncA[6]		
21	50	2	1	12	3	2	1	5	1	1.10%
	422	2	1	5	3	2	5	5	3	3.30%
	3610	2	1	5	88	2	11	5	1	1.10%
42	42	1	2	3	4	5	9	3	1	1.10%
	3676	1	307	3	4	5	9	3	1	1.10%
45	25	4	7	10	1	1	7	1	1	1.10%
	45	4	7	10	4	1	7	1	6	6.59%
	137	4	7	10	4	42	7	1	1	1.10%
	3802	4	319	10	4	1	7	1	1	1.10%
48	38	2	4	2	2	6	1	5	1	1.10%
	474	2	4	1	2	2	1	5	2	2.20%
61	61	1	4	2	2	6	3	17	1	1.10%
177	177	17	2	8	5	8	2	4	2	2.20%
354	1517	8	10	149	2	11	12	6	1	1.10%
403	2026	10	1	16	19	10	5	7	1	1.10%
677	677	10	81	50	99	120	76	52	1	1.10%
692	2584	2	1	57	26	127	29	35	1	1.10%
	3659	37	52	57	26	127	29	1	2	2.20%
	3664	37	52	4	26	127	29	23	1	1.10%
1034	694	2	59	4	105	126	25	23	1	1.10%
	2378	2	15	4	48	356	25	23	1	1.10%
1275	1223	27	33	22	49	43	9	31	1	1.10%
	1225	27	33	22	49	43	7	31	8	8.79%
	3657	27	33	22	104	134	7	31	1	1.10%
	3661	27	33	22	49	134	7	31	1	1.10%
	3662	27	33	22	49	43	110	31	3	3.30%
	3674	27	33	22	49	43	350	31	1	1.10%

Contd.....

....contd

CC	ST	House-keeping genes							Total no.	Relative frequency
		aspA[6]	glnA[6]	gltA[6]	glyA[6]	aspA[6]	glnA[6]	gltA[6]		
U/A[7]	436	7	21	5	62	4	61	44	1	1.10%
	526	2	15	4	27	13	80	23	1	1.10%
	992	2	59	4	27	126	29	23	1	1.10%
	996	2	29	84	48	131	25	57	1	1.10%
	1030	37	4	4	48	13	25	57	1	1.10%
	2347	2	4	4	105	10	25	57	1	1.10%
	2354	37	4	4	48	13	25	23	1	1.10%
	2381	175	251	216	282	359	293	102	19	20.88%
	2619	191	251	216	282	359	293	214	2	2.20%
	3538	47	2	4	2	6	5	17	1	1.10%
	3640	1	6	5	4	261	7	3	3	3.30%
	3655	1	6	5	282	261	7	3	2	2.20%
	3656	175	251	216	282	359	293	3	1	1.10%
	3658	1	295	216	282	359	293	102	1	1.10%
	3660	192	295	216	282	359	293	102	1	1.10%
	3663	175	6	216	282	261	7	3	2	2.20%
	3672	236	306	254	339	433	349	255	1	1.10%
	3673	175	6	216	4	434	7	3	1	1.10%
	3675	237	2	254	340	435	349	256	1	1.10%
	3800	175	6	5	282	261	7	262	1	1.10%
	3801	175	318	216	282	359	293	102	1	1.10%
	3803	27	8	34	6	39	88	3	1	1.10%
Total									91	100.00%

[6]*aspA*- aspartase; *glnA*- glutamine synthetase; *gltA*- citrate synthase;

glyA- serine hydroxy methyl transferase; *pgm*- phosphor-glucomutase,

t*kt*- transketolase; and uncA- ATP synthase alpha subunit

[7]Unidefined

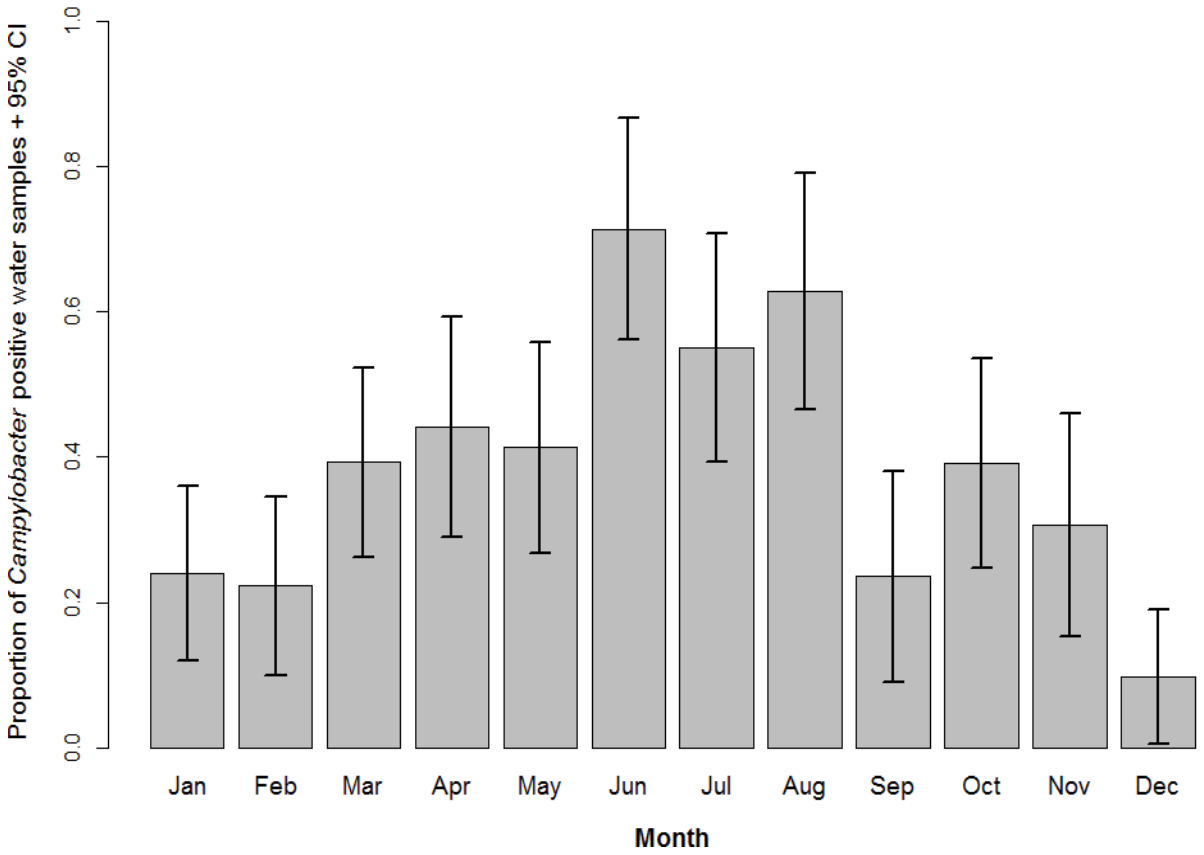


Figure 2.2: A bar plot showing the monthly variations in presumptive *Campylobacter* spp. positive water samples from six recreational water sites in the Manawatu region, combining the data collected between December 2005 and April 2009. The vertical lines represent the 95% confidence intervals.

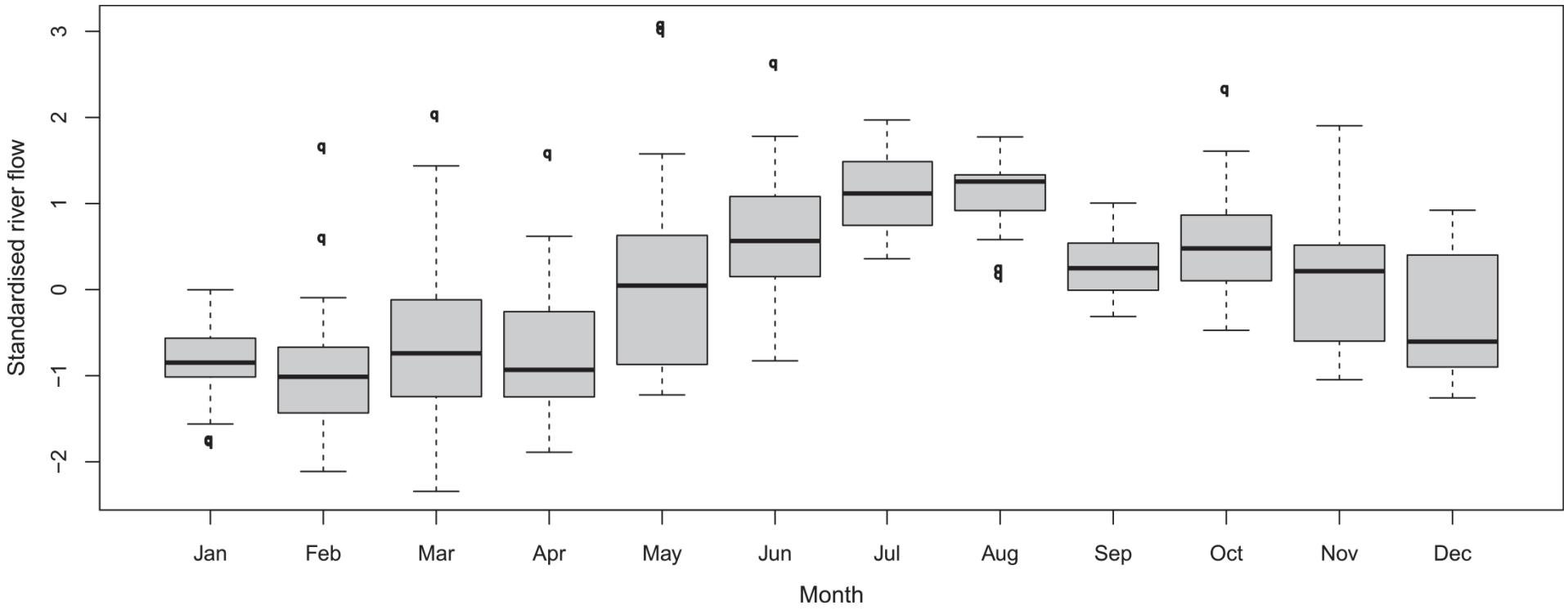


Figure 2.3: Notched Box-and-whiskers plot showing the temporal variation of standardised river flow rates (cubic meter per second) for four recreational water sites in the Manawatu between December 2005 and April 2009. Top and bottom of the box represent 25th and 75th percentile, a notch with a dark line in the box shows median value, and the two ends of the whiskers depict minimum and maximum value of river flow rates. Round dots depict the outliers of the data.

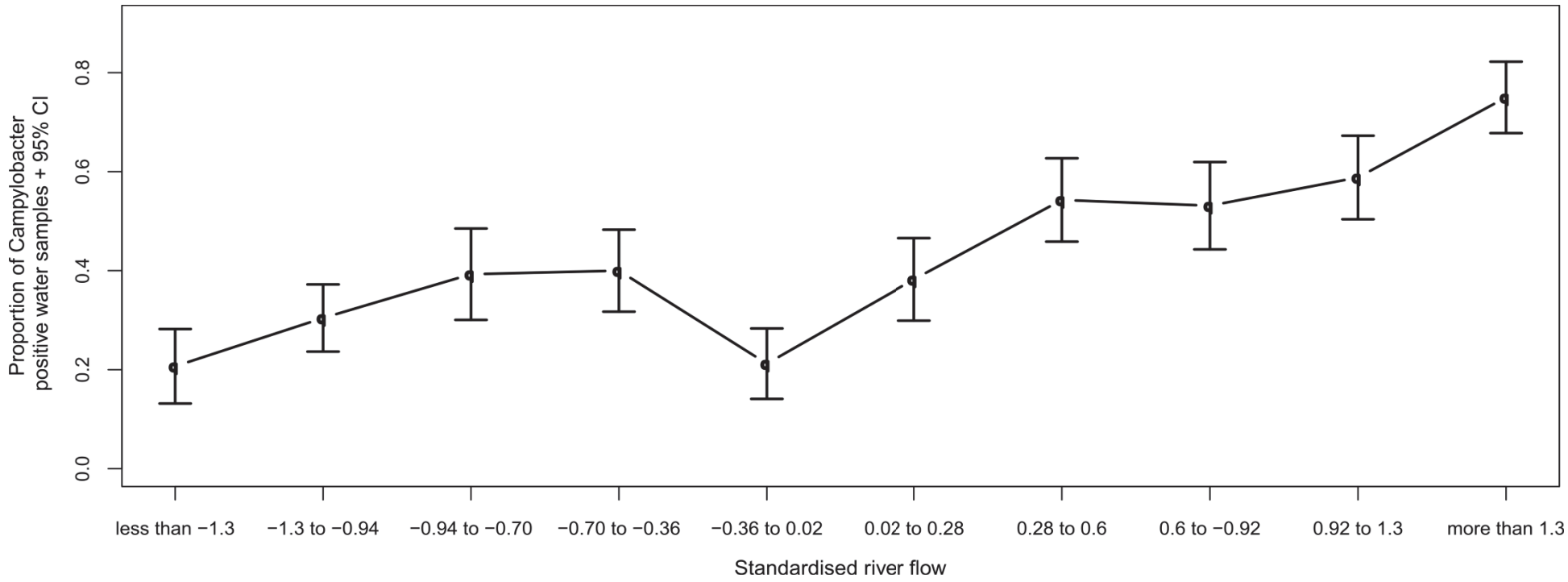


Figure 2.4: Error bar plot illustrating the variation in the proportion of presumptive *Campylobacter* spp. positive water samples (red colour dot) across each decile of the standardised river flow rates for the four recreational water sites. The vertical lines represent the 95% confidence intervals.

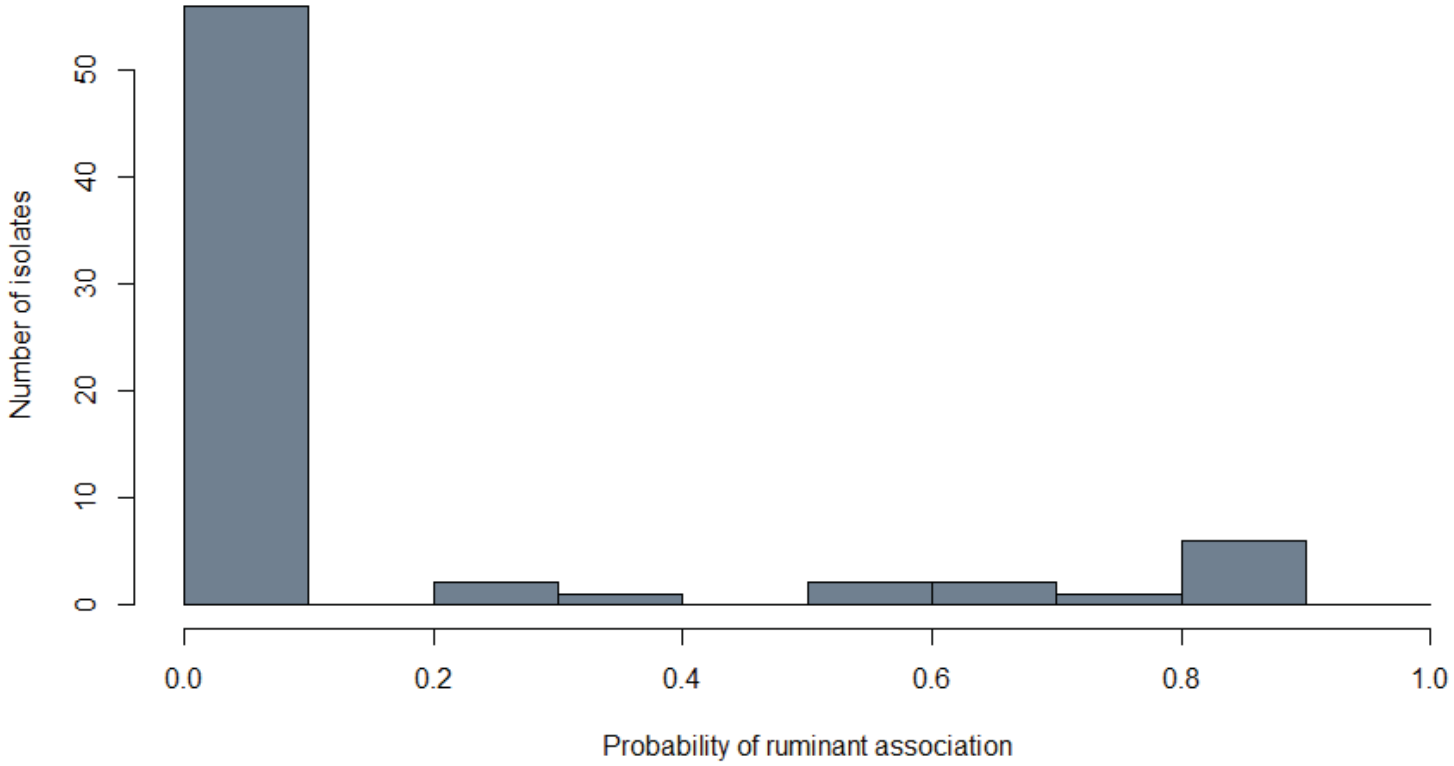


Figure 2.5: A histogram depicting the number of isolates that were attributed to the probability of being ruminant-associated *C. jejuni.*

2.5.2 Relationship of presumptive *Campylobacter* spp. with standardised river flow, month, and river sites

An exploration of the association between the response variable, isolated *Campylobacter* spp., and standardised river flow using the generalised additive model indicated a non-linear relationship (p=0.008) when adjusted for the month and the sampling sites (Figure 2.6). Table 4 shows the results of the two multivariable logistic regression models that fitted the datasets A and B. The two models differed only in the form of the month used for fitting the models. Model A used the second-degree polynomial term for a month while model B used the categorical form of the month. Of the two models, model B fitted better than model A (AIC-A=397 and AIC-B=388). The models were also not over-dispersed ($\hat{c}$ = 1.01).

Both models showed a significant and a positive association between the response variable and standardised river flow. This association implies that for every one-unit increase in the standardised river flow rate, a water sample has 1.77 times the odds of being presumptive for *Campylobacter*. A one-unit increase in the standardised river flow level is a one standard deviation increase in the un-normalised variable (i.e. one standard deviation of the log-transformed data). On examining the river sites variable, both models have indicated there is a higher chance of detecting the presumptive *Campylobacter* in water samples from the Mangapapa Stream than from other sites.

With respect to the effect of month upon the probability of isolating *Campylobacter* spp., model A shows a significantly higher likelihood of isolating *Campylobacter* in the winter month of June compared to January (OR=57, p<0.0001), and model B shows the second-degree polynomial form of month was found to be significant (p<0.0001). The latter results imply that there was a quadratic relationship between these two variables. The probability function curves in Figure 6 portray the same trend for each site, with increased probabilities of isolation during the autumn and winter months April to August. Figure 2.7 also highlights the differences between sites, with lower probabilities of isolating *Campylobacter* from the Manawatu River at Hopelands and the Oroua River compared to the Mangapapa Stream and Manawatu River at Albert Street.

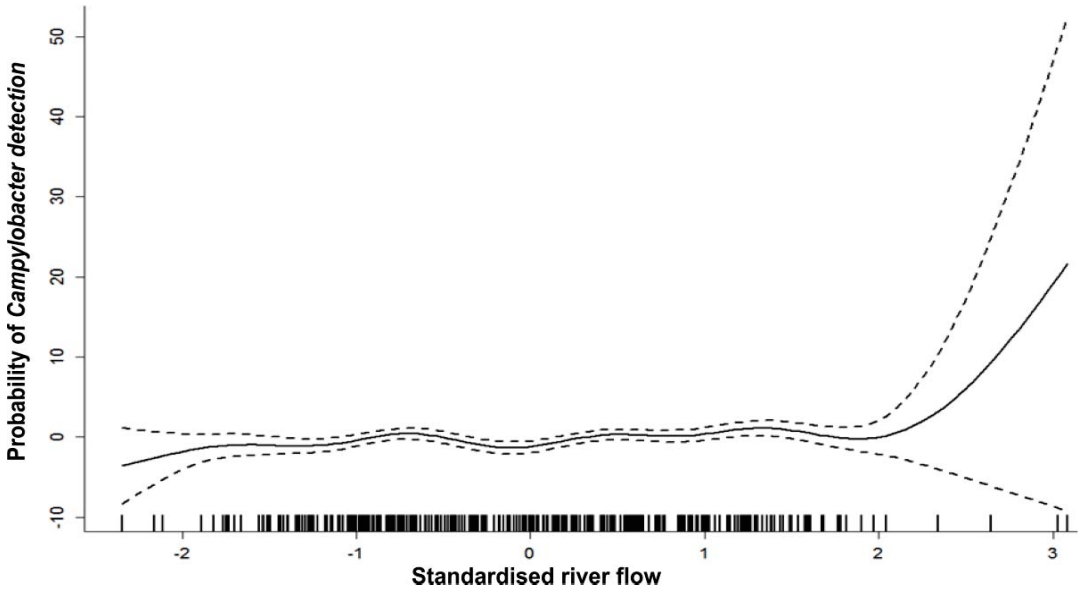


Figure 2.6: Generalised additive model plots demonstrating the relationship between standardised river flow and the response (the presence of *Campylobacter* spp. in the water samples) when adjusted for sampling sites and sampling months. The solid line shows the fitted model and the dashed lines represent the 95% confidence intervals. The rungs at the X-axis indicate the individual data points.

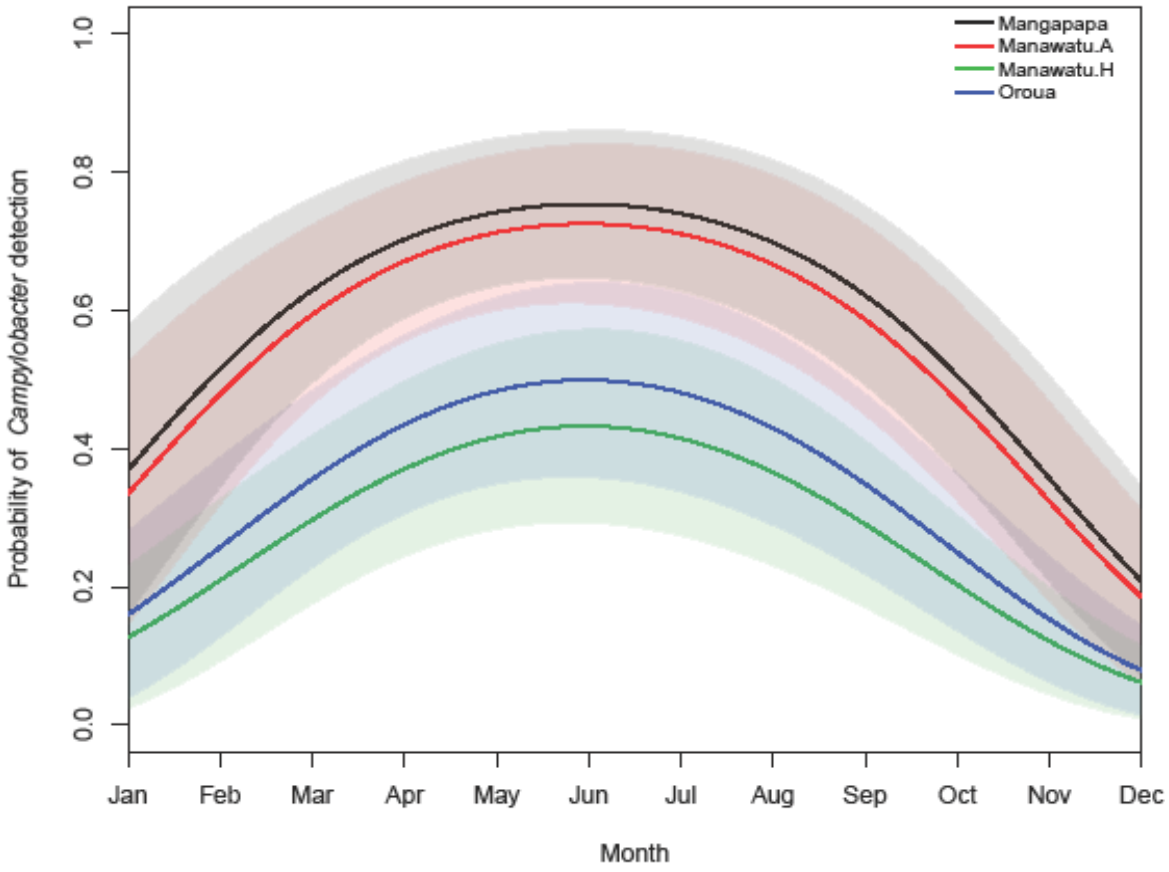


Figure 2.7: A model-fitted relationship of the probability of detecting presumptive *Campylobacter*-positive water samples in each month from four rivers, adjusted for the site and river flow. The solid coloured line with shade is the mean probability and 95% confidence interval. Manawatu A refers to the Albert Street section of the Manawatu River, and Manawatu H refers to the Hopelands Picnic Reserve section of the Manawatu River.

Table 2.4: Results of two multivariate logistic regression models (A and B) of culturing presumptive *Campylobacter* spp. from recreational water samples; the two models differed only in the measures of month use that were fitted.

Variables and parameters	Model A			Model B				Chi-square (p-value)
	Polynomial (Month^2)[8]			Factor (Month)[9]				
	β[10]	SE[11]	p-value	β[10]	SE[11]	p-value	Odds Ratio with 95% CI	
Intercept	-1.28	0.25	0.22	0.31	0.54	0.57	0.74 (0.24 - 2.06)	
Standardised flow rate (m3/s)	0.53	0.16	<0.000	0.57	0.18	<0.000	1.77 (1.25-2.56)	
Rivers								19.7 (0.0001)
Mangapapa		Ref[12]				Ref[12]		
Manawatu at Hopelands	-1.39	0.36	<0.000	-1.50	0.38	<0.000	0.22 (0.1-0.46)	
Manawatu at Albert St	-0.15	0.34	0.67	-0.15	0.35	0.67	0.86 (0.43-1.72)	
Oroua	-1.12	0.35	<0.000	-1.21	0.37	<0.000	0.3 (0.14-0.61)	
Months								68.3 (<0.0001)
Month^1	0.80	2.82	0.13	-	-	-	-	
Month^2	-0.07	2.78	<0.000	-	-	-	-	
January	-	-	-			Ref[12]		
February	-	-	-	-0.21	0.70	0.76	0.81 (0.20-3.22)	
March	-	-	-	0.89	0.59	0.13	2.43 (0.78-8.25)	
April	-	-	-	1.67	0.62	<0.000	5.33 (1.63-19.19)	
May	-	-	-	1.09	0.63	0.08	2.97 (0.89-10.70)	
June	-	-	-	4.05	1.16	<0.000	57.41(8.37-1187.95)	
July	-	-	-	1.00	0.73	0.17	2.71 (0.68-11.81)	
August	-	-	-	0.99	0.73	0.18	2.68 (0.66-11.84)	
September	-	-	-	-0.33	0.71	0.64	0.72 (0.18-2.93)	
October	-	-	-	0.50	0.65	0.44	1.64 (0.47-6.12)	
November	-	-	-	0.38	0.68	0.57	1.47 (0.39-5.73)	
December	-	-	-	-1.11	0.80	0.17	0.33 (0.06-1.51)	

[8]Model A fitted the polynomial form of the month; [9]Model B fitted the categorical form of the month; [10]β = Co-efficient; [11]SE = Standard error of co-efficient; [12] Ref = Reference.

2.5.3 Relationship of ruminant-associated *Campylobacter jejuni* with standardised river flow, month, and river sites

In the univariable linear regression model, no significant relationship was observed between the probability of isolating a ruminant-associated[8] *C. jejuni* and standardised river flow rate (p= 0.38) or month (p= 0.18). However, there were significant differences between sites (p= 0.005). Pair-wise interactions were also investigated between these variables. However, no relationship was established. Multivariable analysis showed that ruminant-associated *C. jejuni* were more likely to be found in the Mangapapa Stream (OR: 7.5; 95% CI: 1.03-54.4) than the Manawatu River at Hopelands, when adjusted for river flow and month. There was no significant difference in the likelihood of ruminant-associated samples between the Oroua River and the Manawatu River at Hopelands (OR: 0.36; 95% CI: 0.04-3.6). A weak quadratic relationship (p= 0.06) of ruminant-associated *C. jejuni* was found with a second-degree polynomial term of the month. However, a closer examination of the relationship with month using a generalised additive model did not show evidence of a non-linear relationship with months (p= 0.24), when adjusted for the sampling site and river flow (Figure 2.8).

2.5.4 Population structure analysis

Figure 2.9 is a minimum spanning tree showing the genetic relationships between STs of *C. jejuni* isolates from natural, recreational water sites. Four major clusters were found to be related to each other. A distinct larger pie with adjoined smaller pies is a wild bird cluster with the majority of them being ST- 2381, while other smaller clusters of ST-45 and ST-1225 were related to human and poultry (http://pubmlst.org/campylobacter/).

In addition, there is a widely dispersed cluster of many single isolates of different sequence types. Some of these STs have been found in both animals and humans, for example, ST-42 and ST-61. Not all the STs were isolated from all rivers. For example, ST 2381 was found in all rivers, while ST-61 was found only in the Kaikokopu Stream, and ST-474 and ST-42 were present only in the Mangapapa Stream. The wild bird associated STs were found more frequently in the Mangapapa

[8] Ruiminant-associated *C. jejuni* were determined by combining cattle and sheep attribution for that STs using an island model as described by Wilson *et al.,* 2008.

Stream (10 of 26) and less frequently in the Manawatu River at Alberta street (2 of 26).

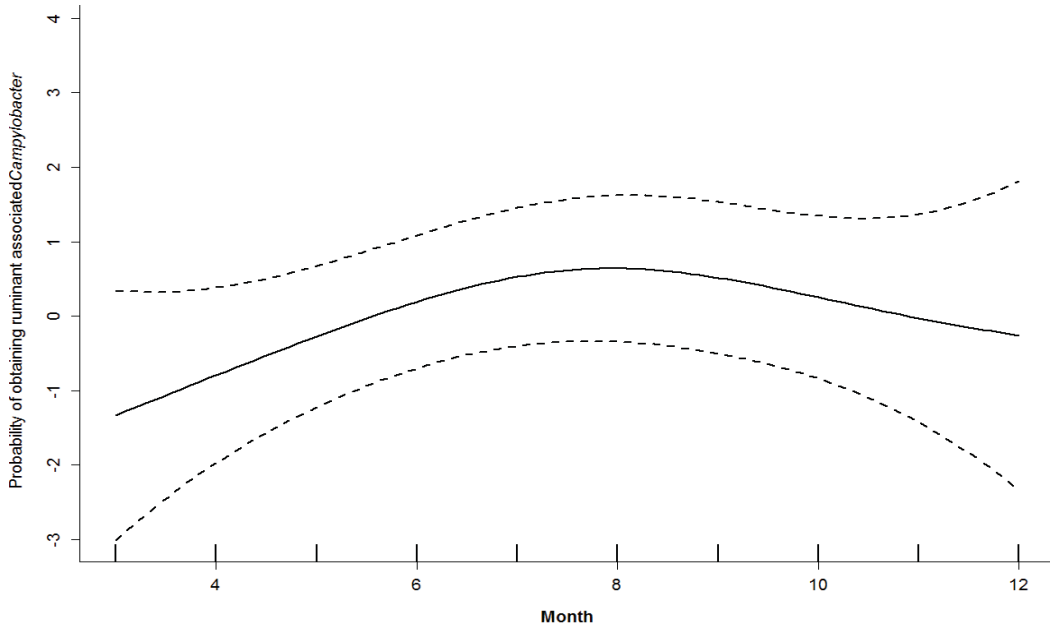


Figure 2.8: Generalised additive model plots demonstrating the relationship between sampling months and the response (the probability of obtaining ruminant-associated *Campylobacter* in the water samples) when adjusted for river flow and sampling months. The solid line shows the fitted model and the dashed lines represent the 95% confidence intervals. Each line and number at the X-axis denote the month from March (3) to December (12).

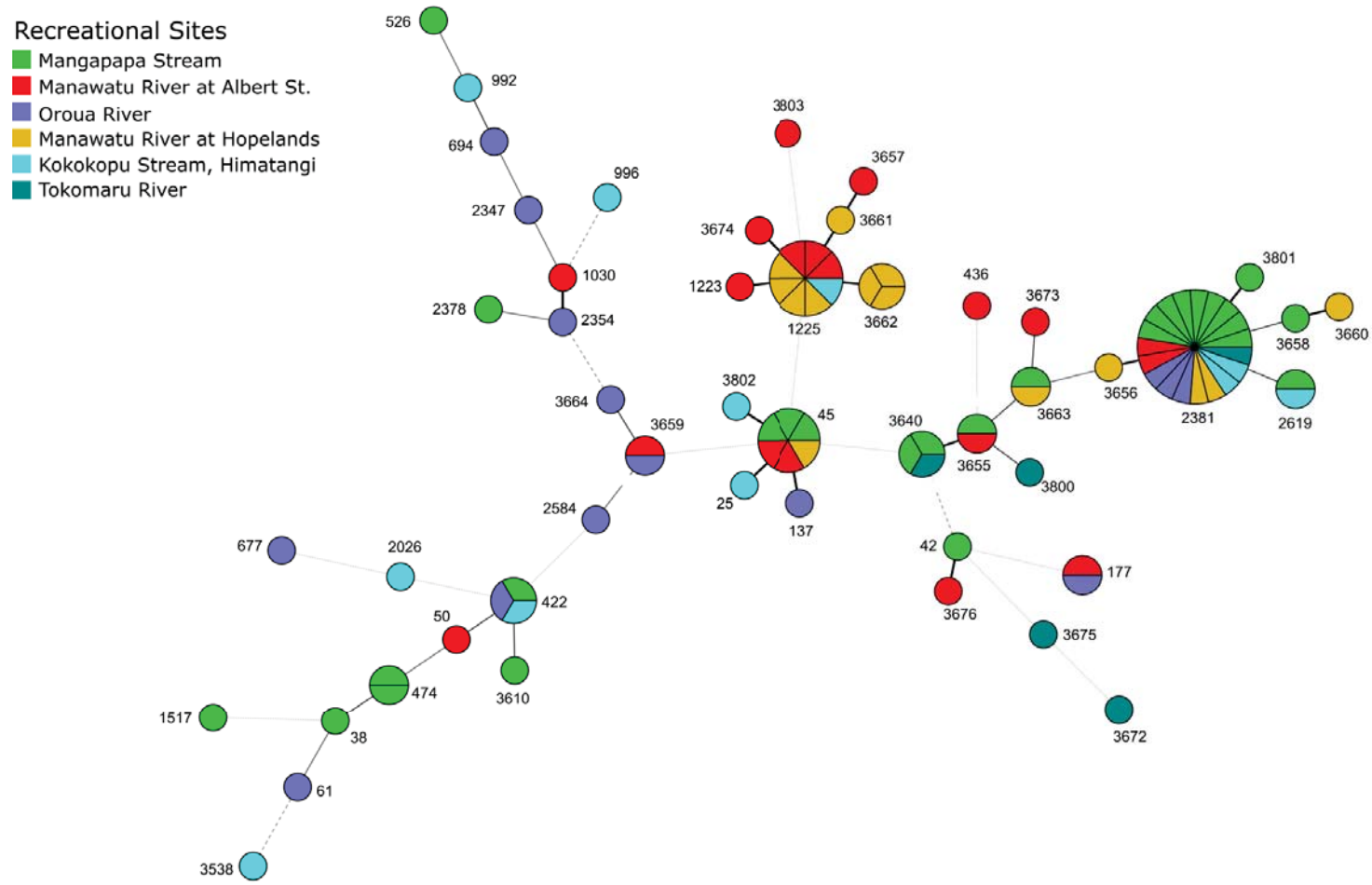


Figure 2.9: A minimum spanning tree of subtypes of *C. jejuni* found in the six recreational rivers in the Manawatu region of New Zealand. Each pie chart represents one subtype and the colours represent different river sites. Larger pie charts indicate a greater number of isolates of that subtype present. The thicker connecting lines show greater similarities between subtypes. For example, a thick black solid line between ST 2381 and ST 3656 illustrate that they are different in only one allele of seven housekeeping genes.

2.6 Discussion

Data were collected from a three-year longitudinal study in the Manawatu-Wanganui region to explore the temporal and spatial patterns of isolations of *Campylobacter* spp. and *C. jejuni* from river water, as well as the genetic population structure of *C. jejuni* in recreational water. The presence of *Campylobacter* spp. and *C. jejuni* in water samples from six recreational river sites (Table 2.1) shows the presence of faecal contamination in the rivers. Although genetic studies revealed that wild birds were the major sources of *C. jejuni* in river water, the presence of human-associated (ST-42) and ruminant-associated (ST-61) strains indicate the possible dissemination of waterborne diseases via the recreational use of rivers, and these risks to human health are highest in the winter months and at times of high river flow.

The results of studies in other countries show wide variation in the isolation of *Campylobacter* spp. from recreational river water. Some of these variations are due to differences in the isolation methods used and some are related to differences in the river catchments. In New Zealand, *Campylobacter* spp. isolation rates of 60% have been reported from recreational water (Till et al., 2008; Savill et al., 2001); whilst overseas studies report between 0 to 87.5% positive samples (Van Dyke et al., 2010; Kemp et al., 2005; Moore et al., 2001). In our study, *Campylobacter* spp. were found in 38% and *C. jejuni* in 21% of the water samples (n=507). This isolation frequency is relatively low compared with the previous freshwater studies (60%) in New Zealand by Savill *et al.* (2001) and Till *et al.* (2008). These two studies were conducted in different geographic regions between 1998 and 2000, before the implementation of national freshwater recreation water quality guidelines (2003). These guidelines have helped in the implementation of fencing land and restricting stock grazing around the waterways to reduce direct faecal deposition into the water. These actions could be an influencing factor in finding the different isolation frequencies between their studies and ours (Federighi et al., 1998; Abulreesh et al., 2006). In addition, both Savill's and Till's studies utilised a most probable number method to detect and quantify the *Campylobacter* spp. in water samples while in our study, *Campylobacter* spp. was detected using selective media culture and PCR techniques (Till et al., 2008; Savill et al., 2001). Moore *et al.* (2001) indicated that direct PCR assay is a better method for detecting *Campylobacter* spp. in water than traditional culture techniques due to its ability to detect viable but non-culturable (VBNC) *Campylobacter* in water. It is important to

detect organisms in a VBNC state as they have been demonstrated to be able to revert to a culturable pathogenic state after 30 days of incubation in microcosm water (Federighi et al., 1998) and thus may present a potential risk factor for humans.

The occurrences of *Campylobacter* spp. in freshwater samples are relatively low in this study compared with previous studies conducted in New Zealand and our analysis has shown a marked monthly seasonal variation with a distinct peak in the winter months, between June and August (Figure 2.2). The regression analyses also showed a higher likelihood of obtaining *Campylobacter* spp. in June compared to January. This result is similar to that reported by Obiri-Danso *et al.* (2001) and Abulreesh *et al.* (2006) who demonstrated a higher prevalence of *Campylobacter* spp. in winter months in UK rivers. These seasonal differences in *Campylobacter* spp. isolation could be related to factors such as larger amounts of rainfall increasing the agricultural runoff during winter months (Roig et al., 2011), and the inability of *Campylobacter* spp. to survive in water in the summer months due to greater levels of UV radiation (Obiri-Danso et al., 2001). Therefore, it is likely that other sources are the cause of human campylobacteriosis during summer as the peaks of *Campylobacter* spp. detected in our study are discordant with the seasonal peak of human campylobacteriosis cases, which is in summer, in New Zealand (ESR, 2011b; Eyles et al., 2003). There was no relationship established between ruminant-associated *C. jejuni* and a month, although weak association was shown by regression analysis, and is related to low study power (N=70).

Heavy rainfall events initiate agricultural runoff, which leads to water contamination and increased river flow (Roig et al., 2011). In New Zealand, it is evident that the heaviest rainfall occurs during winter months (NIWA, 2010), which would account for the higher river flows seen during those months (Figure 2.3). Our analyses suggest that the likelihood of retrieving *Campylobacter* spp. positive water samples are mostly higher when river has flood flow rather than base flow, and runoff may be the cause of *Campylobacter* spp. contamination in water. On the contrary, Eyles *et al.* (2003) reported two main peaks of *Campylobacter* flux: one during a high flow with moderate *Campylobacter* levels and the other when the flow is low but with high *Campylobacter* levels. They suggested that the abundance of *Campylobacter* spp. in water during summer when river flow is low could be due to continuous faecal contamination of water. Nevertheless, it implies that high river flow likely increased *Campylobacter* spp. contamination in river water.

Our study also identified that there were significant differences in the probability of obtaining *Campylobacter* spp. positive samples between different sampling sites. The lowest occurrences of *Campylobacter* spp. were found in Tokomaru River and the highest in the Mangapapa Stream. Other researchers have also observed the variation in *Campylobacter* spp. presence between locations, and have suggested that these variations are attributed to differences between the catchments. For example, Kemp *et al.* (2005) examined the possible risk factors for the presence of *Campylobacter* spp. in water and found that soil types and farm types in the catchments are significant contributors to the differences. A large-scale freshwater study conducted in New Zealand over 15 months reported higher concentrations of *Campylobacter* spp. in water obtained from the catchments containing ruminants than from other catchments (Till et al., 2008). Among our study sites, more than 60% of the land in the Mangapapa catchment had been utilised for sheep, beef, and dairy farming while the Tokomaru catchment contained more bush than farmland (Agribase™, Assure Quality 2012). In addition, there was also a higher likelihood of finding the ruminant-associated *Campylobacter* in the Mangapapa Stream (OR: 7.5). This supports the hypothesis that land use within a river catchment needs to be considered when conducting risk assessments and fitting prediction models for the presence of *Campylobacter* spp. in water.

In this study, 118 species-specific *C. jejuni* PCR-positive isolates from the six recreational waters were assigned to sequence types using multilocus sequence typing. Our results indicate the presence of diverse *C. jejuni* subtypes in recreational water. The majority of these isolates (51.6%) were assigned to unknown clonal complexes that also contain ST-2381, the most prevalent sequence types (21% of 91) recovered from all rivers. This result provides evidence that the ST-2381 is prevalent in the majority of New Zealand's river water. This subtype has only been isolated from the native wild birds pukeko and takahe but not from human cases of campylobacteriosis, implying that ST-2381 is possibly non-pathogenic to humans (Motarjemi, 2014; French et al., 2011). In this study, 69% of total sequence types recovered are STs that are known to be associated with wild birds (http://pubmlst.org/ campylobacter/), which provides evidence that recreational water contamination is often due to the wildlife inhabiting the area. However, the presence of *C. jejuni* belonging to CC-21, CC-42, CC-45, CC-48, CC-61, CC-177, CC-1517, CC-2026, and CC-677 represents the possibility of zoonotic transmission because many sequence types of these clonal-complexes have been isolated from human campylobacteriosis outbreaks (http://pubmlst.org/ campylobacter/).

Interrogation of the PubMLST website also shows the evidence of these clonal-complexes have been isolated from sporadic cases of human *Campylobacter* infections, ruminants, poultry, dogs, and various meat (lamb, chicken, and beef) samples in the UK, Canada and the Netherlands (http://pubmlst.org/campylobacter/).

Some *C. jejuni* sequence-types isolated in this study were only found in New Zealand water when it is queried in the PubMLST database. The unique geographical features of New Zealand might have been responsible for the fact that ST-3656, ST-3659, ST-3664 and ST-2381 have only been identified in New Zealand waters as reported in the *Campylobacter* PubMLST database (Muellner et al., 2011). Possibly, recombination of *C. jejuni* might have led to a new sequence type arising in New Zealand or, alternatively, these sequence types may not have been present at the time of sampling, or are present only in very low numbers thus unidentified in water samples from other countries. Nevertheless, further investigation is required before making any conclusive remarks.

2.7 Conclusion

Taken together, the findings of our study suggest that cattle and sheep may contribute to water contamination, and may therefore act as a source of infection to humans. The population structure of *C. jejuni* from swimming sites demonstrated the dominance of wild bird strains of *C. jejuni* in these waters. Nonetheless, findings of livestock- associated strains of *C. jejuni* and strains that have been isolated from human infections in the same region demonstrate that there is a risk of contracting campylobacteriosis from swimming in natural freshwaters. This study also highlights the spatial and temporal variations and effect of river flows in isolating *Campylobacter* spp. from recreational river water in the Manawatu-Wanganui region. Therefore, the model will be useful for drinking water suppliers in generating an effective treatment plan by measuring the river flow to determine the risks of getting faecal contamination into the source water. In addition, public health officers can use this model to mitigate the *Campylobacter* exposure by assessing the risks, times and sources of potential water contamination.

2.8 Highlights of the study

- *Campylobacter jejuni* is also determined in the rivers of Manawatu-Wanganui region.
- Wild bird associated *C. jejuni* (ST-2381) was dominant in the river water samples.
- Cattle and sheep associated *C. jejuni* that are potentially zoonotic (e.g. ST-61) were also found.
- Months of sampling, sampling sites and river flows may affect the *Campylobacter* isolation from surface water samples.

2.9 Acknowledgements

I would like to thank all the members of the Molecular Epidemiology and Public Health Laboratory (mEpiLab) Team and Horizons Regional Council for their valuable contributions in generating the data for the Manawatu Sentinel Surveillance Program.

Chapter 3

Epidemiology and molecular characterisation of *Campylobacter*, *Cryptosporidium*, and *Giardia* species in farmed ruminant faeces and pre-treatment drinking water in two rural town water catchments in New Zealand

3.1 Preamble

In the previous chapter (2), we described the presence of *Campylobacter* spp. and potentially zoonotic *C. jejuni* in recreational river water samples. Rivers/streams in New Zealand are also used as source water for drinking purposes. Bacterial and protozoan infections have been reported in people living in rural areas of New Zealand, where drinking water suppliers use the rivers and streams. Therefore, this study was designed to determine the contamination sources of pre-treatment drinking water by collecting faecal and water samples from two rural town water catchments in the Manawatu-Wanganui region of New Zealand, and analysing those samples for the presence of *Cryptosporidium*, *Giardia* and *Campylobacter* spp.

3.2 Abstract

Protozoan parasites, *Cryptosporidium* and *Giardia*, and a bacteria belonging to genus *Campylobacter* are among the most widespread waterborne zoonotic pathogens in many developed countries, including New Zealand. Yet, the public health significance and distributions of these three pathogens in ruminant faeces and pre-treatment drinking water, particularly in rural town areas in New Zealand

were unknown. Therefore, the repeated cross-sectional study was conducted to determine the epidemiology of *Campylobacter, Cryptosporidium* and *Giardia* species in faeces from cattle and sheep, and in river/stream water in two rural town catchments, Dannevirke and Shannon, in the Manawatu-Wanganui region. Altogether, 24 water samples from river/stream water and 499 faecal samples from dairy cattle, beef cattle, and sheep on 8 farms were collected between 1st August and 30th November 2011, from the two catchment areas. *Campylobacter, Cryptosporidium* and *Giardia* spp. were detected in 45.1, 30.3 and 3.6% of 499 faecal samples and in 4, 3 and 18 of 24 water samples, respectively. *Campylobacter* and *Giardia* spp. were identified in a majority of the farms (7/8 and 8/8, respectively), whereas *Cryptosporidium* spp. were identified in calves from one Dannevirke farm only. *Cryptosporidium* and *Giardia* spp. were found more commonly in young animals (≤3 months), whereas *Campylobacter* spp. were most commonly found in juvenile (≥ 3 to <12 months) followed by adult animals. Zoonotic potential *C. jejuni*, *Cryptosporidium parvum, G. duodenalis* were identified in the faecal samples. In water samples, all the four isolates of *Campylobacter* were confirmed as *C. jejuni*, and 7/18 *Giardia* isolates were genotyped as *Giardia duodenalis.* Identification of *C. parvum* gp60 allelic types IIA18G3R1 (7/18) and IIA19G4R1 (7/18) and *G. duodenalis* assemblages AII, BII, BIII, BIV in cattle and sheep and assemblages AII, BI, BII, BIV and E in surface water indicate a possible public health risk through both ruminants and water.

3.3 Introduction

Waterborne diseases are a major cause of human morbidity and mortality worldwide (WHO, 2004; Hrudey and Hrudey, 2007a; 2007b). A variety of enteric and non-enteric microorganisms and parasites has been associated with waterborne diseases. The protozoa organisms *Cryptosporidium* and *Giardia*, and a bacterium *Campylobacter* are among the most widespread waterborne pathogens in many developed countries (ESR, 2011; Fletcher et al., 2012; Tam et al., 2012). These pathogens usually cause sporadic cases of gastroenteritis in people but are also associated with outbreaks of disease. Whilst in most immunocompetent patients infections with these pathogens are self-limiting, some people can develop chronic forms and occasionally long-term sequelae, such as Guillain-Barre' syndrome following acute campylobacteriosis (Baker et al., 2012; Nachamkin et al., 2008). Ruminants are considered to be the main reservoir for all three pathogens.

Cryptosporidium and *Giardia* species were not considered to be of significant public health importance before the Milwaukee, Wisconsin waterborne *Cryptosporidium* outbreak in 1993, in which 400,000 people were infected with *Cryptosporidium* spp. from the same drinking water source (Mackenzie et al., 1994). This and several other waterborne outbreaks in many countries that followed led public health officials to consider *Cryptosporidium* parasites to be important waterborne human zoonotic pathogens. Worldwide, *Cryptosporidium* and *Giardia* parasites have been associated with sporadic and outbreak cases of diarrhoea and nutritional disorders in both humans and animals, including cattle and sheep (Slapeta, 2013). To date, *Cryptosporidium parvum* and *Cryptosporidium hominis* and *Giardia duodenalis* assemblages A and B are all considered to be a major cause of cryptosporidiosis and giardiasis in people all over the world (Bouzid et al., 2013). Other species and assemblages including *Cryptosporidium bovis, Cryptosporidium andersoni, C. meleagridis, C. cuniculus, C. felis, C. canis,* and *G. duodenalis* assemblages C and F have also been reported from human cases, though rarely (Bouzid et al., 2013).

In New Zealand, human cases of *Cryptosporidium* and *Giardia* spp. infections were included in the list of notifiable diseases from mid-1996 (Snel et al., 2009a; 2009b). Since then, the cryptosporidiosis and giardiasis annual notification cases have shown an unstable but increasing pattern (Figure 1.1). During the 10-year period, from 1997 to 2006, the average annual incidence rate of cryptosporidiosis was 22 cases, and of giardiasis was 44.1 cases per 100,000 population (Snel et al., 2009a; 2009b). Similar figures were obtained in 2010, which were higher when compared to those reported in other countries (Table 3.1). In addition, these two enteric protozoa were implicated in 37.3% (19/51) of the total number of waterborne disease outbreaks in 2012, and these outbreaks were commonly linked to untreated or inadequately treated drinking water supplies (ESR, 2012a).

Campylobacter species also infect both humans and animals worldwide (Nachamkin et al., 2008; Stanley and Jones, 2003). Two species, *C. jejuni* and *C. coli,* are the most commonly identified pathogens causing gastroenteritis in humans (Nachamkin et al., 2008). Contact with farm animals, live birds, drinking contaminated water or raw milk, eating raw and improperly cooked meat, are commonly identified exposure risk factors of *Campylobacter* infection in humans (Mullner et al., 2009; Wilson et al., 2008; Wong et al., 2006). Campylobacteriosis is a notifiable disease in New Zealand and the rate of reported cases in this country is high relative to other developed countries (158.6 per 100,000 population in 2012) (ESR, 2012b).

However, campylobacteriosis is not a notifiable disease in some other countries, and comparing notification rates between countries is though possible but may not be meaningful. The annual rate of *Campylobacter* infections reported in New Zealand has been reduced by half since the implementation of processing plant control measures by the poultry industry in 2007 (Sears et al., 2011). In addition, *Campylobacter* spp. were associated with 19.6% of waterborne outbreak cases in New Zealand in 2012, and these *Campylobacter*-associated waterborne cases were also commonly linked to untreated or inadequately treated water supplies (ESR, 2012a). Therefore, these persistently high numbers of campylobacteriosis cases validate the need to conduct further research on sources of infections other than poultry, such as livestock, farm run-off and contaminated water.

Cattle and sheep are considered amplifier species for *Cryptosporidium* and maintenance reservoir species for *Campylobacter* and *Giardia*. Although calves and lambs between 1 and 8 weeks of age shed a higher number of organisms than adult animals, shedding of all three of these pathogens has also been reported in adult cattle and sheep that appeared to be clinically normal (Fayer and Xiao, 2008; Rapp et al., 2012).

Table 3.1: Campylobacteriosis, cryptosporidiosis and giardiasis cases reported in New Zealand and other developed countries in the year 2010. 'NA' indicates data not available.

Countries	**Cases per 100,000 population**			**References**
	Cryptosporidiosis	**Giardiasis**	**Campylobacteriosis**	
New Zealand	21.8	45.4	158.6	ESR, 2011; 2012a
Australia	6.6	NA	112.3	NNDSS, 2012
United Kingdom	7.4	6.5	113.3	ECDC, 2012
Scotland	8.5	4.0	127.0	HPA, 2011; 2012a; 2012b
United States	2.9	7.6	13.6	Yoder et al., 2012; CDC, 2013b

In New Zealand, 36.0 (58/161), 21.1 (33/156) and 4.5% (7/155) of cattle (Grinberg et al., 2005; Moriarty et al., 2008), and 30.4 (66/220) and 3.6% (8/220) of sheep faecal samples (Moriarty et al., 2011a) were found to be positive for *Campylobacter,*

Cryptosporidium and *Giardia* spp., respectively by using a culture method for bacteria and immune-fluorescence assays for protozoa. Although *Giardia* prevalence in sheep have not been reported previously, our study provides evidence that cattle and sheep are likely to play an important role in water contamination with these agents by direct deposition of faeces into the water source, or by indirect means such as surface runoff, either from farmland or from improper management of animal waste (Bezirtzoglou et al., 2011; Smith & Grimason, 2003). In addition, increases in cattle density and land use for cattle and sheep farming are thought to be risk factors for increasing water contamination with enteric pathogens (Snel et al., 2009a; 2009b; Spencer et al., 2012).

Campylobacter, Cryptosporidium and *Giardia* spp. have been isolated from marine water, ponds, lakes, streams, rivers and other waterways (Fayer, 2004; Horman et al., 2004; Feng et al., 2011; Helmi et al., 2011). Lakes, streams, rivers and groundwater are sources of drinking water in New Zealand. Some drinking water sources have been found to be contaminated with faecal bacteria, with coliform counts exceeding the maximum acceptable concentrations (Anonymous, 2011). *Campylobacter, Cryptosporidium* and *Giardia* spp. were detected in 23, 23, and 9% respectively of 80 water samples collected between July 2009 and June 2010 before undergoes for treatment (Prattley et al., 2010). These water samples were collected from 20 large water plants supplying drinking water to New Zealand communities with populations greater than 10,000. These large water treatment plants were automated water treatment facilities that used processes including chlorination, ozonation and UV radiation or a combination of them. In contrast, a manually controlled single treatment process such as chlorination is routinely used in small, often rural, water treatment plants (WINZ, 2011). Consequently, the smaller water treatment plants are unable to cope with extreme conditions, such as heavy rainfall and flooding, and this might lead to *Campylobacter, Cryptosporidium* and *Giardia* spp. infections in consumers who depend on smaller water suppliers (Craun et al., 2010; Roig et al., 2011). New Zealand has an open grazing system where ruminants are kept on pasture and stock often graze around rural water sources. Therefore, an improvement in the understanding of the epidemiology and transmission dynamics of zoonotic pathogens from ruminants to sources of drinking water could assist the development of control measures aimed at reducing waterborne cases of gastroenteritis and potential sequelae in smaller communities.

The Tamaki River is the source of drinking water for the town of Dannevirke, and the Mangaore Stream is the major source of drinking water for the town of Shannon. Both are rural towns in the Manawatu-Wanganui region of New Zealand, and the water catchment areas include cattle and sheep farms, as well as areas of bushland and forest. The drinking water in these areas was treated by chlorinating gravity feed water in 2010/2011. The New Zealand Ministry of Health has assessed the drinking water quality of the Tamaki River and the Mangaore stream and the supply infrastructure, and graded both catchments in the "Ee" category, indicating an unacceptable level of risk to human health (WINZ, 2011). Therefore, analysis for the presence of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. in ruminants' faeces around the catchment areas and in water samples could help to estimate the likelihood of transmission of pathogens from ruminants into the source water, and elucidate the human health risks associated with the consumption of contaminated or ineffectively treated water. Furthermore, understanding the risk factors associated with water contamination by zoonotic pathogens from ruminants and farms would assist with the development of appropriate control strategies. Finally, genotyping of the microbial isolates collected in this study can provide data required for source attribution studies. Therefore, a repeated cross-sectional study was conducted in order to:

a) investigate the occurrence of *Cryptosporidium*, *Giardia* and *Campylobacter* spp. in ruminant faeces and in surface water at drinking water extraction points
b) explore the temporal distribution of these pathogens in livestock faeces and water during lambing and calving season, and
c) identify associations between the presence of *Cryptosporidium*, *Giardia* and *Campylobacter* spp. in faeces and variables such as animal age and species and other farm-related factors.

3.4 Materials and methods

3.4.1 Study design and sampling period

Surface water and ruminant faecal samples collected from two drinking water catchment areas were analysed for the presence of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. To collect the samples, a repeat cross-sectional study was conducted in the Manawatu region of New Zealand for a four-month period in 2011. The timing of the study (2nd August through to 28th November) was chosen to coincide with the calving and lambing seasons in the area, which were identified

through consultation with local veterinarians. The reason for selecting this study period was to allow for the collection of faecal samples not only from adults, but also from calves and lambs that were <1 month of age, and this period is the short and concentrated calving and lambing season characterising dairy cattle and sheep farming in New Zealand. Ruminants in this age group are known to shed these three pathogens at high frequency (Giacoboni et al., 1993; Xiao and Herd, 1994; Grinberg et al., 2005), and thus calving and lambing time may represent a time of greater risk of drinking water contamination (Keeley and Faulkner, 2008).

3.4.2 Drinking water catchment area and farm selection

Two small towns in the Manawatu region (Dannevirke and Shannon) were selected for this study. These towns were selected because they had been given a low drinking water quality grading (grade Ee) in 2010 (WINZ, 2011). In addition, the local public health unit reported receiving persistent complaints regarding poor water quality and repeated reports/notification of diarrheal diseases in the two towns [personal communication from public health officers at the Public Health Unit, Mid-Central District Health Board (MDHB)].

In Dannevirke, the drinking water for the town is abstracted from the Tamaki River at Armstrong Road, is gravity fed, and stored in two reservoirs of 16,644m^3 and 4,546m^3 capacities (Anonymous, 2011b). In 2010, residents of Shannon received their drinking water from the Mangaore Stream after gravity feeding and storage in a 1,150m^3 reservoir (Anonymous, 2011a). All the reservoirs were situated at a lower altitude than the water abstraction points, and water was distributed to residents of both sites from the respective reservoir after chlorination (Appendix C).

The Agribase™ dataset holds information on categories of land use, farm co-ordinates and owner details (AsureQuality Ltd., 2011); whilst a river environment classification database held by the National Institute of Water and Atmospheric Research (NIWA, 2010) provides geospatial information about the rivers in New Zealand. These two databases were used to categorise the river catchment areas upstream from the drinking water intake points and the farms within the catchment areas (Figure 3.1). The data from these databases were linked with the co-ordinates of the river water intake points, and the river catchment areas were defined in ArcView GIS 3.2a software (ESRI, 2002). The Agribase™ dataset showed

the presence of 15 farms in the Dannevirke catchment and 3 farms in the Shannon catchment with an overall mean farm size of 400 (range: 200-800) animals.

There were no deer farms present within the catchment areas; therefore, only cattle and/or sheep farms were considered in this study. Initially, letters were mailed to the farmers in the Shannon and the Dannevirke catchment areas inviting them to participate in the study. Subsequently, a phone call was made to each farmer to discuss the objectives of the study, and to ask for their oral consent to participate. In total, 3/3 and 5/15 farms within the Shannon and the Dannevirke catchment areas, respectively, consented to participate in this study. Among the eight farms, both beef cattle and sheep were present in four farms and sheep only in two farms in both catchment areas, whereas beef cattle were present only in one farm and dairy calves were present in one farm in the Dannevirke catchment area.

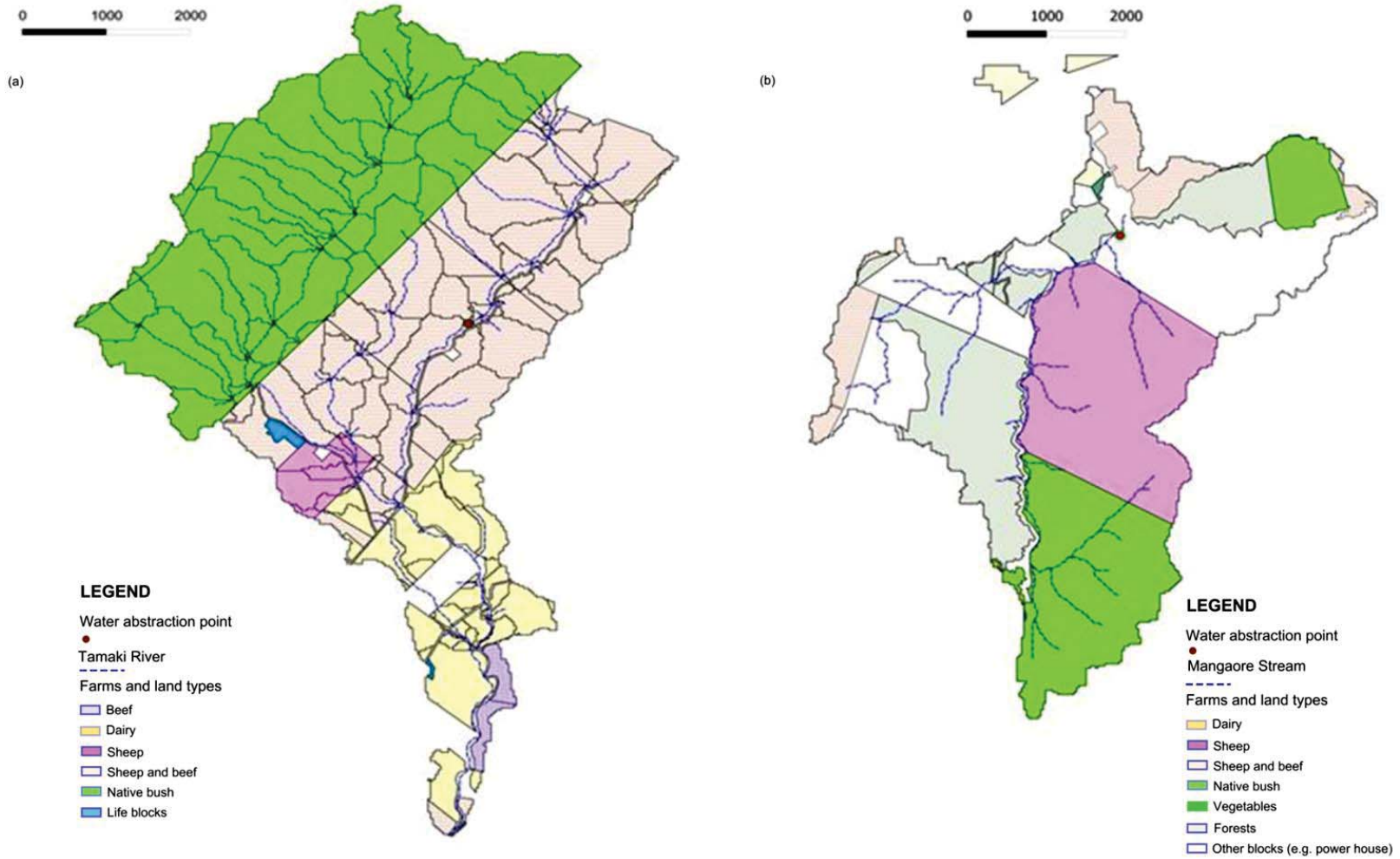


Figure 3.1: The Dannevirke (a) and Shannon (b) catchment areas along with river/streams, water abstraction (intake) points and farms within the catchment areas.

3.4.3 Sample size computations

3.4.3.1 Faecal samples

The minimum number of faecal samples required to detect at least one positive animal for any of the three pathogens per farm from both the catchment areas were estimated using the Agribase™ dataset, published reports on the prevalence of the three pathogens in livestock, and the sensitivity and specificity of the diagnostic tools used.

As published reports suggested that *Cryptosporidium* and *Giardia* spp. are less likely to be detected in ruminant faeces than *Campylobacter* spp. (Dorner, et al., 2004; Grinberg et al., 2005), we used estimated values for the within- and between-farm prevalence of *Cryptosporidium* and *Giardia* spp. to calculate the required sample size. Previously reported animal-level (within-farm) prevalences for *Cryptosporidium* spp. ranged from 20 to 56% of animals within a farm, and for *Giardia* spp. from 4 to 57% (Grinberg, et al., 2005; Ng, et al., 2011; Yang, et al., 2009). At the farm level (between-farm), reported *Cryptosporidium* spp. prevalence ranged from 4.5 to 42%, and *Giardia* spp. prevalence from 11 to 38% (Grinberg, et al., 2005; Ng, et al., 2011; Yang, et al., 2009).

Hence, for our sample size calculations, we assumed a within-farm prevalence of 20% and a between-farm prevalence of 10%. The sensitivity and specificity of the immuno-fluorescence assay (IFA) technique to detect *Cryptosporidium* and *Giardia* spp. are both 90% (USEPA, 2005). The number of faecal samples required (n) to be 95% certain of detecting at least one positive animal for if pathogen was present was estimated using the epiR library package (Stevenson, et al., 2011) and the following formula:

$$n = \left\{1 - (\propto)^{\frac{1}{D}}\right\}\left\{N - \frac{D-1}{2}\right\}$$

Where:

n = the number of faecal samples required

α = 1-confidence level (α is 0.05)

N = mean farm animal size (i.e. 400)

D = estimated number of organisms detected in animals in the group

*(mean population size * minimum expected prevalence)*

Based on this calculation, it was estimated that a minimum of seven farms and 14 faecal samples per farm would be required. Therefore, we included all the eight farms (Dannevirke: 5; Shannon: 3) that are happy to be involved from both catchment areas, and collected 15 faecal samples per farm on a monthly basis between 1st August and 28th November 2011 for a total of four visits per farm.

3.4.3.2 Water samples

For detecting *Campylobacter*, 200 mL water is required; whereas for protozoa detection, 10 to 100 L water is passed through a membrane filter (USEPA, 2005). It is also important to establish an optimal water sampling frequency, because frequent (e.g. weekly) water sampling may lead to over-representation of isolated organisms, due to the strong serial correlation with the previous week's results (Anonymous, 2007). On the other hand, with monthly water collection, transient water contamination could remain undetected. In addition, weekly water sampling and processing is expensive to perform. Therefore, as a compromise, water sampling was performed fortnightly from each catchment area during the study period. In addition, storm-water samples were collected two days after heavy rainfall events in the catchment areas, to determine whether there were differences in detecting the three pathogens after heavy rainfall. Both regular and storm water samples were collected upstream of the abstraction point.

3.4.4 Farmer interviews

A questionnaire developed to collect farm management information so that the association between human or animal diseases on the farms and the three pathogens of interest could be evaluated.

Nine people who farmed or who had worked on farms assessed the performances of questions within an initial questionnaire designed using a cognitive testing approach. Cognitive testing of questionnaires is a method of measuring the comprehensiveness of questions, a respondent's ability to retrieve answers, his or her judgement of questions and answers, and responsiveness to prior answers (Collins, 2003). In light of the results of the cognitive testing, minor alterations were made to the questionnaire prior to delivering the final version to the participating farmers in the two catchments. The final questionnaire is presented in full in Appendix C.

The finalised questionnaire was delivered to each farm owner during the first visit to his or her farm. During each subsequent farm visit, a shorter questionnaire was completed with the farmer to elicit the presence of diseases (if any) and any animal movements' onto the farms during the previous 30-day period. The data generated from the questionnaires were stored in password-protected Excel spreadsheets (MS Office 2007, Microsoft Corporation).

3.4.5 Ethics approval

In this study, faecal samples were collected from the ground and animals were not manipulated in any way; therefore, it was not necessary to seek approval from an animal ethics committee. However, a 'low-risk notification' was submitted to the Massey University Human Ethics Committee, because of the collection of personal information using the questionnaire, including information on human illnesses potentially caused by any of the three organisms analysed (low-risk notification number PN623, submitted on 17 May 2011). Participating farmers signed a written consent form before the delivery of the questionnaire and it was explained to them that they were free to choose whether to answer a question or not. The invitation letter to participate in the research and consent forms is shown in Appendix C.

3.4.6 Faecal sample collection and processing

3.4.6.1 Collection of faecal samples

Freshly deposited (during sampling time or ≤1 hour), wet faeces were collected from the ground in the paddocks that contained the youngest stocks that were being grazed near to the river or its tributaries (paddock adjacent to the river) on the day of sampling. The faeces were placed into sterile, plastic containers and the date and time of sampling, and the species and age of the animals were recorded for each collected sample. The co-ordinates of the sampling points for each sample were also recorded using a GPS navigator (Garmin, Kansas, USA). The faecal samples were transported to the laboratory on ice in a cool box and were stored at 4 ^{o}C. Upon arrival at the laboratory, faecal samples were analysed for *Campylobacter* within two hours of sample collection in order to prevent sample deterioration impeding *Campylobacter* detection, whereas analysis for the presence of *Cryptosporidium* and *Giardia* spp. was performed within 72 hours.

3.4.6.2 Detection of pathogens in faecal samples

3.4.6.2.1 Detection of *Campylobacter* spp.

For *Campylobacter* spp., approximately 1 g of each faecal sample was inoculated into 3 mL Bolton's enrichment broth, and incubated for 48 hours at 42 ^{o}C in a micro aerobic (85% N_2, 10% CO_2 and 5% O_2) incubator (VAIN, Don Whitley Scientific, Yorkshire, UK). Aliquots of incubated broth were sub-cultured onto modified cefoperazone-charcoal-deoxycholate (mCCDA) agar plates (Fort Richards Laboratories, Auckland, New Zealand) agar plates, and incubated in the same environment for 48 hours. Plates were examined for growth of *Campylobacter* colonies of typical grey, sticky and muddy appearance. A single colony was mixed in a drop of sterile distilled water on a slide and was examined by dark-field microscopy (Olympus, BH2). If bacteria were present that displayed motile corkscrew movement and were spiral in shape, then they were recorded as presumptive *Campylobacter*. Two colonies of presumptive *Campylobacter* from each plate were sub-cultured onto 5% horse lysed blood agar (BA) (Fort Richards Laboratories, Auckland, New Zealand), and incubated at 42 ^{o}C for 48 hours in a micro aerobic chamber as above. After 24 hours, one *Campylobacter* colony was recovered from the BA plate for DNA preparation. The remaining presumptive *Campylobacter* colonies in the plate were suspended in 15% glycerol broth (Oxoid, Basingstoke, England) and stored in a -80 ^{o}C freezer for future reference.

3.4.6.2.2 Detection of Cryptosporidium and Giardia spp.

Testing for *Cryptosporidium* oocysts and *Giardia* cysts in faecal samples was performed using an immunofluorescence microscopy assay (IFA). Approximately 50-100 mg of each faecal sample was suspended in eppendorf tubes containing 700 µL of phosphate buffered saline, pH 7.4, (PBS). The eppendorfs were vortexed until the faeces were thoroughly suspended, and then allowed to settle. Then, 50µL volume of the supernatant were dried on fluorescence microscopy slides (Marienfield, Germany) in a 37 ^{o}C incubator for 30 minutes, fixed with 20 µL absolute methanol, and covered with 50 µL of a diluted (1:5) fluorescein isothiocyanate conjugated anti-*Cryptosporidium* and anti-*Giardia* antibody solution (AquaGlo™ G/C kit, Waterborne Inc., LA, USA) in molecular grade water (MGW). A freshly diluted antibody solution was prepared for each testing. The slide was incubated at 37 ^{o}C for 45 minutes in a humid chamber in the dark and the slide was briefly washed twice using 50 µL of PBS; air-dried; mounted using mounting fluid and covered with a coverslip. The slides were examined for the presence of apple-

green, round *Cryptosporidium* oocysts of size 4-6 µm or elliptical *Giardia* cysts of size 8-14 µm using an epifluorescence microscope (Olympus BX 60, Tokyo, JAPAN) with 460-490 nm excitation wavelengths and at 20x and 40 x magnification.

3.4.6.3 DNA extraction of pathogens from faecal isolates

3.4.6.3.1 DNA isolation of Campylobacter spp.

Campylobacter DNA was extracted immediately from freshly grown BA cultures. A 5 µL loop of culture was added to a microtube containing 2% Chelex (Biorad) and the mixture was boiled for ten minutes. After boiling, the mixture was centrifuged at 10,000 x *g* to remove both cell debris and the Chelex. The supernatant containing the DNA was transferred to a new microcentrifuge tube, and the tube was stored at -20 ^{o}C for up to 2 years and used it prior to molecular typing.

3.4.6.3.2 DNA isolation of Cryptosporidium spp.

Cryptosporidium DNA extractions from the faecal samples were performed using a faecal DNA kit (Bioline, Sydney, Australia) following the manufacturer's instructions. In brief, the faecal samples in which oocysts had been detected microscopically were added directly to a bashing beads lysis tube (Bioline, Sydney, Australia) and centrifuged rapidly at 10,000 x *g* for 1 minute without the use of organic denaturants or proteinases. The lysate was then filtered to recover DNA, and the DNA was bound, washed, isolated and purified using the spin columns provided.

3.4.6.3.3 DNA isolation of Giardia spp.

Several attempts were made to extract DNA using the same kit used for *Cryptosporidium* (Bioline, Sydney, Australia), which did not work for our samples. Therefore, *Giardia* The NucleoSpin® Soil DNA isolation kit (MACHEREY-NAGEL GmbH & Co. KG, Germany) was used to extract *Giardia* DNA from the IFA-positive faecal samples following the manufacturer's instructions. Approximately 300 mg of faeces and lysis buffer (700 µL SL1 + 150 µL enhancer SX) were added to a NucleoSpin® Bead Tube, and the tube was vortexed horizontally for 5 minutes. The tube was centrifuged (11,000 x *g* for 2 minutes), and 150 µL SL3 buffer was then added to the tube. Prior to being centrifuged (11,000 x *g* for 1 minute) to precipitate the contaminants, the tube was briefly mixed in a vortex and chilled for 5 minutes at 4 ^{o}C. The supernatant was loaded onto a NucleoSpin® inhibitor removal column and centrifuged at 11,000 x *g* for 1 minute. The mixture of filtrate and 250 µL SB buffer was added into a NucleoSpin® Soil column and was centrifuged (11,000 x *g*

for 1 minute) to bind the DNA. The DNA bound column was then washed with the washing buffers provided, centrifuged (11,000 x *g* for 30 sec), and finally, DNA was eluted, using elution buffer. The eluted DNA solution was stored at -20 ^{o}C in a freezer for use in a PCR assay.

3.4.7 Water sampling and processing

3.4.7.1 Collection of water sample

Separate water samples were collected for the recovery of *Campylobacter* and *Cryptosporidium/Giardia* spp. In the Dannevirke catchment, the samples were taken in the vicinity of the abstraction point, at a distance of 1 m from the riverbank. In Shannon, it was not possible to reach the abstraction point due to blockage of the road. Therefore, water samples from the Shannon catchment area were collected from a tap at the water treatment plant. The tap was situated in the pipeline that passes raw river water to a reservoir tank. The surface of the tap was disinfected with 70% alcohol, and the first 2-3 litres of water were discarded prior to water collection.

For the isolation of *Campylobacter* spp., 200 mL water were collected from the river/stream in a 250 mL polypropylene sterile bottle and kept in a cool-box. For protozoa detection, the initial plan was to collect a standard volume of 100 L of water. However, increased water-turbidity during the winter period often clogged the filters, which prevented the collection of this desired volume of water. Therefore, on any occasion, water was collected until the desired volume was reached or for 90 minutes, regardless of the volume. Overall, between 50 L and 300 L source-water were filtered through a Filta-Max® (IDEXX, Westbrook, Maine, USA) foam filter system. Before water sampling, the filtration equipment, including the filter housing, pipes and peristaltic pumps, was rinsed with reverse osmosis (RO) water and dried in a 37 ^{o}C incubator. The filtration equipment was connected upstream to the filtration housing; whilst downstream, the housing was joined to a flow metre through the pipe (Appendix C). In the Tamaki River at Dannevirke, the filtration equipment was securely placed on the bank of the river. To avoid clogging of the filter by sucking up soil nearby the bank, the pump was placed into the water at a distance of at least of one metre from the riverbank. In Shannon, the filtration equipment was connected to the disinfected tap at the pre-treatment reservoir. The water bottles and used water filters were kept in a cool box with ice and were

transported to the laboratory at the Hopkirk Research Institute for further processing.

In addition, to determine the how well the USEPA method 1623 performs with the particular type of sample matrix expected to be obtained in this study, 100 L water samples were collected from both sampling sites once every three months during the year 2012 and these samples were subjected to an oocyst spiking experiment, as described in Appendix C.

3.4.7.2 Detection of pathogens in water samples

3.4.7.2.1 Detection of Campylobacter spp.

Water samples were processed for *Campylobacter* spp. isolation immediately upon arrival at the laboratory. Each 200 mL of water sample collected was filtered through a 0.45 µm filter (Sartorious AG, Germany) using vacuum filtration with vacuum pressure kept between 50 and 100 mmHg. The filter was removed aseptically using forceps pre-soaked in 70% alcohol, and immersed into 20 mL Bolton broth. The broth was incubated and processed using the methods described for the faecal samples.

3.4.7.2.2 Detection of Cryptosporidium and Giardia spp.

Detection of *Cryptosporidium* and *Giardia* in water samples was performed according to the United States Environmental Protection Agency 1623 procedure (USEPA, 2005) except for the elution procedure. As there was not automatic elution system in the laboratory, the elution procedure of the USEPA method was modified. Briefly, the filter was placed in a stomacher bag (Stomacher® 3500 series standard bag, Seward Limited, West Sussex, UK), and the filter module was removed allowing the foam discs to expand. The discs were rinsed with a 500 mL eluting solution (PBS powder, pH 7.4, Sigma-Aldrich, USA) in the stomacher bag, and the bag was placed in the Stomacher to homogenise for 10 minutes on the normal setting. The eluate in the bag was decanted into a 2 L beaker and the foam disks in the bag were wrung by hand to remove any further eluate, which was poured into the beaker. The eluate was decanted into a 500 mL conical centrifuge tube, and centrifuged at 3000 x *g* for 15 minutes at 10 ^{o}C in a bench-top centrifuge (Sorvall RT7, GMI Inc., Minnesota, USA). The pellet volume was recorded and a Venturi vacuum unit was used to aspirate off the supernatant to 50 mL. The tube was vortexed to resuspend the pellet, and the solution was decanted into a 50 mL centrifuge tube. This was centrifuged as before, and 40 mL supernatant solution was aspirated off. The

concentrated sample was processed according to the manufacturer's instructions using an anti-*Cryptosporidium* and anti-*Giardia* immunomagnetic separation kit (Dynabeads® GC-Combo IMS kit: Invitrogen, California, USA). Finally, the 50 µL solution was transferred to a labelled fluorescence microscopy slide (Marienfield, Germany), and was examined using a similar procedure to that previously described for faecal samples. However, an additional step of adding 50 µL of 4-, 6-diamidino-2-phenylindole (DAPI) stain solution (Life Technologies™, CA, USA) to a well-slide for one minute was applied prior to the mounting of the slide to counterstain the nuclei present in the (oo)cysts. All positive slides were stored in a dark box at 4 ^{o}C to allow for genotyping at a later date.

3.4.7.3 DNA extraction of pathogens from water isolates

3.4.7.3.1 DNA isolation of Campylobacter spp.

The *Campylobacter* DNA from water samples was extracted using a procedure similar to that used for the faecal samples described in section 3.4.6.2.1.

3.4.7.3.2 DNA isolation of Cryptosporidium and Giardia spp.

The Chelex Freeze-Thaw Lysis method described by Di Giovanni *et al.* (2010) was used to extract *Cryptosporidium* and *Giardia* DNA from all the eluted slides originating from IFA-positive water samples, and control slides. The control slides were prepared by spiking a known number of oocysts in molecular grade water in the laboratory.

Firstly, (oo)cysts present on the microscopic slides were removed using a procedure described by Di Giovanni *et al*., 2010. Briefly, the slide was placed on a clean paper towel and the nail polish on the slide was removed using a non-acetone-based nail polish remover. Then, the cover slip was removed using a sterile blade by gently lifting it up from a corner, and then it was inverted onto a clean paper towel. After removing the cover slip, the slide was washed with 50µL of molecular grade water (MGW), and the water was aspirated and collected in a 1.5mL sterile tube. An additional 15µL MGW was added to the slide and the surface of the slide was scraped using closed cell foam (2 mm thick; 3 x 3 mm square). The slide was scraped twice with a 90^{o} rotation between scrapings. The same foam cell was used to scrape the cover slip after moistening it with 15µL MGW. The scrapped MGW was aspirated and transferred to the tube. The foam scraper was also added to the tube. Both the scraped slide and cover slip were dried and immediately examined under the fluorescence microscope to verify the complete removal of

(oo)cysts. Finally, the tube was centrifuged briefly at 10000 x *g* up to 15 seconds to bring the water down to the bottom of the tube before DNA extraction.

Twenty µL of 1:1 Chelex/MGW slurry were added to each sample tube containing the foam used to scrape the slide, and the tube was briefly vortexed at 2500 rpm for up to 15 seconds (Labnet, NJ, USA) and then centrifuged. Then, the contents of tubes were lysed using 8 1-minute cycles of freezing in liquid nitrogen and thawing on a heating block (95 ^{o}C). After each thaw, the samples were shaken to allow the liquid to fall to the bottom of the tube. Thereafter, the lysed samples were centrifuged briefly at 10000 x *g* up to 15 seconds and were transferred to a spin column. The spin columns were then centrifuged at 12470 x *g* for 30 seconds with the filter hinge in the 12 o'clock position. The sample tubes were rinsed with 5-10 µL MGW and vortexed briefly. Then, the rinsate was added to the spin column and was centrifuged at high speed for 30 seconds with the filter hinge in the 6 o'clock position to elute the DNA. The eluted DNA was stored at -20 ^{o}C in a freezer for PCR analyses.

3.4.8 Polymerase chain reaction (PCR) and DNA sequence analysis

The genotypes and subtypes of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. present in both faecal and water samples were analysed using PCR and DNA sequence analysis.

3.4.8.1 *Campylobacter* PCR

Firstly, PCR of the presumptive *Campylobacter* isolates was performed to confirm the genus *Campylobacter*. Those isolates that were *Campylobacter* genus-positive were further identified to detect the genes associated with either *C. jejuni* or *C. coli*. All the samples that appeared negative were re-analysed two more times using the same PCR protocols.

For confirmation of *Campylobacter* genus, 16*S rRNA* gene within the *Campylobacter* DNA was targeted, whereas *Campylobacter* species were typed by amplifying a map*A* gene that is found only in *C. jejuni* and the ceuE gene of *C. coli* (Table 3.2). Each PCR reaction, which was made to a final volume of 20 µL using sterile water, was amplified in a thermocycler (SensoQuest Labcycler, Goettingen, Germany) (Table 3.2). The primers and amplification conditions are presented in Table 3.2. The amplicons were visualised after electrophoresis using a 1% agarose gel in Tris-

borate-EDTA (TBE) buffer by staining the gel with ethidium bromide and exposing it to ultraviolet illumination. The isolates were confirmed as *Campylobacter* if a ~816 bp product was present. For the species typing PCR, a ~603 bp or ~462 bp product indicated the presence of *C jejuni* or *C. coli*, respectively.

3.4.8.2 *Cryptosporidium* and *Giardia* PCR

The highly polymorphic regions of *Cryptosporidium*, 18S small subunit rRNA (18S SSU RNA) and 70 kDa heat shock protein (hsp) genes, were amplified to identify the *Cryptosporidium* species present in the samples, and the polymorphic 60 kDa glycoprotein (gp60) locus was targeted to determine the subtypes of *Cryptosporidium parvum*. The glutamate dehydrogenase (GDH) locus within the *Giardia* cysts was amplified for detecting the *Giardia* species and subtypes (assemblages and subassemblages)present in the samples. Details of the reaction mixtures and the conditions used for amplification of *Cryptosporidium* are shown in Table 3.3. The amplifications were carried out in a SensoQuest Labcycler Thermocycler. The PCR products were visualised using ethidium bromide and ultraviolet light.

Positive amplicons for *Cryptosporidium* or *Giardia* spp. were purified using an in-house Polyethylene glycol (PEG)-ethanol precipitation method. Briefly, 20 µL of 20% PEG was added to each PCR tube, mixed and incubated at 37°C for 15 minutes. Thereafter, tubes were spun at 1,000in the bench centrifuge for 30 minutes. After centrifugation, the supernatant was removed with a pipette without disturbing the deposit. Then, 150 µL of 80% ethanol was added to each tube to precipitate DNA, and the mixture was centrifuged at 12470 x *g* for 10 minutes. The supernatant in the centrifuged tube was carefully aspirated and discarded, and the pellet was air-dried at room temperature overnight.

The pellet was resuspended in 20 µL of molecular grade water, and the DNA present was quantified using NanoDrop spectrophotometer (Thermo Fisher Scientific) and Qubit® fluorometer (Life Technologies). Then, the DNA template was submitted to the Massey Genome Service for bidirectional sequencing of an internal segment of the amplicon using forward and reverse internal primers stated in Table 3.3. Forward and reverse sequences obtained were aligned and edited manually using Geneious software version 5.6.5 (Biomatters Ltd., http://www.geneious.com). The two ends of sequence that could not be verified were trimmed and the resulting edited sequences aligned with sequences deposited

in the Genbank using the alignment algorithm BLAST (http://blast.ncbi.nlm.nih.gov).

Table 3.2: Primers and conditions used for PCR amplification of *Campylobacter* spp. loci in this study.

Typing	***Campylobacter* genus**	***Campylobacter* species**	
Locus	16SrRNA	MapA (*C. jejuni*)	ceuE (*C. coli*)
Primers	C412F: GGATGACACTTTTCGGAGC C1288R: CATTGTAGCACGTGTGTC	MapA-F: CTTGGCTTGAAATTTGCTTG MapA-R: GCTTGGTGCGGATTGTAAA	COL3: AATTGAAAATTGCTCCAACTATG MDCOL2: TGATTTTATTATTTGTAGCAGCG
PCR mixture in each 20 µL reaction	1 x PCR Buffer 0.2 mM dNTP 2.5 mM MgCl2 1 U Platinum Taq 4 pmol of each primer 2 µL DNA template	Same to 16SrRNA mixture except 1.5 mM MgCl2	
PCR conditions	I: 2 minutes at 95 °C D: 30 seconds at 94 °C A: 30 seconds at 56 °C E: 30 seconds at 72 °C Cycle numbers: 40	Same to 16SrRNA mixture except D: 15 seconds at 94 °C A: 20 seconds at 60 °C	
Product length	~816 bp	~603 bp	~462 bp
References	Linton *et al.* 1996	Stucki *et al.* 1995	Gonzalez *et al.* 1997

F= Forward Primer; R= Reverse Primer; Platinum Taq (Invitrogen, Auckland, New Zealand);

dNTP = deoxynucleotide triphosphate (Fermentas, Auckland, New Zealand);

I = Initial heating; D = Denaturation; A = Annealing; E = Extension

Table 3.3: Primers and conditions used for PCR amplification of *Cryptosporidium* and *Giardia* loci in this study

Organisms	***Cryptosporidium* Locus**			***Giardia* Locus**
Locus	**HSP70** **Singe step PCR**	**18S SSU rDNA** **Nested PCR**	**gp60 subgenotype** **Nested PCR**	**GDH** **Nested PCR**
Primer sequences (5'-3')	F':CACCATCCAAGAACCAAAGG R':GCCTAAAGGTAGAGTGTGCTTTC	External		
		F':GTTAAACTGCGAATGGCTCA R':CCATTTCCTTCGAAACAGGA	F':ATAGTCTCCGCTGTATTC R':GGAAGGAACGATGTATCT	F': TCAACGTYAAYCGYGGYTTCCGT R': GTTRTCCTTGCACATCTCC
		Internal		
		F':CTCGACTTTATGGAAGGGTTG R':CCTCCAATCTCTAGTTGGCATA	F':TCCGCTGTATTCTCAGCC R':GCAGAGGAACCAGCATC	F': CAGTACAACTCYGCTCTCGG R': GTTRTCCTTGCACATCTCC
PCR mixture in each 20µL reaction	1 x PCR Buffer 2 mM dNTP 4 mg BSA 1.5 mM MgCl2 2.5 U Platinum Taq 4 pmol of each F'/R 2 µL DNA	External		
		Same to HSP70 except 2.5 mM $MgCl_2$	Same to 18S SSU rDNA	Same to HSP70 except 1 µL of DMSO, 1 U Platinum Taq, 4 µg BSA and 4 µL DNA
		Internal		
		Same to external except 1.5 mM of $MgCl_2$ and DNA replaced with 1µl of external PCR product	Same to 18S SSU rDNA	Same to external except DNA replaced with 2 µl of external PCR product
PCR conditions	I: 2 minutes at 96 °C D: 20 sec at 94 °C A: 20 sec at 57 °C E: 30 sec at 72 °C Cycle numbers: 40	External		
		Same to HSP70	Same to 18S SSU rDNA	I: 5 minutes at 95 °C D: 1 minute at 94 °C x 35 cycles A: 1.30 minutes at 60 °C x 35 cycles E: 2 minutes sec at 72 °C x 2 cycles FE: 5 minutes at 72 °C
		Internal		
		Same to external except annealing temperature of 60°C	Same to 18S SSU rDNA	Same to external
Product length	~400 bp	~850 bp	~850 bp	~420 bp
References	Grinberg *et al.* (2008)	Learmonth *et al.* (2004)	Peng *et al.* (2001) & Alves *et al.* (2003)	Read *et al.* (2005)

F'= Forward Primer; R'= Reverse Primer; dNTP = deoxynucleotide triphosphate (Fermentas, Auckland, New Zealand); Platinum Taq (Invitrogen, Auckland, New Zealand); BSA = non-acetylated bovine serum albumin (Invitrogen, Auckland, New Zealand);
I = Initial heating; D = Denaturation; A = Annealing; E = Extension; FE = Final Extension

3.4.9 Statistical analysis

The laboratory results from the faecal and water samples were entered into separate spreadsheets and were linked to R.2.14.1 (R Development Core Team, 2012) for statistical analysis. Summary measures and 2 x 2 contingency tables were created to explore the presence of any missing observations and outliers. The primary aim was to detect the pathogens in faeces on the paddock and in sources of drinking water over the period of calving or lambing in 2011. Along with this outcome, farm management information obtained from the collected questionnaire was explored using descriptive statistics.

3.4.9.1 Analysis of faecal samples data

For each pathogen, the apparent sample-level prevalence was computed by dividing the number of samples having at least one positive pathogen with the total number of samples analysed. Bar plots were produced to show the proportion of samples positive for *Campylobacter*, *Cryptosporidium* and *Giardia* spp. among farm-, age- and species-specific faecal samples. Line plots were also produced to investigate temporal variation in the proportions of positive samples obtained. Confidence intervals for proportions were calculated using the following formula described by Fleiss (1981).

$$Confidence\ interval\ (CI) = p \pm z * \sqrt{[\{p\ (1- p)\}/n]}$$

Where:

$CI = set\ at\ 95\%;\ z = 1.96$

$p = proportion\ of\ interested\ pathogen$

$n = sample\ size.$

3.4.9.2 Analysis of water samples data

The proportion of water samples that tested positive for each pathogen was calculated for each site. The line plot was produced to investigate the temporal variation between the proportions of positive samples obtained in water compared to faeces.

Sensitivity, also known as recovery rate or efficiency, of the USEPA method 1623 for New Zealand river/stream water was estimated using the matrix spiking method. For budget reasons, the matrix spiking of surface water samples from each site was conducted once every three months for the 2012 year. Using this data, Bayesian modelling was performed in the WinBUGS version 1.4.3 software to estimate the

sensitivity of the USEPA 1623 method used during the actual sampling period of this study (Lunn et al., 2000). Bayesian modelling is a statistical procedure that endeavours to estimate parameters of an underlying posterior distribution based on the observed prior distribution. The sensitivity was estimated after fitting the model:

$$Y_{[i]} \sim binomial\ (sensitivity_{[i]}, N)$$

$$logit\ (sensitivity[i]) = \beta_0 + \beta_1 * X_{1[i]} + \beta_2 * X_{2[i]}$$

$$\beta_0 \text{ or } \beta_1 \text{ or } \beta_2 \sim dnorm\ (0.0, 0.001)$$

Where:

$Y_{[i]}$ = number of ColurSeed (oo)cysts recovered during matrix spiking

N = total number of ColurSeed (oo)cysts spiked

β = regression coefficient (β_0 = intercept; β_1 / β_2 = variable coefficients)

$X_{1[i]}$ = explanatory variable sampling site (either Shannon or Dannevirke)

$X_{2[i]}$ = explanatory variable pathogen type (either Giardia or Cryptosporidium)

$sensitivity_{[i]}$ = sensitivity of method USEPA for Giardia or Cryptosporidium

Here, we assumed a uniform distribution of pathogens in a sample, and 100% specificity of the method (Personal communication Anthony Pita, IVABS). The WinBUGS software also provided 95% credible intervals around the estimated values.

Similarly, we assumed that the number of (oo)cysts that could be present in the river/stream during the sampling time was between 1 and 10,000 (oo)cysts per sample, and that they are uniformly distributed in each site (Prior: True $counts_{[i]}$). Thereafter, probable number of (oo)cysts present during the sampling occasions was also estimated by fitting a Bayesian model:

$$Y_{[ij]} \sim binomial\ (sens_{[cat[i]]}, true\_counts_{[cat[i]]})$$

$$True\ counts_{[i]} = dunif\ (1, 10000)$$

$$sens_{[cat[i]]} \sim dnorm\ (means\ sens_{[cat[i]]}, sd\_sens_{[cat[i]]})$$

Where:

[i] = Cryptosporidium oocyst or Giardia cyst

$cat_{[i]}$ = Categories of [i] as Cryptosporidium oocyst in Shannon, or

Cryptosporidium oocyst in Dannevirke, or

Giardia cyst in Shannon, or

Giardia cyst in Dannevirke

$Y_{[i]}$ *= probable number of oocyst/cyst present in each site during the sampling occasions*

$sens_{[cat[i]]}$ *= estimated sensitivity for each category*

$true\_counts_{[cat[i]]}$ *= true number of Cryptosporidium oocyst or Giardia cyst present in each site*

$True\ counts_{[i]}$ *= prior assumption of probable number of oocyst/cyst present*

means $sens_{[cat[i]]}$ *= mean of sensitivity for each category*

sd_sens$[cat_{[i]}]$ *= standard deviation of sensitivity for each category*

The concentrations of (oo)cysts that were present in the actual water samples and the limit of detection for at least one *Cryptosporidium* oocyst or one *Giardia* cyst recoverable in the collected samples were then determined by using the equations below.

(oo)cysts concentration = No. of (oo)cysts detected/ (Sample volume x Sensitivity)

Limit of detection = one oocyst or cyst / (Sample volume and sensitivity)

3.4.10 Genotyping and subgenotyping analysis of pathogens

Results of *Cryptosporidium* and *Giardia* genotyping and subtyping in this study were then compared with the human and bovine results available in the "protozoa research unit (PRU)" database within mEpiLab in the Massey University of New Zealand.

3.5 Results

In total, 499 faecal samples and 24 water samples were collected between 2nd August and 28th November 2011. Three hundred and twelve faecal samples were collected from the five Dannevirke farms, and 187 faecal samples from the three Shannon farms. The faecal samples originated from three different age groups of cattle or sheep: ≤3 months (n = 59), >3 to ≤12 months (n = 205), and >12 months old (n = 235) age groups. Animals in the age group of ≤ 3 months from 3/8 farms (SF2, DF3 and DF5), >3 to ≤12 months from one Shannon farm (SF1), and >12 months from one Dannevirke farm (DF5) could not be sampled due to their absence

during the sampling periods. Of the collected faecal samples, 242 were from cattle (121 from dairy cattle and 121 from beef cattle) and 257 were from sheep.

A total of 13 water samples from Tamaki River, Dannevirke and 11 water samples from Mangaore Stream, Shannon were collected. Of the 24 water samples collected, five (Dannevirke: 3 and Shannon: 2) were storm water samples collected 1-2 days after heavy rain.

3.5.1 Questionnaire survey

For the study year of 2011, the calving period on the Dannevirke farms was between 26th July and the end of November, and the lambing period between 1st of August and the end of October. In Shannon, the calving period was between the middle of September and the end of November, and the lambing period between the 2nd of August and the 1st of November. The median farm size in the Dannevirke catchment area was 230 ha (range: 59-485.6), and it was 884 ha (range: 270-909) in the Shannon catchment area. At the time of the first sampling in the Dannevirke farms, the median cattle and sheep populations were 180 (Range: 36-460) and 2300 (range: 180-3500) respectively, whilst the Shannon farms had a median of 253 cattle (range: 2-440) and 663 sheep (Range: 60-4180). At the time of first sampling in the month August, the 147 calves present were only on Dannevirke farms, and the 345 lambs present were only on Shannon farms. During the sampling period, scours or diarrhoea were reported in 45 calves on one Dannevirke farm, and 6 calves and 10 lambs on one Shannon farm. Detailed information on the management of the Dannevirke and Shannon farms are shown in Table 3.4.

In summary, housing was found on only one farm, used for newborn calves until weaning, that is, up to 28 days age. Animals drank water directly from the source; therefore, no livestock water was treated. Many farms were unfenced, and stocks were grazing near the waterways for more than 20 days per month. All farms were growing ryegrass/clovers swards and 3/8 farms had crops. Half of the farms had drainage systems that discharged waste into rivers/streams and had experienced 3-1 times drainage failure in 2010. In addition, only one farm had a sewage tank, and the effluent was used to irrigate pasture/crops of that farm.

Table 3.4: Information on farm management variables from Dannevirke and Shannon farms that were obtained through questionnaire interview of farmers.

Variables	Farm management	
	Dannevirke farms (n=5)	Shannon farms (n=3)
Housing for livestock (Present)	1/5	-
Animal water source:		
river/stream	4/5	3/3
town supply	-	-
ground water	-	-
other (dam)	1/5	-
Animal water treatment (Yes)	-	-
Farm fencing (Present)	3/5	-
Fence types:		
Electric wire	2/3	-
Hedges and wire	-	-
Hedges	1/3	-
Others	-	-
Stock grazing near waterways		
1-5 days	-	1/3
6-10 days	-	-
11-20 days	-	-
> 20 days	5/5	2/3
Pasture or crops planted		
ryegrass+clovers+crops	2/5	-
Ryegrass + clovers	3/5	1/3
Ryegrass + pine tree	-	1/3
Crops	-	1/3
Drainage system (Present)	3/5	1/3
Drainage types:		
Tiled	-	1/3
Untiled	3/5	-
Drainage outlets into:		
river/streams	3/5	1/3
Ditch	-	-
other (dams)	-	-
Drainage failure:		
1-3 times	3/3	1/1
4-6 times	-	-
7-10 times	-	-
> 10 times	-	-
Wastewater treatment facilities (Present)	3/5	-

3.5.2 Pathogens in faecal samples

Four hundred and two of the 499 (80.6%) faecal samples were found positive for at least any one of the three pathogens, with 38.1% of 402 had a co-occurrence of two or more pathogens (Table 3.5).

3.5.2.1 *Campylobacter* spp. prevalence

Overall, 225 of 499 (45.1%; 95% CI: 40.6-49.5%) faecal samples from 7/8 farms tested positive for presumptive *Campylobacter* spp (Table 3.6). The highest proportion of samples that yielded presumptive *Campylobacter spp.* bacteria was found in Dannevirke Farm 5 (24% of 225), followed by Shannon Farm 2 (17% of 225) and Dannevirke Farm 2 (14% of 225). The prevalence of presumptive *Campylobacter* spp.-positive samples varied widely between the seven positive farms, ranging from 26.5 to 90% of total samples (n=499) collected (Figure 3.2 (a)). Between locations, there are differences (Chi-square test, P<0.001) in findings of presumptive *Campylobacter* spp. (Dannevirke: 168/312; Shannon: 57/187). The highest proportion of presumptive *Campylobacter* spp. were detected in faecal samples obtained from the >3 to 12 month-old animals (62.1% of 205, 95% CI: 55.7-68.4%) followed by >12 month-old animals (7.1% of 235, 95% CI: 3.8-38.4%) and ≤3 month-old animals (30.8% of 59, 95% CI: 24.8-36.8%) (Figure 3.2 (b)). Similarly, among the species, the highest proportion of presumptive *Campylobacter* spp. was found in dairy faecal samples (68.6% of 121, 95% CI: 60.3-76.9%), followed by beef (56.2% of 121, 95% CI: 47.4-65%) and sheep (28.8% of 257, 95% CI: 22.9-33.9%) faecal samples (Table 3.6). Between sampling months, there were also differences in the *Campylobacter* spp. detection (Figure 3.3), the highest number being detected in November (59.7% of 124) followed by the month of September (42.3% of 149).

3.5.2.2 *Cryptosporidium* and *Giardia* spp. prevalences

In total, *Cryptosporidium* oocysts were detected on 3/8 farms and in 18/499 (3.6%; 95% CI: 2-5.2%) faecal samples. Fifteen of 18 *Cryptosporidium* spp. positive samples obtained were from Dannevirke Farm 4, and the remaining samples were from Shannon Farm 2 (2/3) and Farm 3 (1/3) (Figure 3.2 (a)). The 15/18 *Cryptosporidium*-positive faecal samples were collected from the ≤3 month-old dairy calves, and the remaining *Cryptosporidium*-positive samples originated from >3 to ≤12 months-old beef and sheep faecal samples (Figure 3.2 (b)). Like *Campylobacter* spp., the highest proportions of *Cryptosporidium* spp. were detected

in faecal samples from dairy cattle (12.4% of 121; 95% CI: 6.5-18.3%; Table 3.6), and were detected only in the faecal samples collected in the month of August (Figure 3.3).

All 8 farms had at least one *Giardia* positive, with 159/499 (32%; 95% CI: 27.8-36.0%) faecal samples being positive. Among the *Giardia*-positive samples, 21.4% of 159 were from Dannevirke Farm 4 followed by Shannon Farm 3 (15.1% of 159) and Shannon Farm 1 (13.8% of 159) (Figure 3.2 (a)). *Giardia* cysts were found in 30.1% of 312 faecal samples from the Dannevirke catchment area; and 34.8% of 187 faecal samples from the Shannon catchment areas (Chi-square test, P=0.33). *Giardia* cysts were detected at a greater frequency in faecal samples from young calves (≤ 3-months: 59.3% of 59, 95% CI: 46.5-72.1%) than from juvenile (>3 to ≤12: 29.8% of 235, 95% CI: 23.4-36.2%) and adult (>12-months: 26.8% of 205, 95% CI: 21-32.6%) animals' faecal samples (Figure 3.3).

The highest percentage of *Giardia*-positive samples was detected in dairy cattle faecal samples (36.4% of 121; 95% CI: 27.8-45.0%) compared to sheep (32.7% of 257; 95% CI: 27-38.4%) and beef (25.6% of 121; 95% CI: 17.8-33.4%) faecal samples (Table 3.6). In addition, there were no differences in findings of IFA positive *Giardia* samples among months (Figure 3.3), with the highest positive faecal samples being from the month of August (38.9% of 108) followed by September (33.6% of 149).

Table 3.5: Percentage of presumptive *Campylobacter*, *Cryptosporidium* and *Giardia* species detected alone or co-occurred in ruminant faecal samples collected from the two catchment areas.

Pathogens	**Percentage of faecal samples positive for pathogens (n=402)**
Campylobacter only	40.8
Cryptosporidium only	0
Giardia only	21.1
Campylobacter + Cryptosporidium	0.3
Campylobacter + Giardia	14.4
Cryptosporidium + Giardia	3.7
Campylobacter + Cryptosporidium + Giardia	0.5

Table 3.6: Apparent prevalence of presumptive *Campylobacter*, *Cryptosporidium* and *Giardia* in three different species' faecal samples collected from the ground on farms in the two catchment areas: Dannevirke and Shannon of the Manawatu region in New Zealand (N: number of samples positive and CI: 95% confidence interval). A "-" denotes the pathogens not detected in the samples.

Pathogens	**Percentage of samples positive for pathogens (N; CI)**				
	Dannevirke (n=312)			**Shannon (n=187)**	
	Dairy cattle (n=121)	**Beef cattle (n=53)**	**Sheep (n=138)**	**Beef cattle (n=68)**	**Sheep (n=119)**
***Campylobacter* spp.**	68.6 (83; 60.3-76.9)	52.8 (28; 39.4-66.3)	40.6 (57; 32.9-49.7)	58.8 (40; 47.1-67.6)	14.3 (17; 8-20.6)
***Cryptosporidium* spp.**	12.4 (15; 6.5-18.3)	-	-	1.5 (1; <0-3.6)	1.7 (2; <0-4)
***Giardia* spp.**	36.4 (44; 27.8-44.9)	20.8 (11; 9.8-31.7)	28.3 (39; 20.7-35.8)	29.4 (20;18.6-37.5)	37.8 (45; 29.1-46.5)

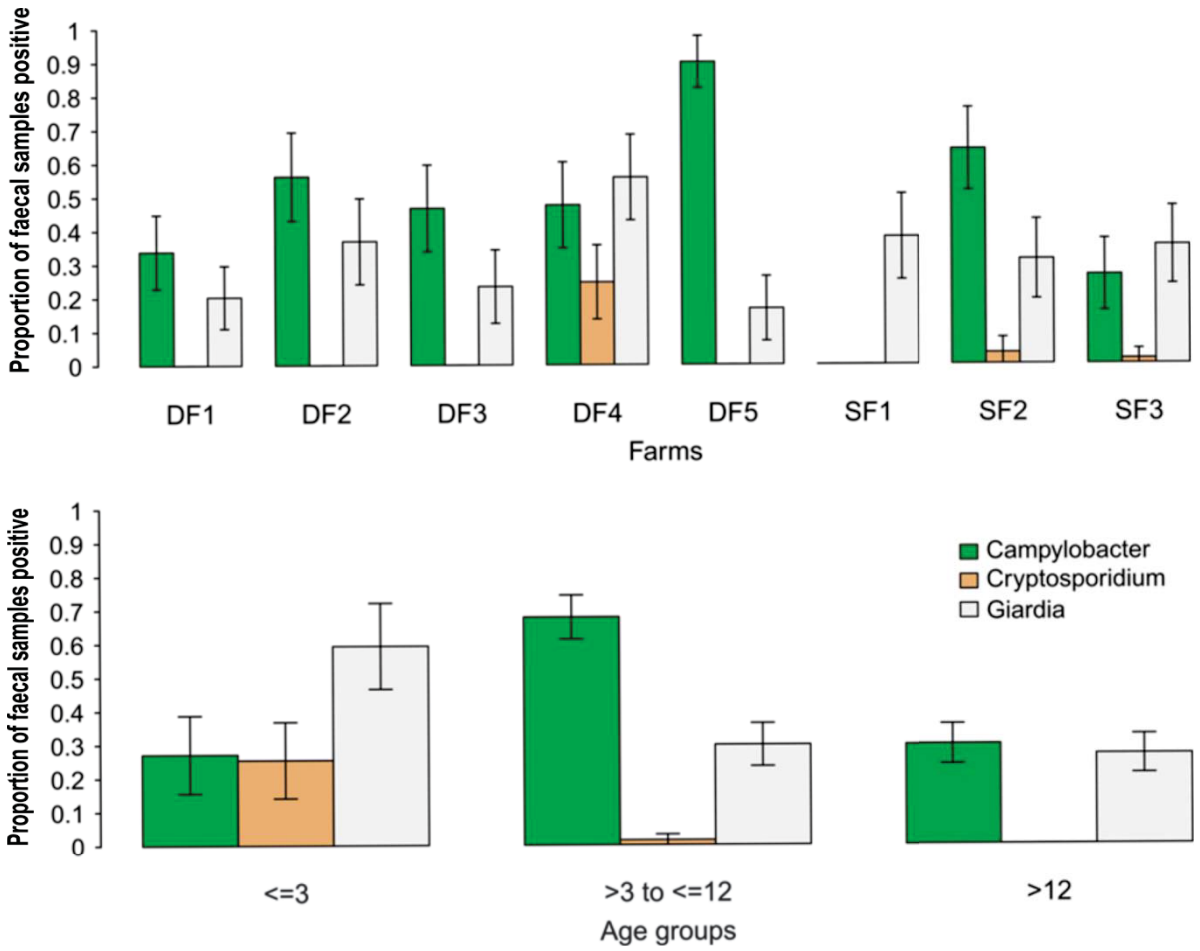


Figure 3.2: Distribution of *Campylobacter*, *Cryptosporidium* and *Giardia* detected in a) each farm from location Dannevirke (DF) and Shannon (SF) (Upper graph) and b) faecal samples from different age groups of animals (lower graph).

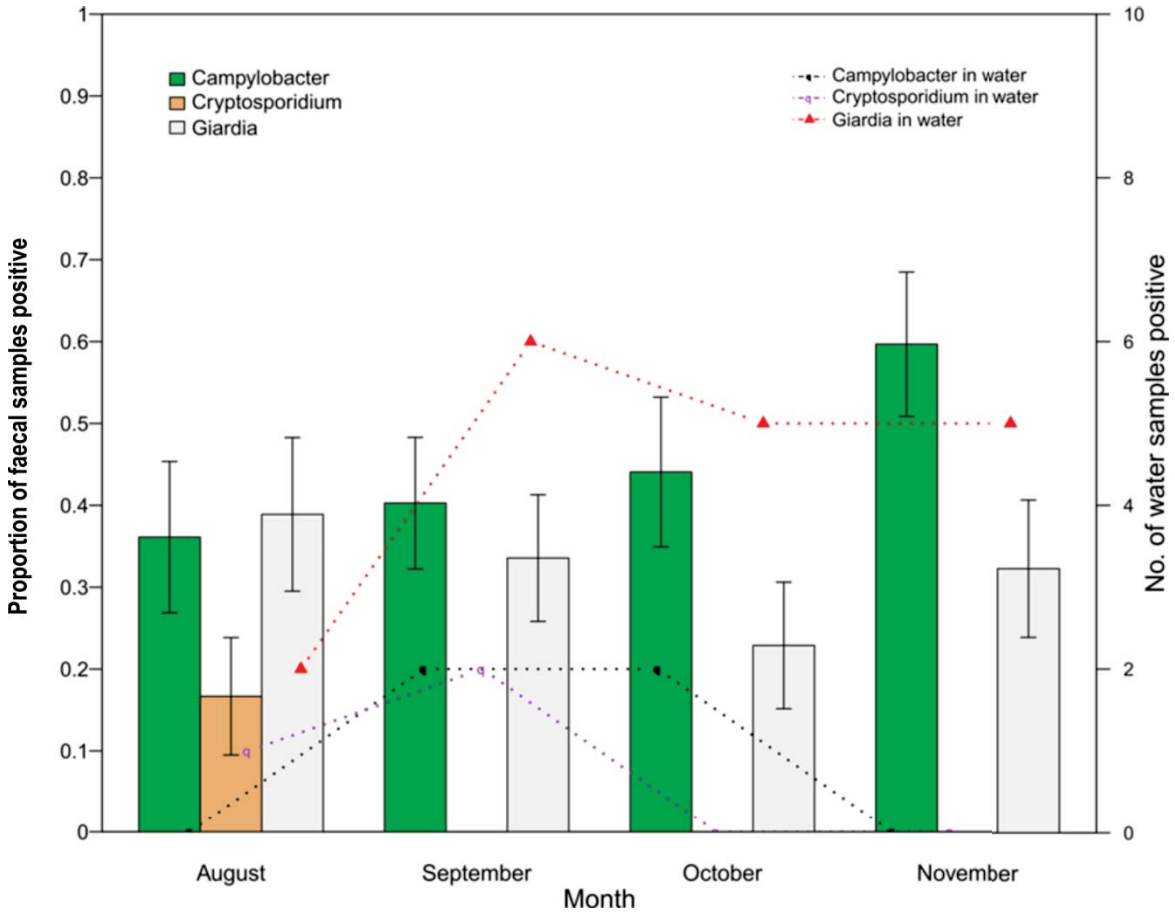


Figure 3.3: Proportion of *Campylobacter*, *Cryptosporidium* and *Giardia* detected in faecal samples collected in different months from August to November 2011. The lines show the number of water samples positive for each pathogen on that farms.

3.5.3 Pathogens in water samples

Presumptive *Campylobacter*, *Cryptosporidium* and *Giardia* spp. were detected in 4/24 (17%), 3/24 (13%) and 18/24 (75%) of the water samples collected from Tamaki River and Mangaore Stream in the two catchment areas. The number (%) of samples positive for *Campylobacter*, *Cryptosporidium* and *Giardia* spp. is presented in Table 3.7. The distribution of the pathogens in water and faecal samples across each month of the sampling period is shown in Figure 3.3.

Table 3.8 shows the estimated average sensitivity (recovery efficiency) of the USEPA method 1623 in detecting *Cryptosporidium* oocysts and *Giardia* cysts on the surface of the water samples used in this study. The sensitivity in detecting *Cryptosporidium* oocysts and *Giardia* cysts in water samples from both sites was relatively low compared to the USEPA method. The possible number of (oo)cysts that could be present in Tamaki River and Mangaore Stream water samples during the sampling period is shown in Table 3.8. The sensitivity for *Giardia* spp. detection in Shannon water samples was higher than in Dannevirke and the estimated average probable number of cysts in the Shannon water sample was 5 times more than that of the Dannevirke sample.

Based on the estimated sensitivity (recovery efficiency), the median concentration of *Cryptosporidium* oocysts in the positive water samples was 0.15 (Range: 0.11 to 0.38) oocysts/L while the median concentration of *Giardia* cysts was 0.16 (Range: 0.09 to 1.54) cysts/L. The median concentration of cysts in the storm water samples was 0.47 (Range: 0.12 to 1.54) cysts/L.

At the estimated recovery efficiency, the detection limit of the USEPA 1623 method for detecting *Cryptosporidium* and *Giardia* in Tamaki River water was 0.13 oocysts/L and 0.08 cysts/L, and in Mangaore Stream was 0.08 oocysts/L and 0.06 cysts/L.

Table 3.7: Percentage of regular and storm water samples found positive for presumptive *Campylobacter*, *Cryptosporidium* and *Giardia*. The water samples were collected from an abstraction point of the Tamaki River in Dannevirke and from the Mangaore Stream at the Shannon water treatment plant.

Organisms	**No. of samples positive (%)**			
	Tamaki River (n=13)		**Mangaore Stream (n=11)**	
	Regular water (n=10)	Storm water (n=3)	Regular water (n=9)	Storm water (n=2)
***Campylobacter* spp.**	4/10 (40%)	0/3 (0%)	0/9 (0%)	0/2 (0%)
***Cryptosporidium* spp.**	2/10 (20%)	1/3 (33.3%)	0/9 (0%)	0/2 (0%)
***Giardia* spp.**	8/10 (80%)	3/3 (100%)	6/9 (67%)	1/2 (50%)

Table 3.8: Sensitivity (recovery rate) and a probable number of *Cryptosporidium* oocysts and *Giardia* cysts present in the Tamaki River, Dannevirke and Mangaore Stream, Shannon during water sampling.

Locations/ organisms	Recovery rate ± sd [a] (Credible Interval)[b]	Probable Number of (oo)cysts ± sd (Credible Interval)
Dannevirke		
Cryptosporidium spp.	7.6 ± 0.009 (6.00 – 9.00)	6.5 ± 2.6 (2.7 - 12.7)
Giardia spp.	13 ± 0.012 (11.00 – 15.00)	25.6 ± 4.4 (18.1 – 35.2)
Shannon		
Cryptosporidium spp.	11.7 ± 0.012 (9.00 – 14.7)	1.7 ± 0.76 (1.0 - 3.8)
Giardia spp.	17.3 ± 0.013 (15.00 – 20.00)	9.7 ± 2.1 (6.3 – 14.4)

[a] sd : standard deviation

[b] 95% Credible interval

3.5.4 Genotyping and subtyping analysis

3.5.4.1 *Campylobacter* genotyping

Table 3.9 reports the percentages of *C. jejuni* and *C. coli* that were isolated from the faecal samples of different ruminant species collected from the two catchment areas. Of the 220 *Campylobacter* PCR genus positive isolates, the highest number of *C. jejuni* were isolated from faecal samples from the >3 to ≤12 month age group (34.5%) followed by samples from the >12 month age group (12.3%). Conversely, *C. coli* was isolated at a greater frequency from samples from the >12 month age group (9.1% of 235) followed by samples from the >3 to ≤12 age group (5.5% of 205). Both *C. jejuni* (3.2%) and *C. coli* (1.4%) isolated were the lowest in the ≤3 month age group samples (n=59). Among the farms, the highest numbers of *C. jejuni* were identified in the faecal samples from Shannon Farm 2 (12.3% of 110), followed by Dannevirke Farm 5 (12/110; 10.5%) and Farm 4 (11/110; 10.0%). *C. coli* were detected at a greater frequency in the faecal samples from Dannevirke Farm 2 (38.6% of 57) than other 7 farms. The majority, 27 of 35, of *C. coli* isolated were originated from sheep faecal samples.

All the four water samples positive for *Campylobacter* genus PCR were identified and confirmed as *C. jejuni* by PCR.

Table 3.9: Total number and percentage of presumptive *Campylobacter, C. jejuni*, *C. coli,* in three different species' faecal samples collected from the ground on farms of the two catchment areas: Dannevirke and Shannon of the Manawatu region in New Zealand. A "-" denotes the pathogens are not detected in the samples.

Location/ Species	Number of faecal samples	Presumptive *Campylobacter*	PCR confirmed[d] *Campylobacter*	*C. jejuni*[c]	*C. coli*[c]
Dannevirke	312	168 (53.8%)	165 (52.6%)	77 (24.7%)	32 (10.3%)
Dairy cattle	121	83 (68.6%)	83 (68.6%)	37 (30.6%)	6 (5%)
Beef cattle	53	28 (52.8%)	26 (49.1%)	19 (35.8%)	-
Sheep	138	57 (41.3%)	56 (40.6%)	21 (15.2%)	26 (18.8%)
Shannon	187	57 (30.5%)	56 (29.9%)	33 (17.6%)	3 (1.6%)
Beef cattle	68	40 (58.8%)	39 (57.4%)	19 (27.9%)	2 (2.9%)
Sheep	119	17 (14.3%)	17 (14.3%)	14 (11.8%)	1 (0.8%)
Grand Total	**499**	**225 (45.1%)**	**221 (44.3%)**	**110 (22.0%)**	**35 (7.0%)**

[a]A sample was considered presumptive *Campylobacter* positive when samples cultured on mCCDA and BA showed colonies with typical *Campylobacter* morphology.

[b]A sample was confirmed positive for *Campylobacter* if at least one presumptive *Campylobacter* isolate was positive for genus PCR.

[c]A sample was confirmed positive for *C. jejuni* or *C. coli* positive if at least one genus positive *Campylobacter* isolate was positive for *C. jejuni* or *C. coli* by PCR

3.5.4.2 *Cryptosporidium* and *Giardia* genotyping and subtyping

The *Cryptosporidium* genotypes found in this study are illustrated in Table 3.10 and the *Giardia* genotypes and assemblages in Table 3.11. PCR analysis of 18S SSU RNA and HSP70 genes confirmed *Cryptosporidium parvum* presence in 5/18 and 14/18 *Cryptosporidium* isolates, respectively. The 18S SSU RNA analysis, however, also found *C. bovis* in the three specimens in which *C. parvum* were identified using the HSP70 gene analysis. Three *C. parvum* 18S SSU RNA sequences were 100% identical to accession No: DQ010952 sequences, whereas two *C. parvum* 18S SSU RNA sequences were new and deposited in the GenBank (Accession number KF840580 andKF840581). Similarly, 18S SSU RNA *C. bovis sequences* were 100% identical to accession number AY120911.1 and AY741305 sequences in the GenBank. All the 14 *C. parvum* sequences of HSP70 gene were 100% identical to the *C. parvum* HSP70 gene sequence in the conserved region (Genbank Accession number U11761.1). Only *C. parvum* were detected in the 3 isolates obtained from faecal samples of sheep and beef cattle in the 3 to12 month age group, whereas, 15 *C. parvum* and 3 *C. bovis* isolates were originated from dairy calves in the ≤3 months age group.

Of the 159 IFA positive *Giardia* samples, *Giardia duodenalis* presence was confirmed in 70.4% of the samples using a glutamate dehydrogenase (GDH) locus PCR. The highest percentage of *G. duodenalis* was detected in the faecal samples collected from Dannevirke Farm 4 (23.2% of 112), and the isolates from the adult animals' faeces (40.2% of 112). In contrast, the lowest percentage of *G. duodenalis* was identified in the faecal samples from Dannevirke Farm 5 (3.6% of 112) and the isolates from the young animals' faeces (24.1% of 112). Of the 112 *G. duodenalis* isolates, 80 (73.2%) were assigned to assemblages A, B and E, and the remaining isolates were not amplified. Assemblage E was found at a greater frequency (87.5% of 80), mainly originating from sheep faeces (62.9% of 70).

The comparison between different gp60 subtypes of *C. parvum* and *Giardia* assemblages found in this study and New Zealand databases are shown in the Table 3.12. Two very common *C. parvum* gp60 allelic types, IIA18G3R1 and IIA19G4R1, found in bovines and humans of New Zealand were also isolated in this study. On the Dannevirke farms, both the *C. parvum* subtypes IIA18G3R1 and IIA19G4R1 were found on, whereas, *C. parvum* subtype IIA18G3R1 was only identified on the Shannon farms. In addition, the very common *Giardia* assemblages AII and BIV isolated from New Zealand human samples were also detected in bovine and ovine samples from this study (Table 3.12).

No *Cryptosporidium* isolates from water samples were successfully genotyped whereas only seven *Giardia* isolates from surface water samples were successfully genotyped as *G. duodenalis*. Subtyping of *G. duodenalis* isolates showed the presence of four types of subassemblages belonging to three assemblages A, B and E (Table 3.12).

Table 3.10: Number and percentages of *Cryptosporidium* genotypes determined by PCR and DNA sequencing of immunofluorescence assay (IFA) positive isolates detected in dairy, beef cattle, and sheep faecal samples collected from the ground on farms within the two catchment areas: Dannevirke and Shannon of the Manawatu region in New Zealand. A "-" denotes the pathogens were not detected in the samples (in brackets, %).

Locations/Species	**Number of faecal samples**	**IFA positive *Cryptosporidium***	***C. parvum*[a]**	***C. bovis*[b]**
Dannevirke	312	15 (4.8%)	15 (4.8%)	3 (1%)
Dairy cattle	121	15 (12.4%)	15 (12.4%)	3 (2.5%)
Beef cattle	53	-	-	-
Sheep	138	-	-	-
Shannon	187	3 (1.6%)	3 (1.6%)	-
Beef cattle	68	1 (1.5%)	1 (1.5%)	-
Sheep	119	2 (1.2%)	2 (1.2%)	-
Grand Total	499	18 (3.6%)	18 (3.6%)	3 (0.6%)

[a] A sample was confirmed positive for *C. parvum* if at least one isolate was positive for either HSP70 or 18S SSU rDNA PCR.

[b] A sample was confirmed positive for *C. bovis* if at least one isolate was positive for 18S SSU rDNA PCR.

Table 3.11: Number of *Giardia* assemblages identified through PCR and DNA sequencing analysis of by immunofluorescence assay (IFA) positive *Giardia* isolates detected in dairy, beef cattle, and sheep faecal samples collected from the ground on farms within the two catchment areas: Dannevirke and Shannon of the Manawatu region in New Zealand. A "-" denotes the pathogens were not detected in the samples (in brackets, %). Only numbers are shown for the identified assemblages of *Giardia*.

Locations/ Species	Number of faecal samples	IFA positive *Giardia*	PCR positive *Giardia*[a]	*Giardia* Assemblages[b]							
				Number determined	AI	AII	B	BII	BIII	BIV	E
Dannevirke	312	94 (30.1%)	65 (20.8%)	49 (15.7%)	-	-	1	1	-	2	45
Dairy cattle	121	44 (36.4%)	30 (24.8%)	21 (17.4%)	-	-	1	1	-	-	18
Beef cattle	53	11 (20.8%)	5 (9.4%)	2 (3.8%)	-	-	-	-	-	1	2
Sheep	138	39 (28.3%)	30 (21.7%)	26 (18.8%)	-	-	-	-	-	1	25
Shannon	187	65 (34.8%)	47(25.1%)	33 (17.6%)	1	4	1	1	1	1	25
Beef cattle	68	20 (29.4%)	14(20.6%)	10 (14.7%)		1	1	1	1	1	6
Sheep	119	45 (37.8%)	33(27.7%)	23 (19.3%)	1	3	-	-	-	-	19
Grand Total	**499**	**159 (31.9%)**	**112 (22.4%)**	**82 (16.4%)**	**1**	**4**	**1**	**2**	**1**	**3**	**70**

[a] A sample was confirmed positive for *Giardia* if at least one isolate was positive for GDH PCR.

[b] A *Giardia* assemblages were determined for the GDH PCR positive isolates that were sent for DNA Sanger sequencing.

Table 3.12: Number of *Cryptosporidium* parvum gp60 allelic types and *Giardia* assemblages identified in this study, and that are found in bovine, and humans of New Zealand (retrieved from New Zealand *Cryptosporidium* and *Giardia* sequence database). The A "-" denotes the specified subtypes not detected yet.

C. parvum gp60 allelic types	**In this study**		**New Zealand database**		***Giardia* Assemblages**	**In this study**		**New Zealand database**	
	Bovine	**Ovine**	**Bovine**	**Human**		**Bovine**	**Ovine**	**Bovine**	**Human**
IIaA14G1R1	-	-	-	3	AI	-	1	2	2
IIaA17G1	-	-	-	2	AII	1	3	1	31
IIaA15G2R1	-	-	2	-	AIII	-	-	-	-
IIaA16G3R1	-	-	3	9	AIV	-	-	-	-
IIaA18G2R1	-	-	-	2	B	1	-	-	-
IIaA18G3R1	6	1	48	168	BI	-	-	-	-
IIaA18G4R1	-	-	-	1	BII	2	-	-	-
IIaA19G3R1	-	-	-	7	BIII	1	-	4	15
IIaA19G4R1	7	-	11	81	BIV	2	1	2	132
IIaA20G3R1	-	-	6	-	E	26	44	22	-
IIaA20G4R1	-	-	-	5	F	-	-	-	1
IIaA20G5R1	-	-	-	14					
IIaA21G4R1	-	-	-	5					
IIdA23G1	-	-	-	8					
IIdA25G1	-	-	6	1					

Table 3.13: Number of *Campylobacter*, *Cryptosporidium* and *Giardia* genotypes identified through PCR and DNA sequencing analysis of isolates from the collected water samples (sites: Tamaki River, Dannevirke and Mangaore Stream, Shannon). A "-" denotes the pathogens were not detected in the samples (in brackets, %).

Organisms	Number of samples positive (%)		
	Tamaki River (n=13)	**Mangaore Stream (n= 11)**	**Total (n=24)**
Presumptive *Campylobacter*[a]	4 (30.8%)	-	4 (16.7%)
PCR confirmed *Campylobacter*[b]	4 (30.8%)	-	4 (16.7%)
C. jejuni[c]	4 (30.8%)	-	4 (16.7%)
C. coli[c]	-	-	-
IFA positive *Cryptosporidium*	3 (23.1%)	-	3(12.5%)
Cryptosporidium PCR[d]	-	-	
IFA positive *Giardia*	11 (83.6%)	7 (63.6%)	18 (75%)
PCR positive *Giardia*[e]	5 (38.5%)	3 (27.3%)	8 (33.3%)
Giardia Assemblages[f]	4 (30.8%)	3 (27.3%)	7(29.2%)
AII	1	-	1
BI	1	-	1
BII	-	1	1
BIV	2	1	3
E	-	1	1

[a] A sample was considered presumptive *Campylobacter* positive when samples cultured on mCCDA and BA showed colonies with typical *Campylobacter* morphology.

[b] A sample was confirmed positive for *Campylobacter* if at least one presumptive *Campylobacter* isolate was positive for genus PCR.

[c] A sample was confirmed positive for *C. jejuni* or *C. coli* positive if at least one genus positive *Campylobacter* isolate was positive for *C. jejuni* or *C. coli* by PCR

[d] A sample was confirmed positive for *Cryptosporidium* if at least one isolate was positive for either HSP70 or 18S SSU rDNA PCR.

[e] A sample was confirmed positive for *Giardia* if at least one isolate was positive for GDH PCR.

[f] A *Giardia* assemblages were determined for the GDH PCR positive isolates that were sent for DNA Sanger sequencing. The only number of samples positive was shown.

3.6 Discussion

Cattle and sheep have been identified as an important reservoir for the zoonotic pathogens *Campylobacter*, *Cryptosporidium* and *Giardia* spp., and may contribute significantly to contamination of watersheds (Stanley and Jones, 2003; Grinberg et al., 2005; Karanis et al., 2007; Fayer and Xiao, 2008; Xiao and Feng et al, 2008; Lal et al., 2015; Al Mawly et al., 2015a; 2015b). They may also provide a source of human and animal infection through direct contact with infected animals or through contaminated environments. This chapter has added further evidence of the statement that dairy calves are important carriers of *Cryptosporidium* spp. in the Dannevirke water catchment areas; whereas both cattle and sheep, particularly juveniles, are important carriers of *Campylobacter* and *Giardia* spp. in two of the watersheds of the Manawatu region of New Zealand. Findings of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. in the water samples also revealed the faecal contamination of surface water that is a source for drinking water. This is the first study that provides information about the prevalence and genetic diversity of *Giardia* spp. in beef cattle and sheep in New Zealand.

The aim of this study was to detect the three pathogens in the farm environment and surface water within the Dannevirke and Shannon catchment areas during the spring calving and lambing season in New Zealand. A collection of faecal samples from the ground was used by both Hunt *et al.* (2000) and Winkworth *et al.* (2008), and no attempt was made to collect the faecal samples from known individuals. Therefore, the results of this study need to be extrapolated with care as variables such as stock density, concentration of types and age of stock, the topography of the farm and climatic conditions may influence the findings of the prevalence of these pathogens in other areas of New Zealand. The majority of faecal samples collected originated from juvenile (>3 to ≤ 12 months old) (n=205) and adult (>12 months old) (n=235) animals, which could be the reason for overestimating *Campylobacter* and *Giardia* spp. prevalences, and for underestimating *Cryptosporidium* spp. Each test was repeated at least two times, however, which could have helped in cutting the bias towards false positives and negatives.

Like many New Zealand farms, the two catchment areas of this study have a pastoral farming system. In addition, many farms (7/8 farms) have their stock near waterways for more than 20 days a month, and there is unfenced land on 4/8 of the farms. Consequently, stock crossing the streams or river and loading pathogens in the water is probably unavoidable, and provides evidence that these catchment

areas possibly have diffuse-source pollution (Bagshaw, 2002 ; Wilcock, 2006). Access to contaminated surface water sources may also have infected the ruminants grazing on the pasture (Hanninen et al. 1998; Humphrey and Beckett, 1987). Therefore, there is a need to consider the improvement of farm management through stream-bank (riparian) planting, and by excluding stock from waterways using bridging and fencing, ultimately providing cleaner water for drinking purposes for both humans and animals (Collins et al., 2007; Winkworth et al., 2008).

3.6.1 Pathogens in faeces

3.6.1.1 *Campylobacter* spp.

In the present study, the prevalence of *Campylobacter* spp. (45.1%) detected lies within the broad range of prevalence reported in the literature. However, results of prevalence studies are difficult to compare because of the different isolation methods, farmtypes, age of animals, farm-management systems, geography, and sampling time. This study found a significantly higher *Campylobacter* spp. and *C. jejuni* prevalence in cattle (62.4%; 31%) than in sheep (28.4%; 13.6%) faeces respectively. This finding is in agreement with many other studies (Bailey et al., 2003; Devane et al., 2005; Milnes et al., 2008; Moriarty et al., 2011a; Sproston et al., 2011). Oporto *et al.* (2007) reported a 55% *Campylobacter* spp. prevalence (n=120) in sheep faecal samples collected from the rectum, and Stanley *et al.* (1998) reported 91.7% prevalence in small intestine samples (n=320) from lambs for slaughter. However, Stanley *et al.* (1998) demonstrated the substantial reduction in the *Campylobacter* spp. prevalence (29.3%; n=420) in sheep on pasture. This provides evidence that *Campylobacter* spp. loads on pasture in our study are comparatively lower than *Campylobacter* spp. colonised in the intestine, and reflected the fact that *Campylobacter* spp. on pasture die off or enter a viable but non-culturable state. Jones *et al.* (1999) suggested that ewes before lambing do not shed the *Campylobacter* spp. up to the detection limit. This might be the reason why *Campylobacter* spp. were not detected in Shannon Farm 1 (contains sheep only) of this study. Therefore, absence of *Campylobacter* spp. in Shannon Farm 1 supports the hypothesis that waterways near farms containing cattle or a mix of cattle and sheep are potentially more vulnerable to faecal contamination than those containing sheep only (Close et al., 2008; Davies-Colley et al., 2004). Nevertheless, the findings of a higher proportion of *C. coli* in sheep (10.5% of 257) than in cattle (3.3% of 242) should not be overlooked because ~10% of human

campylobacteriosis cases have been attributed to *C. coli* (Nachamkin et al., 2008). In addition, *C. coli* were frequently isolated from water samples from various sources including river, stream, ponds etc. Cattle and sheep are considered the second most important source of *Campylobacter* spp. transmission to humans in New Zealand after poultry (~18-40% cases attributed to ruminants) (Mullner et al., 2009b; French and Marshall, 2012), and have also been implicated in the environmental spread of *Campylobacter* to water (Clark et al., 2003).

More dairy cattle faecal samples tested positive for *Campylobacter* spp. and *C. coli* than beef cattle samples. Conversely, *C. jejuni* prevalence was higher in the beef samples than the dairy samples. A Spanish study reported similar results with 67.1% and 3.7% of dairy cattle and 58·9% and 2.4% of beef cattle samples positive for the *Campylobacter* spp. and *C. coli,* respectively, whereas *C. jejuni* were isolated from 14.6% and 20.2% samples of dairy and beef samples respectively (Oporto et al., 2007). On the contrary, a New Zealand study demonstrated a higher prevalence of *Campylobacter* spp. and *C. jejuni* in dairy cattle than beef cattle (Devane et al., 2005). The cattle are known to shed *Campylobacter* spp. intermittently, which might be one of the reasons for the differences in *Campylobacter* prevalence in faeces. Beef cattle are not considered as important a reservoir for human campylobacteriosis as dairy cattle. This finding was concluded based on the finding of low *Campylobacter* prevalence (<5%) in retail beef samples (Bohaychuk et al., 2006; Wong et al., 2007; Llarena et al., 2014). However, Inglis *et al.* (2004) demonstrated that beef cattle could chronically shed *Campylobacter* spp. In New Zealand, around 3.7 million beef cattle and 6.6 million dairy cattle graze in the open pasture (Meat and Wool New Zealand, 2013; Statistics New Zealand, 2013). Thus, an environmental load of *Campylobacter* through beef cattle should not be ignored, particularly if animals are grazing near waterways, as research shows that *Campylobacter* spp. can be detected for up to 14 days in bovine faeces on pasture (Gilpin et al., 2008), and can transform itself into a viable but non-culturable state in cool weather, prolonging *Campylobacter* spp. viability.

The prevalence of *Campylobacter* spp. was higher in juvenile calves (67.8%) than in adults (29.8%) and young (≤3 month) calves (27.1%) in this study. The results contradict other studies stating that young calves have a higher proportion of *Campylobacter* spp. positive. For example, Nielsen (2002) reported a higher prevalence of *Campylobacter* spp. in calves of ≤ 4 months than in young and adult cows. She also demonstrated the increase in the *Campylobacter* spp. prevalence from 14% in calves ≤31 days old to 80% in calves aged 91- 120 days. Similarly,

Jones *et al.* (1999) found lambs colonised with *Campylobacter* spp. 5 days after birth. Young animals might be more susceptible to colonisation with *Campylobacter* spp. in the intestine than adult animals (Giacoboni et al., 1993).Studies show that healthy adult cattle were found to shed *Campylobacter* spp. and ewes were found to shed up to 5 $\log_{10}$ g^{-1} *Campylobacter* spp. after lambing (Jones et al., 1999; Fayer and Xiao, 2008; Rapp et al., 2012) which might be the reason for the higher *Campylobacter* spp. levels in the juvenile and adult animals in our studies.

There was a significant difference between the Dannevirke and Shannon catchment areas concerning isolated thermophilic *Campylobacter* spp. In this study, we focused entirely on faeces found in the environment upstream from the catchment areas. In the Shannon catchment areas, faecal samples originated only from sheep and beef cattle that have been known to shed less *Campylobacter* spp. than dairy cattle (Close et al., 2008; Davies-Colley et al., 2004). This finding shows that the surface water of the Shannon catchment area is less likely to be contaminated with *Campylobacter* species. In New Zealand, both spatial and temporal variations were found in human campylobacteriosis cases (Hearnden et al., 2003; Rind and Pearce, 2010). The South Island has higher notifications rates than the North Island. Within the North Island, the cases are more concentrated in the urban areas and in the intensive dairy farming areas. Additionally, infections are low in winter (June to August), increasing in spring (September to November) and peaking sharply in summer (December to February) (ESR, 2014). Our study has a similar pattern of *Campylobacter* spp. prevalence in faeces, with the lowest month being August and the highest being November. However, this result must be treated with caution because it might be influenced by the inclusion of more young calves or lambs samples in the latter sampling occasion months. No seasonal pattern was observed (data not presented) when the sampling occasion (range: 45-47.5%) was taken into consideration rather than the month of sampling.

PCR confirmed the presence of thermophilic *Campylobacter* spp. in 97.8% of presumptive *Campylobacter* samples in this study. However, only 50% and 16% of PCR-confirmed *Campylobacter* samples were identified as *C. jejuni* and *C. coli,* suggesting the possibility of other species of *Campylobacter* being present in the remaining isolates. The identification of other *Campylobacter* spp. was out of the scope of this study, and thus was not performed. *C. jejuni* and *C. coli* identification in the current study suggests the likelihood of zoonotic transmission to people within the catchment areas through water or contact with animals, Without subtyping the

C. jejuni and *C. coli*, however, the actual risk of *Campylobacter* spp. remains unknown.

3.6.1.2 *Cryptosporidium* spp.

Cryptosporidium spp. prevalence reported in this study (3.6% from 3/8 farms) was low when compared to many New Zealand and international studies (Oslon et al., 1997; Learmonth et al., 2003; Grinberg et al., 2005; Moriarty et al., 2008; Izzo et al., 2011). *Cryptosporidium* spp. were reported in 21.2% calves from 10 out of 24 farms in a 2002 cross-sectional survey of calves conducted in the Manawatu region of New Zealand during calving season (Grinberg et al., 2005). Similarly, a national level survey of 97 farms in New Zealand reported IFA-positive *Cryptosporidium* spp. in 6% (25/429)of the specimens from calves aged 1-5 days and in 22% (178/797) of the specimens from calves aged 9-21 days (Al Mawly et al., 2014). These studies examined samples collected from dairy calves aged <21 days, unlike the current study that includes different age groups of dairy, beef and sheep faecal specimens. Cattle are known to shed a higher frequency of *Cryptosporidium* spp. than sheep (Kváč et al., 2006; Fayer and Xiao, 2008; Caccio and Widmer, 2014). Typically, the highest prevalence was observed in calves aged between 2 and 6 weeks, declining as the age of the calves increased (Santín et al., 2004, Feng et al., 2007; Maikai et al., 2011; Rieux et al., 2013). In our study, too, 15 out of 18 *Cryptosporidium*-positive specimens originated from 5 to 6 -week -old dairy calves, two from 4- month- old lambs and one from a 1-year- old beef cow. Surprisingly, *Cryptosporidium* spp. were not detected in any of the subsequent sampling specimens, indicating that the animals might be shedding oocysts below the detection limit or might not be shedding any oocysts at the time of sampling. This may lead to an underestimation of the *Cryptosporidium* spp. prevalence (Silverlås et al., 2013; Blewett et al.,1993; Ortega-Mora and Wright, 1994). Longitudinal studies showed that infected cattle shed oocysts over an average of 12 days (Castro-Hermida et al. 2002; Santı´n et al. 2008). Other factors that may affect the prevalence are the shorter prepatent period of *Cryptosporidium* spp.; the study design such as the fact that samples were collected only from animals near waterways; diagnostic methods used; climatic conditions; and farm management systems (Geurden et al., 2007; Giadinis et al., 2012; Ryan et al., 2005; Santín et al., 2008; Silverlås et al., 2010; Britton et al., 2010; Budu-Amoako et al., 2012a). Further research on *Cryptosporidium* spp. prevalence in ruminants on pasture in New Zealand is warranted, however, as indicated by the low *Cryptosporidium* spp. prevalence reported in faecal specimens collected from pastures of 2 out of 4 farms (5.2%; n=155) during the autumn of 4 sampling

seasons (Moriarty et al., 2008), and from 354 dairy cows (0.6%) in the Waikato region of New Zealand (Learmonth et al., 2005).

PCR analysis of 18S SSU RNA and HSP70 genes confirmed the presence of *Cryptosporidium parvum* in all *Cryptosporidium* isolates, with *C. bovis* co-existence in three calves' specimens. PCR analysis of HSP70 genes to identify *C. bovis* in the samples was unknown prior to analysing the data in this study. Therefore, two *Cryptosporidium* spp. in one specimen suggest the potential presence of a mixed infection. Mixed infections with more than two *Cryptosporidium* species have been reported in many studies (Rzeżutka and Kaupke, 2013; Silverlås et al., 2013; Tanriverdi et al., 2003). Such infections are usually considered an exception rather than the rule, however, so further molecular epidemiology studies are required to make conclusive remarks (Kurniawan et al., 2013; Silverlås et al., 2013; Yang et al., 2009).

C. *parvum* is mostly identified in pre-weaned and *C. bovis* in post-weaned dairy cattle (Santı´n et al. 2004, 2008; Brook et al., 2009). In sheep, *C. bovis* and *C. cervine* have commonly been isolated (Pritchard et al., 2008, Mueller-Doblies et al., 2008), and recently *C. ubiquitum*, *C. xiao* and *C. parvum* have been frequently reported in sheep (Fayer et al., 2010; Fayer and Santı´n, 2009; Pritchard et al., 2007; Dı´az et al., 2010). *C. parvum* was recognised as a primary pathogen of neonatal diarrhoea in both cattle and sheep and responsible for economic losses because of mortality and retarded growth after recovery. (Tzipori et al. 1980; Baker and Carbonell 1974; Castro-Hermida et al. 2008). Therefore, *Cryptosporidium* spp. circulating in the cattle and sheep of the study areas exhibits not only a possible economic burden to farmers but also on-going *Cryptosporidium* spp. contamination of the environment.

Additionally, isolation of *Cryptosporidium parvum* in this study suggests the potential for zoonotic transmission of *Cryptosporidium* from cattle and sheep in the catchment areas. *C. parvum* is also known to infect humans (Xiao, 2010). In New Zealand, cryptosporidiosis is the third most notified gastrointestinal disease, often associated with recreational water and contact with farm animals (ESR, 2014). Genetic characterisation of *Cryptosporidium* spp. from human isolates (n=423) demonstrated the presence of 53% *C. parvum* and 47% *C. hominis*, the former related to isolates from rural areas and the latter from urban areas (Learmonth et al., 2004). The New Zealand protozoa database also showed that the majority (68%; 381/560) of human isolates were *C. parvum*. Two *C. parvum* gp60 allelic types, IIA18G3R1 and IIA19G4R1 found in this study were also identified in human and

bovine samples in New Zealand, Australia, United Kingdom and Canada (Xiao, 2010). This further confirmed the potential risk of zoonotic transmission in the study areas.

3.6.1.3 *Giardia spp.*

In this study, there was not a significant difference between *Giardia* spp. prevalences in cattle (31%; 75/242) and in sheep (32.7%; 84/257), and the prevalence is within the range found elsewhere in the world. *Giardia* spp. have been reported worldwide in both cattle and sheep, and point prevalences vary considerably. In cattle, the reported prevalence ranged between 9 and 73%, and in sheep between 1.5% and 43% (Geurden and Olson, 2011). In New Zealand, only a handful of *Giardia* spp. studies have been performed in ruminants, particularly in dairy cattle, with a prevalence ranging from10.5 to 41% in calves and 4.5% in adult cattle (Hunt et al., 2000; Moriarty et al., 2008; Winkworth et al., 2008; Learmonth et al., 2004). The current study found a 36.4% (44/121) *Giardia* spp. prevalence in dairy cattle, which was at similar levels to those found in dairying regions in the North (41%: Hunt et al., 2000) and South (31%: Winkworth et al., 2008) Islands. The comparable prevalence findings between the two catchment areas of this study and between Islands in previous studies support the argument that behavioural factors influence the prevalences (Winkworth et al., 2008; Becher et al., 2004). Nevertheless, similar management practices such as maintaining cows in fields the entire year, starting spring-calving seasons at similar times and transferring calves to farm paddocks from pens within a few weeks of being born may have influenced the prevalences. Within cattle, point prevalences of *Giardia* spp. were reported between 9 and 93% whereas it was between 8.7 and 37.3% in beef cattle within North America (Olson et al., 1997a; 1997b; O'Handley et al., 2000; Fayer et al., 2000; Appelbee et al., 2003; Trout et al., 2004; 2005; McAllister et al., 2005; Trotz-Williams et al., 2005; Gow and Waldner, 2006; Coklin et al., 2007; 2009). Our studies also found lower *Giardia* spp. prevalences in beef cattle compared to dairy cattle. Although *Giardia* spp. prevalences have been reported in beef cattle and sheep in international studies (Geurden et al., 2008; Robertson et al., 2010), it is surprising that no studies have been conducted in the New Zealand context, given the potential of *Giardia* spp. to exert a pathogenic effect on beef cattle and sheep and to harbour zoonotic genotypes of this parasite.

The excretion of *Giardia* cysts is highly variable between animals, chiefly during a chronic infection. Thus, farm prevalence is suggested to be more informative than

animal prevalence in the cross-sectional study (Geurden and Olson, 2011). In cattle, farm prevalence of *Giardia* spp. was reported between 45 and 100% and in sheep between 0 and 100% (Geurden and Olson, 2011). In this study, as in the North American, European and New Zealand studies, *Giardia* spp. was found in all the farms similar to other studies (Hunt et al., 2000; McAllister et al., 2005; Maddox-Hyttel et al., 2006; Trout et al., 2006). This result implies that every animal on the farms of the two catchment areas will get infected at some point. This further leads to a conclusion that there is a potential risk of surface water contamination in those two catchment areas at any point in time.

A significant finding of this study was that *Giardia* spp. prevalence decreases with age, the highest being in young animals (≤3 months: 59%) although relatively small numbers of samples from sheep and dairy cattle aged ≤3 months were examined. Several studies have demonstrated that *Giardia* spp. infections occur at the end of the neonatal period of cattle with a higher prevalence of *Giardia* spp. in post-weaned (50%) than pre-weaned (40%) calves and adult cattle (28%) (O'Handley and Olson, 2006; Trout et al., 2005; 2007; Winkworth et al., 2008). These previous reports are consistent with our data, which show that 69% of dairy calves, 21.3% of juvenile cattle, 0% of adult dairy cattle, 33.7% of juvenile, and 8% of adult beef cattle were infected by *Giardia* spp. Similarly, lambs are more often infected with *Giardia* spp. than ewes and adult sheep (Ryan et al., 2005; Santín et al., 2007). However, this study showed a lower prevalence in young lambs (≤ 3 months: 28.6%) than in juveniles (32.6%) and adults (33%), reflecting the effect of the low number of lamb samples (n=14). Nevertheless, *Giardia* spp. infection in ruminants can vary markedly at any age period, implying that there is a constant risk of *Giardia* spp. contamination in the environment.

Giardia duodenalis is only one of six species that infect humans. Three distinct genetic groups or assemblages (A, B and E) of *G. duodenalis* have been reported in cattle and sheep (Ryan and Caccio, 2013). Our study also reported assemblages A, B and E in cattle and assemblages A and E in sheep. Assemblages A and B have a wide host range including humans whereas assemblage E has been found only in hoofed livestock and is frequently reported in cattle and sheep. Assemblage B represents the second *G. duodenalis* genotype found in humans and has only recently been identified in cattle (Lalle et al., 2005; Mendonça et al., 2007: Coklin et al., 2007: Winkworth et al., 2008). Our studies showed a higher prevalence of *Giardia duodenalis* assemblage E in cattle than assemblage A, which is in agreement with other studies (Abeywardena et al., 2012; O'Handley et al., 2000; Trout et al., 2004;

2005; 2007). However, Geurden *et al.* (2008) and Santín *et al.* (2009) reported a higher prevalence of assemblage A than E in dairy calves than previously reported, suggesting that dairy calves may pose a greater risk of zoonotic transmission than previously thought. Similarly, *Giardia duodenalis* assemblage E is more prevalent in sheep compared to assemblage A in this study and other studies (Santin et al., 2007; Feng and Xiao, 2011; Ryan et al*.,* 2005). *Giardia duodenalis* assemblage B is not commonly reported in cattle and sheep, but has been reported in other countries such as in Canada, USA, China, Italy, Spain and Norway (Liu et al., 2012, Uehlinger et al., 2011, Aloisio et al., 2006; Castro-Hermida et al., 2007, 2011; Robertson et al., 2010; Zhang et al., 2012). The previous studies conducted in New Zealand identified *Giardia duodenalis* assemblage B only in cattle (Hunt et al., 2000; Learmonth et al., 2003; Winkworth et al., 2008), which contrasts with a study by Abeywardena *et al.* (2012) that did not find assemblage B in their study. There is no published report on *Giardia* assemblages identified in New Zealand sheep. However, findings of zoonotic potential assemblages A and B in cattle and assemblage A in sheep in the present study showed the heterogeneity of *Giardia* distribution on New Zealand farms. This finding also points out that the significance of New Zealand's dairy herd and sheep flock as a potential reservoir of zoonotic *Giardia duodenalis* is probably greater than that of cattle and sheep in other countries. In addition, *Giardia duodenalis* caused 1729 reported cases in people in New Zealand (Unpublished data). Environmental transmission, including transmission in water, is the third most important mode of *Giardia* spp. transmission in New Zealand, after person-to-person and exposure to faecal material. Therefore, further molecular epidemiological studies on *Giardia* spp. in sheep and cattle are necessary for a better understanding of the dynamics of *Giardia* transmission in the New Zealand context.

3.6.2 Pathogens in water samples

Campylobacter, *Cryptosporidium* and *Giardia* spp. have been detected in samples from source water for drinking, drinking water, recreational water and other water sources (Mackenzie et al., 1994; Slapeta, 2013; Fayer et al., 2004; Feng et al., 2011; Helmi et al., 2011; Horman et al., 2004). These three pathogens were also detected in the present study. However, *Campylobacter* and *Cryptosporidium* spp. were detected in only a few samples (4 and 3) compared to *Giardia* spp.which were detected in 18 water samples. The percentage of samples positive for *Campylobacter* spp. (16.7%) was comparable with our previous study (21%;

Chapter 3). In both studies, samples were isolated from rivers and/or streams. Other studies have reported *Campylobacter* spp. in up to 70% of water samples. *Cryptosporidium* spp. have been known to be present in water samples for a long time; however, detection of these pathogens in water samples from rivers and/or streams varied widely (Fayer and Xiao, 2008). The median concentration of oocysts was comparatively lower than some studies (Scott et al., 2003; Payment et al., 2001; Ongerth and Saaed, 2012; Lee et al., 2014). Similarly, *Giardia* spp. were detected at a greater frequency in studies worldwide but the estimated concentration of *Giardia* cysts was also much lower than studies conducted in other countries (Scott et al., 2003; Payment et al., 2001; Ongerth and Saaed, 2012; Lee et al., 2014). Storm water samples have been investigated in many countries to assess the runoff effect on pathogen concentration (Xiao et al., 2000; Miller et al., 2008; Jiang et al., 2005). This study also collected storm water samples and detected *Cryptosporidium* spp. However, it was possible only to estimate the concentration of *Giardia* cysts as *Cryptosporidium* oocysts were detected in only one sample using the immunofluorescence assay.

The lower prevalences of pathogens and their concentration in this study must be interpreted with caution because many factors such as sampling volume, loss of pathogens during sample processing and sensitivity of the test employed can vary the reported estimate (USEPA, 2005). Use of enrichment culture technique is the standard method for isolating presumptive *Campylobacter*. However, *Campylobacter* under stress can change into a viable but non-culturable state (VBNC), thus giving negative results. Direct PCR of water samples has been suggested for *Campylobacter* detection, but some water properties such as high turbidity may interfere with the detection of *Campylobacter* spp. Although *Cryptosporidium* and *Giardia* spp. have been identified in water for a long time, there are no highly sensitive gold standard methods available for identifying these pathogens in water samples. In addition, the current methods employed are labour-intensive and expensive to perform. Therefore, these pathogens are not regularly monitored by the drinking water industry despite the fact that the (oo)cysts are very small in size, highly resistant, do not get killed easily and that outbreaks of waterborne diseases are extremely common. Although they are monitored, the oocysts lost during sample processing and concentration may limit detection. USEPA 1623 is the internationally accepted latest method that was utilised in this study. Yet, the sensitivity of the test varies widely depending on the laboratory, the person's expertise in using the method and water quality parameters. High

turbidity in water is held responsible for not detecting these two protozoa, as they get lost during sample processing. During some water sampling events, the filters collected in this study were blocked due to turbid water, which could be one of the reasons for detecting *Cryptosporidium* oocysts in fewer samples, besides their loss during processing. In addition, it might be possible that the (oo)cysts present in rivers are below the level of detection during sampling time.

3.7 Conclusion

This epidemiological study provides robust information on the occurrence of the three potentially zoonotic pathogens in ruminant faeces and water in two catchment areas of New Zealand, during calving/lambing period. *Campylobacter* and *Giardia* spp. circulate on the majority of the farms, and there is a possibility of *Giardia* spp. contamination in river/streams during calving and lambing period. Dairy cattle shed these three pathogens at a greater frequency than beef cattle and sheep. *Cryptosporidium* and *Giardia* spp. were the dominant pathogens detected in young animals (≤ 3months) compared to juvenile or adult cattle or sheep; whereas juvenile shed *Campylobacter* spp. at a greater frequency than animals of other age groups. *C. jejuni* was the dominant species identified in both cattle and sheep, whilst the majority of *C. coli* identified originated from sheep. Although *C. parvum* was the dominant species identified, findings of other *Cryptosporidium* spp. warrant further investigation on mixed infections. Similarly, *Giardia duodenalis* was the dominant species detected in both cattle and sheep. Detection of zoonotic potential *C. parvum* subtypes and *G. duodenalis* assemblages in water catchment areas reinforce the diagnostic value of genotyping/subtyping. Although these three pathogens were detected in water samples, further investigation on the zoonotic potential of those pathogens and methods employed are warranted. Nevertheless, the results of this study add useful information to improve understanding of the dynamics of transmission of the three pathogens in farm environments. This information will be of strategic importance in the development of intervention strategies to deal with the contamination of surface water from farm-animal sources, by reducing the burden of *Campylobacter*, *Cryptosporidium* and *Giardia* on farms and by implementing measures aimed at decreasing the contamination of watersheds.

3.8 Highlights of this study

- *Campylobacter* and *Giardia* spp. were prevalent in a majority of the farms.
- Dairy cattle shed *Campylobacter*, *Cryptosporidium* and *Giardia* spp. at a greater frequency than beef cattle and sheep.
- Young ones (<3 months) and juvenile animals shed *Campylobacter*, *Cryptosporidium* and *Giardia* spp. at a greater frequency than adults did.
- Zoonotic potential *Campylobacter*, *Cryptosporidium* and *Giardia* spp. were detected in faecal samples that were also reported in humans in New Zealand and other countries.
- Budget limitations did not allow the collection of large numbers of water samples. There is also a Moreover, the USEPA 1623 method used for detecting *Cryptosporidium* and *Giardia* spp. has a high detection limit. Further investigation is, therefore, required before making concrete conclusions.

3.9 Acknowledgements

I would like to express my gratitude to farmers who were involved in this study, Dr. Anne Midwinter, Lynn Rogers, Sarah Moore, Anthony Pita, Niluka Velathanthri, Angy Reynolds, Errol Kwan, Tessy George and Neville Haack from mEpiLab for their help during sampling and laboratory analysis, and Horowhenua and Tararua District Councils for allowing me permission to take water samples from their premises. Thanks also go to Dr Jonathan Marshall for providing help in analysing the results.

This project was funded by IVABS Post-graduate Research Fund and Ministry of Health Fund for Aquatic Protozoa Research Unit.

Chapter 4

Infections with multiple Cryptosporidium species and new genetic variants in young dairy calves on a farm located within a drinking water catchment area in New Zealand

4.1 Preamble

In the previous study (chapter 3), we described the epidemiology of the three pathogens in water and faecal samples. In that study, during molecular analysis of *Cryptosporidium* spp. isolates, we detected different *Cryptosporidium* species in faeces from the same sample. We repeated the analysis to confirm its veracity, which resulted in this chapter. This chapter highlights the importance of iterative PCR analysis targeting more than two loci.

4.2 Abstract

Several *Cryptosporidium* species are known to infect cattle. However, the occurrence of mixed infections with more than one species and the impact of this phenomenon on animal and human health are poorly understood. Therefore, to detect the presence of mixed *Cryptosporidium* spp. infections, 15 immunofluorescence-positive specimens obtained from the faeces of 6-week- old calves on one dairy farm were subjected to PCR-sequencing at multiple loci. DNA sequences of three *Cryptosporidium* species were identified: *C. parvum* (15/15), *C. bovis* (3/15) and *C. andersoni* (1/15), and two new genetic variants. There was evidence of mixed infections in five specimens. *C. parvum*, *C. bovis* and *C. andersoni* sequences were detected together in one specimen, *C. parvum* and *C. bovis* in two

specimens, and *C. parvum* and *C. parvum*-like variants in the remaining two specimens. Sequencing of gp60 amplicons identified the IIaA19G4R1 (8/15) and IIaA18G3R1 (4/15) *C. parvum* subgenotypes. In conclusion, this study provides evidence of endemic mixed infections with the three main *Cryptosporidium* species of cattle, *C. parvum, C. bovis and C. andersoni,* in dairy calves at the transition age of six weeks. In addition, genetic evidence of the presence of *C. andersoni* and two new *Cryptosporidium* genetic variants is provided here for the first time in New Zealand. These results add to the body of evidence describing *Cryptosporidium* parasites as genetically heterogeneous populations, highlighting the need for iterative genotyping at multiple loci to explore their genetic makeup.

4.3 Introduction

Gastro-intestinal infections with *Cryptosporidium* parasites, in particular the species *C. parvum*, are well recognised in cattle. In addition, cattle are also an important infection source for humans (Fayer and Xiao, 2008). Cattle-to-human transmission of C. parvum occurs through direct contact with infected calves or indirectly, via the contamination of food and water with cattle manure or slurry (Fayer and Xiao, 2008).

Many aspects of the epidemiology of cryptosporidiosis are not well understood due to the inability to differentiate between *Cryptosporidium* taxa by their morphology or phenotype. Therefore, molecular genetics tools targeting taxonomically informative loci are widely used retrospectively, to distinguish between the taxa (Caccio et al., 2005; Sulaiman et al., 2000). To date, about 30 *Cryptosporidium* species have been recognised, many of which are associated with disease in humans, domesticated livestock, companion animals and wildlife (Santín, 2012; Šlapeta, 2013). The species so far identified in cattle are the gastric species *C. andersoni,* and the intestinal species *C. parvum, C. bovis, C. rynae, C. ubiquitum*, *C. suis, C. scrofarum* and *C. hominis* (Tanriverdi et al., 2003; Smith et al., 2005; Fayer, 2010; Ng et al., 2011; Ryan and Power, 2012).

C. parvum is a frank, zoonotic pathogen, whereas the clinical and zoonotic potential of the other species is not well understood. In cattle, infections with the different *Cryptosporidium* species tend to follow specific temporal patterns, with *C. parvum* cycling mostly during the first month of life, and the other species more commonly found in post-weaned calves (*C. bovis*; *C. ryanae*), or yearlings and adult cattle (*C. andersoni)* (Fayer et al., 2007a; 2007b; Santin et al., 2004). The reasons for this age-

specific pattern of infection are not well understood, and the level of cross-immunity between the species is not known.

Mixed infections with multiple *Cryptosporidium* species may originate from co-infections or, in the case of chronic infections, super-infections with multiple taxa (Grinberg et al., 2013). Globally, mixed *Cryptosporidium* infections have been described in humans, in particular in HIV-AIDS patients (Cama et al., 2006; Kurniawan et al., 2013), and also in animals (Rzeżutka and Kaupke, 2013; Silverlås et al., 2013; Tanriverdi et al., 2003). However, such infections are usually considered an exception, rather than the rule, and exhaustive investigations of the nature of mixed infections in animal populations are not commonly pursued in molecular epidemiological studies. Thus, the prevalence and impact of this phenomenon on animal or public health are unknown.

Whereas mixed infections may be identified by PCR-restriction fragment length polymorphism (RFLP) and other methods, they cannot be easily confirmed by Sanger sequencing, as this method tends to detect only the predominant genetic variant present in the sequenced sample (Grinberg et al., 2013; Reed et al., 2002). Thus, the detection of mixed infections is likely to be improved by the use of iterative approaches and multiple molecular diagnostic tools (Xiao, 2010). Therefore, we characterised *Cryptosporidium* parasites from calves using an iterative PCR-sequencing approach on multiple genes, with the aim of detecting mixed *Cryptosporidium* species infections in calves.

4.4 Materials and methods

4.4.1 Study design and faecal samples

This study was part of a larger study aimed at assessing the presence of waterborne enteropathogens in two water catchment areas within the Manawatu region of New Zealand. Fresh faecal specimens were sampled from the ground of a paddock on a dairy farm located in one of the catchment areas. The faecal sampling was performed once a month, from 31 August to 18 November 2011, for a total of four sampling events. This period of the year coincided with the spring calving season in New Zealand, usually starting in July and ending in October. The calves grazing in the paddock were between five and six weeks of age on the first sampling occasion. In New Zealand, calves of this age are usually weaned and held on pasture. The calves were kept in barns and fed with reconstituted milk replacer from day 2 to the

first month of their life, then moved to the paddock. About 35 calves were present on the paddock during the sampling period. Adult cattle were not present on this paddock because, after three months, farmers moved these cattle to another paddock where juvenile and adult cattle are. Fifteen faecal specimens were collected in plastic containers on each sampling event, up to a total of 60 specimens. In order to prevent cross-contamination of specimens, faeces located at least two metres apart were sampled, and disposable gloves were changed between specimens. Specimens were transported on ice to Massey University and stored at 4°C for a maximum of four days until analysed.

4.4.2 Laboratory analysis

4.4.2.1 Identification of *Cryptosporidium* oocysts

Cryptosporidium oocysts were identified by immunofluorescent microscopy (IFA) using a commercial kit (Aqua-Glo G/C Direct, Fl, Comprehensive Kit; Waterborne, Inc., New Orleans, USA). The fluorescent conjugate was diluted five-fold with molecular grade water and the stained smears were observed in epifluorescence microscope using a 460-490 nm excitation wavelength. Specimens showing more than one IFA-positive oocyst on slides were considered positive. In order to identify only parasites cycling in the host, rather than naked DNA originating from the farm environment, only IFA-positive specimens were subjected to genotyping. Genomic DNA was extracted from the specimens using faecal DNA extraction kits (Bioline, Sydney, Australia), following the manufacturer's instructions.

4.4.2.2 Identification of *Cryptosporidium* species and subgenotypes

Cryptosporidium taxa were identified using multilocus PCR-sequencing, followed by subgenotyping. For the taxon identification, a nested PCR targeting a ~825 bp fragment of the small-subunit ribosomal DNA (18S SSU rDNA) was used. In addition, a single step PCR was applied to amplify a ~400 bp fragment of the *Cryptosporidium* 70kDa heat shock protein gene (HSP70). For subgenotyping, a ~850 bp fragment of the 60kDa glycoprotein gene (gp60) was amplified using a nested PCR. The PCR conditions were optimised in-house using the previously described primer sequences (Alves et al., 2003; Learmonth et al., 2004; Grinberg et al., 2008). Details of primers and PCR conditions used are shown in Table 3.3. Agarose gel electrophoresis was used to verify PCR products. Positive amplicons were purified using an ethanol precipitation protocol and submitted to a commercial sequencing provider for bidirectional Sanger sequencing, using the

same PCR primers. Consensus sequences were created by manual editing of forward and reverse sequences using Geneious 6.5 software (Biomatters, Auckland, New Zealand; http://www.geneious.com). The consensus sequences were aligned online with sequences deposited in GenBank (National Institutes of Health, Bethesda, MD, USA) using the Basic Local Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi, National Center for Biotechnology Information, Bethesda, MD, USA).

Initial results indicated that some specimens contained both *C. bovis* 18S SSU rDNA and *C. parvum* HSP70 and gp60 sequences. In order to rule out laboratory cross-contamination of specimens, the PCR-sequencing was repeated using re-extracted DNA with identical results. To corroborate the presence of mixed infections, further investigations were performed using different primers. New primers for a different region of the HSP70 gene (Table 4.1) were retrieved from the literature (Morgan et al., 2001) and were successfully mapped to the published *C. bovis* HSP70 gene sequence (AY741306.1). These primers were used in a new PCR (PCR2), in an attempt to amplify the *C. bovis* HSP70 from specimens yielding the *C. bovis* 18S SSU rDNA sequences. In addition, the *Cryptosporidium* actin gene was also amplified using primers (Table 4.1) shared by *C. parvum* and *C. bovis* (Sulaiman et al., 2002). Amplicons were run on gels, purified and sequenced as above. Whenever previously unreported sequences were identified, the PCR-sequencing was repeated using DNA re-extracted from the specimen and if the same sequence was obtained, the new sequence was deposited in GenBank.

Table 4.1. Primers and PCR conditions used for amplification of *Cryptosporidium* loci

	Locus	
	HSP70 PCR2	**Actin**
	Nested PCR	
Primer sequences (5'-3')	External	
	F':GGTGGTGGTACTTTTGATGTATC R':GCCTGAACCTTTGGAATACG	F':ATGRGWGAAGAGWARYWCAAGC R':AGAARCAYTTTCTGTGKACAAT
	Internal	
	F':GCTGSTGATACTCACTTGGGTGG R':CTCTTGTCCATACCAGCATCC	F':CAAGCWTTRGTTGTTGAYAA R':TTTCTGTGKACAATWSWTGG
PCR mixture in each 20uL reaction	External	
	1x PCR Buffer; 4mM dNTP; 1.75mM $MgCl_2$; 1U Platinum Taq; 4pmol of each F'/R'; 2µL DNA	Same to HSP70 PCR2 except: 3mM $MgCl_2$ 2.5U Platinum Taq
	Internal	
	Same to external except: 1.5mM of $MgCl_2$, DNA replaced with 1µl of external PCR product	Same to external except: DNA replaced with 1µl of external PCR product
PCR conditions	External	
	I: 5 mins at 94 °C D: 30 sec at 94 °C A: 30 sec at 57 °C E: 30 sec at 72 °C FE: 10 mins at 72 °C Cycle numbers: 40	I: 5 mins at 94 °C D: 45 sec at 94 °C A: 45 sec at 57 °C E: 60 sec at 72 °C FE: 10 mins at 72 °C Cycle numbers: 35
	Internal	
	Same to external	Same to external except annealing temperature of 45 °C
References	**Morgan et al., (2001)**	**Sulaiman et al., (2002)**

F': Forward Primer; R'': Reverse Primer

dNTP: deoxynucleotide triphosphate (Fermentas, Auckland, New Zealand);

BSA: non-acetylated bovine serum albumin (Invitrogen, Auckland, New Zealand);

Platinum Taq (Invitrogen, Auckland, New Zealand);

I: Initial heating; D: Denaturation; A: Annealing; E: Extension; FE: Final Extension

4.5 Results

4.5.1 Identification of *Cryptosporidium* oocysts

Round *Cryptosporidium* oocysts consistent in shape and size with intestinal species were detected in 15/60 faecal specimens by IFA. All the IFA-positive specimens originated from the first sampling event, in August. Conversely, all the specimens collected on the other three sampling occasions were IFA-negative for oocysts.

4.5.2 Identification of *Cryptosporidium* species and subgenotypes

Initial analysis of the 18S SSU rDNA, HSP70 and gp60 sequences of IFA positive isolates indicated a number of unusual results. A total of 7/15 amplicons yielded editable 18S SSU rDNA sequences, while 8/15 amplicons yielded unusable chromatograms. Of the seven readable amplicons, two displayed sequences identical to the *C. parvum* 18S SSU rDNA gene sequence (Genbank accession number AB746195.1), whereas three displayed sequences identical to the *C. bovis* gene (Genbank accession number AY741305.1). The remaining two 18S SSU rDNA amplicons displayed sequences 99% similar to the *C. parvum* 18S SSU rDNA gene sequence, but were previously unreported in Genbank or in New Zealand. These new 18S SSU rDNA variants differed from the *C. parvum* 18S SSU rDNA sequence of strain Iowa (Genbank accession number AF164102.1) in the region between nucleotide 634 and 695. Each new 18S SSU rDNA variant was confirmed by two additional rounds of sequencing of new amplicons derived from re-extracted DNA, and deposited in Genbank (Accession number KF840580 and KF840581).

Thirteen out of 15 HSP70 amplicons displayed editable sequences in the first PCR reaction (HSP70 PCR 1, Table 4.2). All these were 100% identical to the *C. parvum* HSP70 gene sequence in the conserved region (Genbank Accession number U11761.1), but were of three different types, varying in the number and positioning of the 12 bp repeat units in the variable region of the gene. One HSP70 sequence type was found in five isolates, and contained 12 repeated units (Type 1, Figure 4.1). The other types were found in 8 isolates and contained 11 repeat units (Type 2, Figure 4.1). There was also variation in the position of the different repeat units in Type 1 (Type 1a, Figure 4.1), defining a total of three HSP70 alleles among 13 isolates.

Re-analysis of the three specimens that yielded the *C. bovis* 18S SSU rDNA using different HSP70 primers (PCR 2) confirmed the presence of the *C. parvum* sequence in two specimens, and a sequence 100% identical to the *C. andersoni* HSP70 gene (Genbank accession number JQ031809.1) in one specimen. Conversely, the actin locus indicated the presence of a *C. bovis* sequence (accession number AY741307) in the three isolates (Table 4.2).

Finally, 12/15 isolates amplified the gp60 gene. All the gp60 sequences belonged to the IIaA19G4R1 (n=8) and IIaA18G3R1 (n=4) subtypes (Table 4.2). All the 18S SSU rDNA, HSP70, gp60 gene sequences obtained were shown in Appendix D.

Table 4.2. *Cryptosporidium* taxa and subgenotypes identified in cattle faecal specimens by sequence analysis of PCR products. In the same column, identical sequences are denoted by superscripts

Specimens/Isolates	Locus				
	18S SSU rDNA	HSP70 PCR1	gp60 subgenotype	HSP70 PCR2	Actin gene
1	UD	*C. parvum*[g]	IIaA18G3R1	NA	NA
2	UD	*C. parvum*[f]	IIaA19G4R1	NA	NA
3	UD	*C. parvum*[g]	IIaA18G3R1	NA	NA
4	UD	*C. parvum*[f]	UD	NA	NA
5	*C. bovis*[a]	UD	IIaA19G4R1	*C. parvum*[b]	*C. bovis*[a]
6	*C. bovis*[a]	*C. parvum*[g]	IIaA18G3R1	*C. andersoni*[h]	*C. bovis*[a]
7	UD	*C. parvum*[f]	IIaA18G3R1	Not applicable	NA
8	UD	*C. parvum*[e]	UD	NA	NA
9	New variant (99% identical to *C. parvum*)[c]	*C. parvum*[g]	IIaA19G4R1	NA	NA
10	UD	*C. parvum*[g]	IIaA19G4R1	NA	NA
11	*C. parvum*[b]	*C. parvum*[g]	IIaA19G4R1	NA	NA
12	New variant (99% identical to *C. parvum*)[d]	*C. parvum*[g]	UD	NA	NA
13	*C. bovis*[a]	*C. parvum*[f]	IIaA19G4R1	*C. parvum*[b]	*C. bovis*[a]
14	UD	*C. parvum*[g]	IIaA19G4R1	NA	NA
15	*C. parvum*[b]	UD	IIaA19G4R1	NA	NA

[a] sequence 100% identical to *C. bovis* (Accession No: AY120911.1 and AY741305);
[b] sequence 100% identical to *C. parvum* (Accession No: DQ010952)
[c] GenBank accession number KF840580; [d] GenBank accession number KF840581
[e] HSP70 Type 1 *C. parvum* (see Figure 1); [f] HSP70 Type 1a *C. parvum* (see Figure 1);
[g] HSP70 Type 2 *C. parvum* (see Figure 1)
[h] sequence 100% identical to *C. andersoni* (Accession No: JQ031809.1)
PCR1: PCR performed with the first set of primers (see text);
PCR2: PCR performed with the second set of primers (see text);
UD = Taxon undetermined due to unusable chromatograms;
NA = Not applicable

Figure 4.1: Variable repeat region of HSP70 (PCR 1) allele found in this study. Type 1 and 1a contain 12 repeat units of 12 bps each, whereas type 2 has only 11 repeat units. Each colour represents a 12-bp repeat unit type. Note also the difference in the position of the repeat unit types between type 1 and 1a.

4.6 Discussion

Using iterative multilocus genotyping, we demonstrate the presence of DNA sequences of the three most common *Cryptosporidium* species infecting cattle on a single dairy farm. We also provide evidence for the presence of *C. parvum* + *C. bovis* in three specimens (5, 6, and 13; Table 4.2), *C. parvum + C. bovis* + *C. andersoni* in one specimen, and *C. parvum* + new genetic variants in two specimens (9 and 12). To limit the possibility of cross contamination between the faecal pats, only specimens found at least two metres apart were collected. Thus, the results most likely represent mixed infections.

The isolates were genotyped at the 18S SSU rDNA and HSP70 genes, which are taxonomically informative loci (Caccio et al., 2005) (the taxonomic value of the actin locus is less understood). These loci are assumed to co-segregate during reductional division, so we assume that the combinations of sequences reflect mixed infections rather than recombinant taxa (this second possibility should be assessed by single-sporozoite analysis, which may be feasible in the future). In addition, we report two new *C. parvum*-like 18S SSU rDNA sequences 99% identical to the *C. parvum* sequence. Sequencing or editing artefacts were ruled out by repeating the PCRs on re-extracted DNA, with consistent results. These new variants add to the long list of 18S SSU rDNA sequences available in the GenBank, and could represent heterogeneous 18S SSU rDNA copies of *C. andersoni* or *C. bovis* (the heterogeneous copy of the *C. parvum* 18S SSU rDNA gene is known). In this study, when mixed infections were suspected due to discordant sequencing results at two loci, we performed additional analyses and obtained a more accurate picture of the genetic makeup of the isolates. Such an iterative approach is not always used in population-based studies.

We analysed only IFA-positive specimens containing cycling parasites, rather than naked DNA. All the IFA-positive specimens were firm in consistency, and were obtained on the first sampling occasion when the calves were between 5 and 6 weeks of age. It is possible that on the subsequent occasions the calves were not shedding, or were shedding low numbers of oocysts (Davies et al., 2003). *C. parvum* is commonly identified in calves during the first month of life, whereas *C. bovis* is more often reported in older, weaned calves and *C. andersoni* in juvenile and adult cattle (Fayer et al., 2006; Santin et al., 2004; 2008; Xiao et al., 2007). The presence of *C. parvum* and *C. bovis* in different animals on the same farm has been previously reported (Amer et al., 2013), and in calves, mixed infections with two taxa have been recently suspected from

the results of PCR-RFLP and subsequently confirmed by sequencing (Rzeżutka and Kaupke, 2013). To our knowledge, the confirmation of endemic mixed infections with the three major *Cryptosporidium* species of cattle in one specimen is novel. As the infections occurred in calves about six-weeks of age, these calves could have been at the transition period, when infections with *C. parvum, C. bovis* and *C. andersoni* were patent. According to the farmer, juvenile or adult cattle had never populated the paddock, so cross-contamination of the pat with manure from adult cattle was unlikely.

Only *C. parvum* gp60 sequences were identified in this study. This may be due to polymorphisms in the primer sites between the species (the sequence of the gp60 genes of *C. bovis* and *C. andersoni* are not well characterised). All the isolates carried potentially zoonotic IIaA19G4R1 and IIaA18G3R1 alleles. The use of cloning or next generation sequencing would have allowed identification of both alleles in individual calves, as previously described for humans (Grinberg et al., 2013).

Finally, this is the first report of the presence of *C. bovis* and *C. andersoni* in New Zealand, although a sequence 99% similar to the *C. bovis* 18S SSU rDNA has previously been reported (Grinberg et al., 2011).

4.7 Conclusion

In summary, we report endemic mixed infections in calves at the transition age of six weeks, with the three main *Cryptosporidium* species of cattle and new genetic variants. Mixed infections may be more common than thought, and could easily be overlooked in population-based studies where ambiguous genetic profiles are eliminated from further analysis. These results add to the body of knowledge on the intra-host genetic diversity of *Cryptosporidium* isolates, and highlight the usefulness of iterative analyses for the description of the genetic makeup of the isolates. The clinical impact of mixed infections remains to be established.

4.8 Highlights of this study

- Iterative PCR analysis targeting two or more loci is important to consider in determining the *Cryptosporidium* species.
- Further analysis of mixed infection should be considered as it may help in determining the clinical impact of mixed infections.

4.9 Acknowledgements

I would like to express my sincere thanks to my supervisor Dr Alex Grinberg for providing valuable suggestion and helping to generate an idea for this chapter.

This project was funded by Ministry of Health Fund for Aquatic Protozoa Research Unit.

Chapter 5

Molecular characterisation of *Campylobacter jejuni and Campylobacter coli* isolates from faecal and water samples using next-generation sequencing technology: Introduction of a novel method "massMLST"

5.1 Preamble

In the previous chapter (4), *Campylobacter jejuni* and *Campylobacter coli* were reported in the faecal and water samples collected from the two catchment areas in the Manawatu-Wanganui region of New Zealand. Zoonotic potential and genetic diversity of *Campylobacter* spp. identified in farms and river water within the catchment areas can be determined through subtyping *Campylobacter* isolates. Multi-Locus Sequence Typing (MLST) that uses the Sanger sequencing method is a widely used method for *Campylobacter* spp. subtyping, although it is costly and time-consuming to use this method to sequence and subtype a huge number of isolates. Therefore, this chapter presents a new method, "massMLST" to determine the multi-locus sequence typing of *Campylobacter* species using a next-generation sequencing technology platform.

5.2 Abstract

Next generation sequencing (NGS) technologies have grown rapidly in recent years. This technology has substantially improved the field of biology including molecular

biology and medicine. Multi-locus sequence typing (MLST) is a gold standard test for typing *Campylobacter.* However, this method incurs sizeable costs to analyse a large numbers of isolates. Therefore, MLST in conjunction with NGS analysis could be a promising approach for cost-effective and high-resolution molecular typing of large sets of *Campylobacter* strains. In this study, we describe a novel method "massMLST" that uses two steps: PCR-based target MLST allele amplification followed by sequencing on an Illumina MiSeq machine. To test the feasibility of the massMLST approach, two fragments of seven *Campylobacter* housekeeping gene sequences were amplified and simultaneously sequenced with barcode tagging for each of 96 *Campylobacter* species isolates. The amplicons were then purified, quantified, normalised, pooled into a single tube, and analysed by Illumina MiSeq sequencing. Analysis of the sequence data, as an extension of multilocus sequence typing (MLST), demonstrated the utility and potential of massMLST as a strain typing method. The massMLST also identified potentially zoonotic *C. jejuni/coli* strains in ruminant faeces from farm environments within the two catchment areas. Automation of this massMLST method together with the increasing sequencing capacity of MiSeq machines would enable this massMLST method to become a cost-effective and high-resolution typing method for multi-locus sequence typing of large collections of *Campylobacter* strains in the near future.

5.3 Introduction

Across the world, *Campylobacter* spp. are the principal bacterial cause of human gastrointestinal infectious diseases (WHO, 2011). In New Zealand, campylobacteriosis is the leading notifiable enteric disease, with reported illness rates of 158.6 cases per 100,000 for 2012 despite a significant decrease in the number of campylobacteriosis cases since the peak in 2006 (from 15,873 to 7,031 cases) (ESR, 2013). Moreover, 2.3% of the 10,491 campylobacteriosis cases reported in 2012 in New Zealand were associated with waterborne (190/10,491) and foodborne (51/10,491) outbreaks. Although these cases were related to the outbreaks it is believed to be sporadic infections (ESR, 2013). Approximately 90% of the campylobacteriosis cases were reported to be associated with *Campylobacter jejuni*, <10% with *Campylobacter coli*, and the remaining with other *Campylobacter* species (Nachamkin et al., 2008).

Poultry are a well-known reservoir of *C. jejuni* and *C. coli* and have been considered the most common source of infection for human cases of campylobacteriosis (Eberhart-Phillips et al., 1997; Moore et al., 2005; Mullner et al., 2009). However,

evidence also exists for the roles of ruminants and contaminated water as sources of *Campylobacter* in human infections (Grove-White et al., 2010; French & Marshall., 2012; Wilson et al., 2008; this thesis). Both *C. Jejuni* and *C. coli* are considered commensal organisms in cattle and sheep, with *C. coli* being isolated more frequently from sheep than cattle (Lazou et al., 2013). Additionally, cattle and sheep are also considered the primary source of water contamination (Chatre et al., 2009; Ramonaite et al., 2013). Therefore, both direct contact with animal faeces and consumption of contaminated food and water are implicated in the total burden of human campylobacteriosis. Yet, defined *Campylobacter* infection routes have yet to clarified because laboratory investigations are limited to outbreak situations due to the limited value of collecting and analysing prospective clinical samples. Consequently, efforts to track sources of *Campylobacter* infection have been inadequate.

Molecular typing has become an essential tool for disease management to understand the diversity and population genetics of microbial pathogens. A number of different molecular subtyping methods, such as pulsed-field gel electrophoresis (PFGE), restriction fragment length polymorphism (RFLP), and DNA sequencing of the flagellin gene short variable region (flaA SVR), have been used to identify strains of *Campylobacter* (Klena & Konkel, 2005). PFGE has previously been considered "the gold standard" for subtyping *Campylobacter* because of its high discriminating power (Klena & Konkel, 2005; Fitzgerald et al., 2001). Although PFGE can track the sources of infection, this method has pitfalls in identifying the degree of relatedness among similar isolates (Fugget et al., 2007; Barret et al., 2007). Due to this shortcoming, the PFGE technique has been superceded by a more recently developed robust method, multilocus sequence typing (MLST).

MLST has emerged as the leading method for subtyping *Campylobacter* species and is based upon sequencing seven housekeeping genes (Dingle et al., 2001). MLST utilises similar concepts and processes as multilocus enzyme electrophoresis (Sealander et al., 1986) by indexing seven chromosomal locations, but MLST determines nucleotide sequence to identify neutral genetic variations directly rather than indirectly through electrophoretic mobility of their gene products (Dingle et al., 2001). The allelic profiles of these seven genes are allotted to define a sequence type (ST) and thereby, in many cases, a clonal complex (CC) of each isolate.

MLST analysis of *Campylobacter* spp. has shown that the species are genetically diverse and have a frequent intra- and inter-species horizontal genetic exchange indicating a weakly clonal population structure, (Dingle et al., 2001; 2005; Sheppard et al., 2005). Some MLST studies of *Campylobacter* spp. have shown *C. jejuni* lineages to be linked to restricted geographical areas or to particular ecological niches, such as wild birds, cattle, or sheep (Mullner et al., 2006; McTavish et al., 2008; French et al., 2009; 2011; Carter et al., 2009). The high discriminatory power, reproducibility, and possibility of interlaboratory comparisons of MLST data have contributed to a better understanding of global epidemiology and the population structure of *Campylobacter* (Dingle et al., 2001; French et al., 2005; Maiden and Dingle, 2008). Yet, MLST is complex and expensive to perform, and there can also be significant genomic differences between isolates that are indistinguishable by MLST. Consequently, genotyping methods with enhanced discrimination are continually required for shedding further light on molecular epidemiologic investigations of *Campylobacter* spp.

MLST provides informative sequence data, unambiguous sequences, and allelic profiles for each isolate using PCR amplification of a single *Campylobacter* colony followed by DNA sequencing using the Sanger method (Dingle et al., 2001). The Sanger sequencing (the chain-termination method) was the first method of nucleic acid sequencing, and was considered the gold standard for more than 25 years (Sanger et al., 1977). Although this Sanger method has also been used widely for MLST of *Campylobacter* in many countries, it has limitations relating to its throughput, speed, scalability, resolution, and expense (Dingle et al., 2001; Haas et al., 2011). With the advent of the technology and demand for cheaper and faster sequencing methods, second-generation or next-generation sequencing (NGS) have been developed that have overcome many of the limitations of the Sanger sequencing method.

NGS has widely been used in the molecular epidemiology of infectious diseases in the past 7-8 years on the strength of several key features such as its ability to sequence long DNA base pairs across entire genomes and to run multiple isolates in one lane flow cells. Additionally, NGS has a high depth coverage[1], an output of up to 24-30 million reads within a 40-hour run, and a low sequencing cost (USD 0.05 to 0.15 per one million bases) (Metzker, 2010, Liu et al., 2013). As a result, researchers

[1] Amount of contiguous length of nucleotide bases generated by a sequencing machine that at least covered 90% of all sample bases once.

are increasingly employing this technology to further their understanding of infectious disease epidemiology. For example, a large food-borne disease outbreak caused by *E. coli* O104:H4 occurred in Germany in May 2011. During this outbreak, there were 36 deaths, and 336 people suffered from severe disease with bloody diarrhoea and haemolytic uremic syndrome (Brzuszkiewicz et al., 2011). In the early stages of this outbreak, prospective whole genome sequencing was employed using NGS technology; this helped scientists and medical personnel to respond faster in determining the underlying causes and helped preventing the further cases (Mellmann et al., 2011). Overall, this technology has provided a better understanding of animal, human and plant diseases in order to prompt alerts and controls, and it has increased in importance as a tool in the fields of molecular and genomic epidemiology and can also be used as a clinical diagnostic tool (Brzuszkiewicz et al., 2011, Mellmann et al., 2011, Sikkema-Raddatz et al., 2013)

Recently, multi-locus sequence typing coupled with NGS has been employed in a high-throughput MLST (HiMLST) approach to generate large-scale sequence data on the 454 Roche Genome sequence machine (Boers et al., 2012). This approach used sequence-specific primers to amplify segments of genes from four bacterial species in a two step-PCR, pooled the bar-coded samples, and sequenced them to generate 400-500 bp long reads. Although this method has been used successfully for some pathogens, such as *Pseudomonas,* it has not been optimised for *Campylobacter* species. During this PhD project, there was a limited budget for the subtyping of *C. jejuni* and *C. coli* isolates obtained from the catchment study in Chapter 3. At the same time, I was also interested to know the subtypes of *C. jejuni* and *C. coli* circulating in the two catchment environments. Therefore, I aimed to use NGS with MLST to investigate *C. jejuni* and *C. coli* diversity within samples collected from cattle, sheep and water sources by simultaneously sequencing amplicons generated from PCR of seven MLST genes from 96 isolates on a single run of an Illumina MiSeq sequencing machine, and performing bioinformatics analysis of the reads produced. For the purpose, this combination of NGS with MLST for the mass-scale study was named the "mass multi-locus sequence typing (massMLST)" method. Demonstrating the feasibility of the massMLST method, could help to validate an alternative high throughput typing method. If successful, this method would provide a more cost-effective approach to typing of large numbers of samples simultaneously than running conventional MLST protocols on individual isolates.

5.4 Materials and methods

Of the 290 isolates described in chapter 3, 90 were randomly selected in a way that allowed for equal representation of *Campylobacter* isolates that had originated from beef cattle, dairy cattle, and sheep. Of the remaining six isolates, four had originated from water samples and two were control strains (Table 5.1). The two control isolates used were *C. jejuni* ST-45 from a poultry source and ST-63 from an isolate derived from a clinical human case in New Zealand. The 90 isolates were collected from five farms in the Dannevirke area (68/90) and two farms in the Shannon (22/90) area. Details about each isolate used in this study are shown in Table E.1 in the Appendix.

Table 5.1. Origin and number of *C. jejuni* and *C. coli* isolates used in this study.

Origin of isolates	No. of isolates		
	C. jejuni	*C. coli*	Total
Beef cattle	23	2	25
Dairy cattle	20	5	25
Sheep	21	19	40
River water	4	0	4
Total	**68**	**26**	**94**

5.4.1 Sequencing concept and massMLST primer design

The current MLST scheme for *C. jejuni* and *C. coli* amplifies ~400-500 bp internal fragments of each of seven housekeeping genes via PCR and subsequent sequence determination using capillary sequencers. However, the idea in this study was to sequence the "traditional" MLST gene sequences for massMLST using an Illumina® MiSeq NGS machine. However, at the time of this study, the MiSeq system could only sequence two reads of up to 250 bp long, which is approximately half of the *C. jejuni/coli* MLST gene fragment used in MLST analysis. Therefore, each targeted MLST gene had to be split into two sequences for the MiSeq system. To obtain two fragments, two sets of primers for each MLST gene were redesigned from a set of published PCR primers such that they would amplify and produce overlapping PCR products (Miller et al., 2005). This amplification allowed the determination of any variation outside the "traditional" MLST genes and allowed assessment of SNPs or any new STs or diversity within the sample. When re-designing MLST genes primers (denoted as a "massMLST" primer), the primers

were further degenerated as little as possible by reviewing the proposed primer locations against a set of *Campylobacter* genomes from our laboratory. For each massMLST primer, unique Illumina tags were also attached to differentiate sequences for producing overlapping PCR products (Figure 5.1). Details about the original Miller et al. primers and the massMLST primers for each housekeeping gene are shown in Table 5.2.

Figure 5.1 illustrates the concept of massMLST for one of the seven MLST genes, *gln*A. For each redesigned primer, Illumina transposon adapter sequences were added to the 5' ends of each primer, in order to make the primer compatible with adding the isolate-specific barcode during a second round of PCR (PCR2). The first set of the PCR products of each gene was termed fragment one (F1), and second fragment two (F2). All the massMLST primers were ordered from Invitrogen™ (Life Technology Corporation, USA). Once the massMLST primers were obtained, the primers were diluted with ultra-filtered water to prepare stock-primer solutions of 1 nmol/µL concentration. From each stock solution, working-primer solutions containing 2 pmol/µL were prepared using molecular biology grade water. Both the stock and working solutions were stored at -20 ^{o}C for future work such as PCR optimisation and testing.

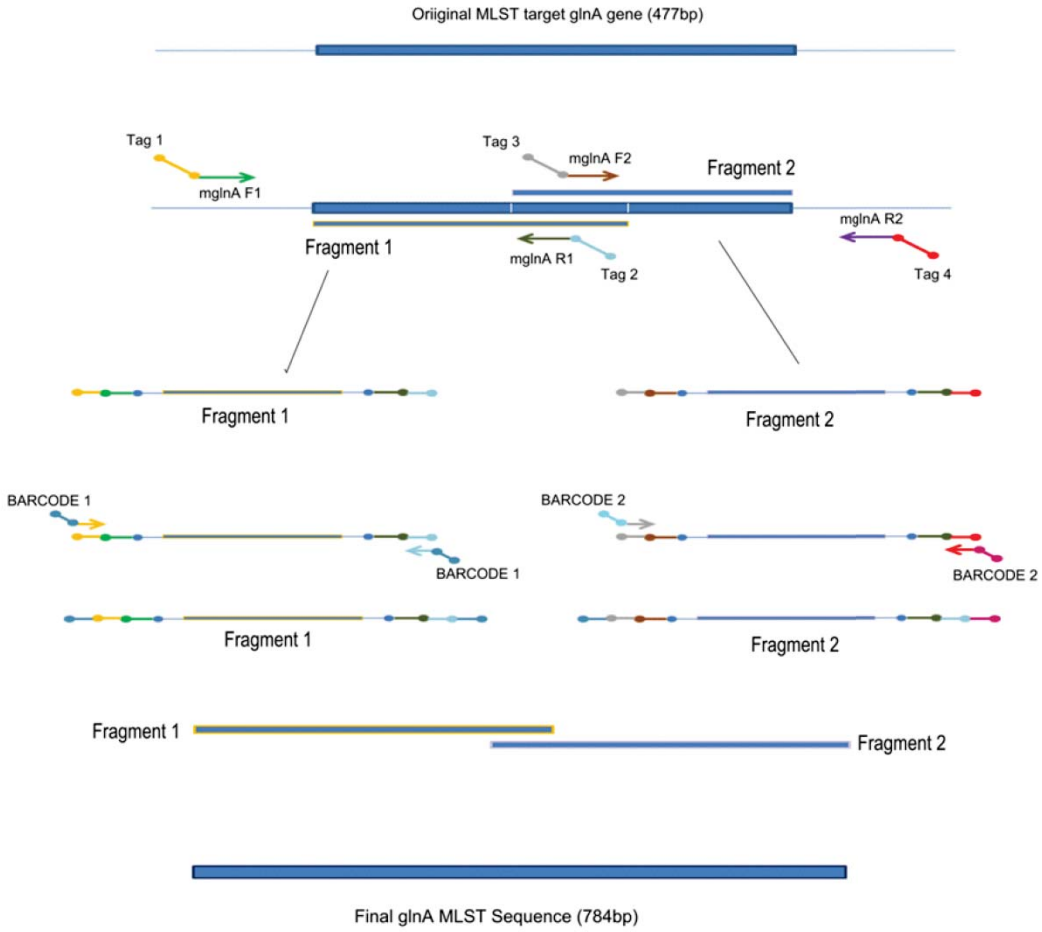


Figure 5.1: The concept behind the re-designed Multi-Locus Sequence Typing (MLST) primers and obtaining the massMLST sequences. The example shown is one of the seven housekeeping genes of Campylobacter, the glnA allele. The dark blue thick line is the length of target gene sequence for the original Miller et al. primers. This sequence is divided into two parts such that the middle sequences partly overlap for each fragment. Arrows in green and dark green colours are the forward and reverse primers for fragment one, and the brown and purple arrows are the forward and reverse primers for fragment two. Angular lines with dots at both ends (yellow, blue, red and grey) are unique Illumina tags attached to the primers. When PCR was run with these primers and tags, fragment one and fragment two sequences were produced. Barcodes (dark blue and pink) were added onto each side of each sequence in the second round of PCR. The final product that acts as the forward sequence is fragment one (green), and that acts as the reverse sequence is fragment two (blue) for that gene. Finally, these sequences were aligned to produce full-length sequences for that gene fragment.

Table 5.2: The primer sequences for seven housekeeping genes for use in massMLST of *Campylobacter* isolates. These primers were modified from the conventional *Campylobacter* MLST primers used to sequence seven housekeeping genes. Each gene's primer sequences contain two fragments along with transposome adapters at the 5' end. The table also shows melting temperature (Tm), sequence size and PCR product size for each locus.

locus	Forward (f)		Reverse (r)		Sequence length	locus spec. Tm	Size with transposome adapters	PCR product size	New primer direction	Start	End	After PCR2
	name	Sequence	name	sequence								
aspA	mAspAF1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGARAGAAAAGCWSAWGAATTTAAAGAT	mAspAR1	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCCYGGCATRATRGARCTDCCYGRTTGC	f-60	51.0=51.7=52.5	455	522	rev	942	968	591
					r-61	57.2=63.8=70.8						
	mAspAF2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGDYCRAAAATGCAAYCRGGHAGYTCY	mAspAR2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTTYTTCATTWGCRSTRATRCCATC	f-58	53.3=59.6=66.9	257	324	for	933	957	393
					r-59	49.9=54.2=58.6						
atpA | uncA	mAtpAF1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGHCAAGGDGTTRTYTGTATHTATGTWGC	mAtpAR1	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGDCKRCHYGRDGRRMGRYGYAAAATYAR	f-61	51.8=55.8=61.0	287	354	rev	823	849	423
					r-61	49.3=61.1=74.9						
	mAtpAF2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGYCKYYCHYCRDGYMGHGARGCTTAYCC	mAtpAR2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTTAADAVYTCWACCATTCTTTGHCC	f-60	54.3=64.6=77.3	429	496	for	834	860	565
					r-60	49.8=53.5=57.8						
glnA	mGlnAF1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGATAGGMACTTGGCAYCATATBAC	mGlnAR1	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTYRTCRYTCCAYYCBCCYTCYTCDGYR	f-58	52.7=56.1=58.9	415	482	rev	499	525	551
					r-61	53.7=64.0=74.1						
	mGlnAF2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGDRAYRCHGARGARGGVGRRTGGARYGA	mGlnAR2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGARRCTCATATHMACATGCATDCCR	f-60	53.7=63.9=74.2	369	436	for	495	521	505
					r-58	50.6=55.4=61.6						
gltA	mGltAF1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGARTGGCTTGCHGAAAAYAARCTTT	mGltAR1a	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGYRAADMCVCDATCYAARTTDGGATAAR	f-58	53.1=56.5=60.7	362	429	rev	603	626	498
					r-62	47.6=56.1=64.7						
	mGltAF2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTCCNAAYTTRGATHGBGKHTTYRC	mGltAR2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTATAAACCCTATGYCCAAARCCCAT	57	45.6=56.0=65.8	368	435	for	603	626	504
					r-59	54.1=56.2=58.4						

glyA	mGlyAF1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGATTCWGGTTCTCAAGCWAATCAAGG	mGlyAR1	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGRVRAAWGVRYYTRDRTGYYCVSHYGCH	f-58	55.0=55.4=55.8	367	434	rev	612	638	503
					r-61	50.4=63.1=76.5						
	mGlyAF2b	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGRYTGTDGCRDSNGRRCAYHYARRY	mGlyAR2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGCYAAATCHGCATCTTTKCCRCTAAA	f-57	48.3=60.6=75.2	375	442	for	612	638	511
					r-60	54.0=57.9=62.3						
pgm | Cj0360	mPgmF1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCWTTRCGYGTDGTTTTAGATGTVGC	mPgmR1	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGYTTBYYRARYACRTAYTTRTCNCC	f-58	51.7=56.7=61.2	422	489	rev	922	945	558
					r-58	43.3=53.8=64.2						
	mPgmF2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGNGAYAARTAYGTRYTYRRVAAR	mPgmR2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAATTTTCHGTBCCWGAATAGCGRAA	f-57	43.3=53.8=64.2	322	389	for	922	945	458
					r-59	53.4=56.1=59.2						
tkt	mTktF1	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGCWAAYTCRGGHCAYCCDGGTGC	mTktR1	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGYYGCYATDGYRAARCYYRYVGCR	f-57	57.9=63.0=69.5	320	387	rev	366	389	456
					r-58	48.6=62.3=75.3						
	mTktF2	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGYGCBRYRRGYTTYRCHATRGCRRC	mTktR2	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTTTTAAYVAVHTCTTCRCCCAAAGGT	f-57	48.6=62.3=75.3	434	501	for	366	389	570
					r-60	51.7=56.5=60.9						

5.4.2 PCR optimisation

PCR reactions, occasionally even established ones, require optimisation to enable better amplification of that particular DNA segment by avoiding the generation of unwanted and undefined PCR products. During PCR, parameters that play a vital role in producing better amplification are the amount and quality of DNA template, primers, magnesium concentration, deoxynucleotide triphosphates (dNTPs), DNA polymerase, buffer, denaturation temperature and duration, annealing temperature and duration, extension time and cycle number.

i. DNA template

The already identified and purified *C. jejuni* and/or *C. coli* were used as a DNA template at 20 ng/μL concentration. Varying amounts of the template starting from 10 ng to 80 ng were used during the optimisation process.

ii. Primers

The prepared primer working solutions (2 pmol/μL) were used for optimisation by varying the primer amounts from 1 pmol to 8 pmol.

iii. Magnesium concentration

Magnesium chloride (a stock solution of 50 mM) of concentrations 1.00, 1.50, 1.75, 2.00, 2.25, 2.50 and 3.00 mM were used to find the optimal concentration to be used for PCR.

iv. deoxynucleotide triphosphates (dNTPs)

dNTPs at a starting concentration of 2.00 mM/μL were used for optimisation by preparing different concentrations of 2.00, 3.00, 4.00, 5.00, and 6.00 mM.

v. DNA polymerase concentration

For optimisation, 0.5 – 2.0 units of Taq DNA polymerase were used for finding the optimal concentration.

vi. Denaturation temperature and duration

Initially, a preset programme for denaturation temperature and duration to perform PCR for MLST in the laboratory was employed for this study and variations were then made on this standard programme to determine the optimal temperature and period.

vii. Annealing temperature and duration

The re-designed primers of Miller et al. were 28 - 33 bp long, but the addition of the 33 or 34 long adapter bases onto the primer increased their resulting length to 61 - 67 bases. This led to a higher melting temperature (T_m). Therefore, the annealing temperature was tested as T_m ±10 ^oC with duration of 15 - 30 seconds.

viii. Extension time

The tested extension time was between 30 to 60 seconds at 68 ^oC and 72 ^oC.

ix. Cycle number

In this study, the PCR cycle conditions included one cycle of initial denaturation and one cycle of final extension, and 35 to 40 cycles of denaturation, annealing and extension for each fragment of the gene.

Therefore, in this study, the combination of all the stated parameters with different concentrations and conditions for 14 sets of primers were optimised before PCR amplification of the isolates. Primers were optimised using four known *Campylobacter* isolates. These isolates originated from poultry, ruminant, human and water sources. Any primers, either reverse or forward, that did not work from each set were tested using cross matching to verify which primer was not working. For example, pgmF1 and pgmR1 are a set of forward and reverse primers for fragment one of the pgm gene, whereas, pgmF2 and pgmR2 are another set of primers for the second fragment of the pgm gene. Both sets were optimised but the first set did not amplify and produce a band in a gel photograph. Therefore, optimisation verification was assessed by the amplification of four control samples using cross-matched primers, the forward primer of fragment one (pgmF1) with a reverse primer of original MLST (pgmR) and vice-versa (Figure 5.2). The amplified products were visualised in a gel photograph, and the problematic primer was identified. Finally, problematic primers were modified by changing their sequence slightly (again in light of the available genome sequences), the variant primers were re-ordered and re-optimised accordingly.

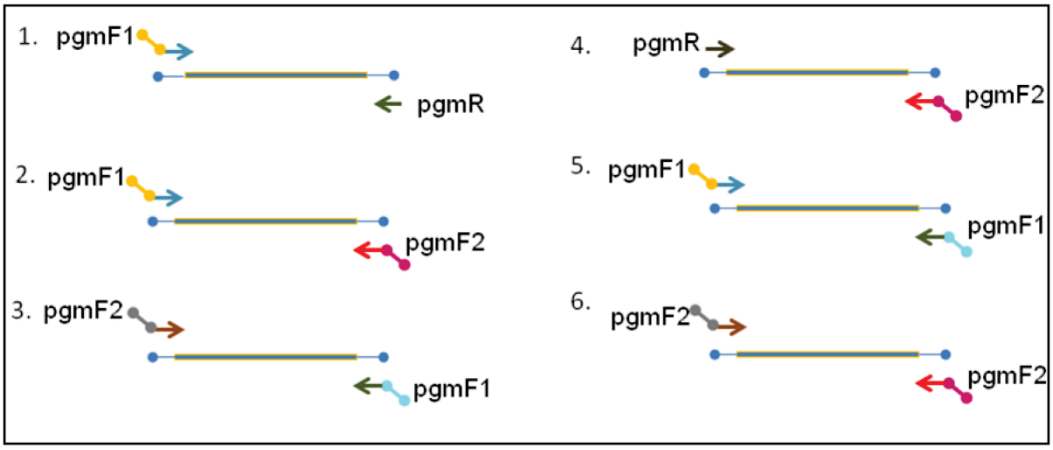


Figure 5.2: Exemplification of the method used to identify unamplified or problematic primers by using different forward and reverse primers combinations (1 - 6), including the original MLST primer, for a new amplification.

5.4.3 Amplification, normalisation, purification and NGS library preparation

A schematic diagram in Figures 5.3 and 5.4 depicts the various steps involved to amplify, purify, and normalise the samples and amplicons for NGS library preparation. The genomic DNA samples were quantified and normalised before amplification.

5.4.3.1 Quantification of genomic DNA

Quantification of DNA used in this study was determined by absorption fluorometry, using a Quant-iT™ PicoGreen® dsDNA assay (Life Technologies Corporation) on a microplate reader (Perkin Elmer Wallace 1420 VICTOR2™) according to the manufacturer's instructions. Each quantified *Campylobacter* DNA sample was normalised by equal molarity to produce the concentration of 20 ng/µL per sample that was used for further sequencing preparation.

5.4.3.2 Illumina library preparation

Two steps of PCR of samples were performed for sequencing library preparation of *Campylobacter* MLST gene amplicons. These amplicons were then used for library preparation using an Illumina MiSeq Sequencer.

5.4.3.2.1 PCR amplification of Campylobacter DNA (PCR 1)

Amplifications of all the isolates were performed in a 20 µL reaction mixture using 96 well plates. The PCR reaction mixes were made up as master mixes containing PCR buffer, Mg^{2+}, dNTPs, Taq polymerase enzyme and water, as listed in Table 5.3. The massMLST primers were added to each PCR reaction separately before adding the DNA, and the final PCR mixture was amplified in a SensoQuest Labcycler Thermocycler using the PCR conditions stated in Table 5.4. In this first PCR run (PCR1), the targeted flank region of 96 isolates were amplified per plate per gene (Figure 5.3). To ensure that all the isolates in a plate were amplified and amplicons were of the expected size, 2 µL of each of the PCR amplicons were run on a 1% agarose gel by electrophoresis at 110 V for 1 hour in 0.5% TBE buffer. Gels were visualised under UV light to determine the presence and size of PCR amplicons. Multiple amplification rounds (at least four) were used for samples that did not amplify well. All positive amplicons from the 14 PCR plate runs were purified, quantified and normalised before submission to the Massey Genome Service (MGS) for the second PCR run.

5.4.3.2.2 Amplicons purification

All PCR-positive amplicons obtained from PCR1 were purified using Agencourt AMPure XP magnetic beads (Beckman Coulter, USA) following the manufacturer's instructions. Briefly, 15 µL of each PCR amplicon was transferred into a new PCR plate with 27 µL AMPure XP beads, gently mixed using pipetting, and then incubated at room temperature for five minutes to bind the amplicons to the paramagnetic beads. The PCR plate was then placed into an Agencourt SPRIPlate Super Magnet Plate (Beckman Coulter, Indiana, USA) for five minutes. The magnetic plate separated the beads and PCR contaminants from solution. Once the solution became clear, the supernatant was carefully removed and discarded while the PCR plate remained on the magnetic plate. Thereafter, the beads were washed twice by adding 200 µL of freshly prepared 70% ethanol into each well, incubating them at room temperature for 30-40 seconds and discarding the alcohol carefully. The plate was left at room temperature to dry for up to five minutes to ensure all traces of ethanol were removed. The PCR plate was removed from the magnet plate and 40 µL of molecular grade water was added to each well of the plate, mixed 10 times with a pipette and incubated for one minute at room temperature. The PCR plate was placed back onto the magnetic plate for another minute to separate the beads from the solution, and the eluted solution was transferred into a new PCR plate. The

purified amplicons were quantified again using PicoGreen and stored at -20 ^{o}C until further processing.

5.4.3.2.3 Quantification and normalisation of purified amplicons

The purified amplicons were quantified using Quant-iT™ PicoGreen® dsDNA[2] (Life Technologies Corporation) on a microplate reader (Perkin Elmer Wallace 1420 VICTOR2™, Massachusetts, USA), following the manufacturer's directions. Briefly, DNA stock solution was diluted in TE buffer, and 1 mL PicoGreen® reagent was added to the diluted solution to prepare two standard solutions. Then, the solutions were incubated for 3-5 minutes at room temperature, before the fluorescence was measured in a fluorescence microplate reader at excitation and emission wavelengths of ~480 and 520 nm, respectively. The fluorescence values obtained were subtracted from the value of the reagent blank to determine a low (25 pg/mL to 25 ng/mL) and a high (1 ng/mL to 1 µg/mL) range standard curve. Fluorescence values for samples were also obtained similar to the standard solution procedure. Finally, the DNA concentration of the sample was determined from the standard curve generated in DNA Standard Curve.

After quantification, each amplicon was diluted with water to a concentration of 1 ng in a volume of 25 µL. The normalised amplicons in the 14 plates were then submitted to MGS for PCR2. Similar to the procedure used for PCR1, the amplicon product was run on an agarose gel as per the conditions above. Thereafter, the quality and/or quantity of each of the 96 amplicons from PCR2 was checked using a PicoGreen® on a plate reader and the amplicons normalised to a concentration of 1 ng in a volume of 25 µL to obtain an equal amount of DNA for each isolate in the library.

5.4.3.2.4 PCR amplification of amplicons from PCR1 (PCR2)

The second run of PCR (PCR2) was performed by the MGS to add the Illumina NexteraXT indexed adapters (synthetic oligonucleotides of known sequence) to each end of the amplicons for library preparation (Figure 5.4). The Illumina-XT adapters contain the 96 combinations of barcodes that allowed for the identification of the reads of each of the 96 isolates pooled together in the same Illumina MiSeq run, by demultiplexing the samples at the end of the run. Similar to

[2] The Quant-iT™ PicoGreen® dsDNA reagent is an ultra-sensitive fluorescent nucleic acid stain for quantitating double-stranded DNA (dsDNA) in solution as little as 25 pg/mL of dsDNA with a standard spectro fluorometer and fluorescein excitation and emission wavelengths.

PCR1, PCR2 amplicons were run on a gel, visualised under UV light, and the correct size of amplicons was determined.

5.4.3.2.5 Pooling of PCR 2 amplicons

Altogether, for each sample, there were 14 amplicons from two fragments of the seven MLST genes. From each of the 14 amplicons per sample, 25 µl of normalised amplicon was pooled together into each well of a 96-well plate such that each well represented an individual sample. The plate with 96 samples (containing a mixture of 14 amplicons per sample) was submitted to MGS for amplification of the library in preparation for sequencing. In order to test the library quality before pooling them into one tube, 12 of the 96 PCR2 amplicons (from the first row of the submitted plate) were run on a DNA High Sensitivity Lab chip using an Agilent 2100 Bioanalyzer.

All the 96 samples were pooled by equal volume from a plate into one tube, and the quality of pooled amplicons was checked again using a Qubit® dsDNA HS Assay kit (Invitrogen, California, USA) with the Qubit Fluorometer (Invitrogen, California, USA), following the manufacturer's instructions. Briefly, 10 µL of each of the two Qubit standards supplied was added to 190 µL of Qubit working solution and mixed by vortexing. Two microliters of each PCR amplicon was added to 198 µL Qubit working solutions and mixed by vortexing. All tubes were incubated at room temperature for two minutes prior to quantification using the dsDNA High-Sensitivity assay type in the Qubit fluorometer. The two standards were used initially to calibrate the fluorometer.

5.4.3.2.6 Next generation sequencing

Finally, amplification of the library in preparation for sequencing was performed on an Illumina MiSeq as a 2 x 250 base paired end run (Figure 5.5).

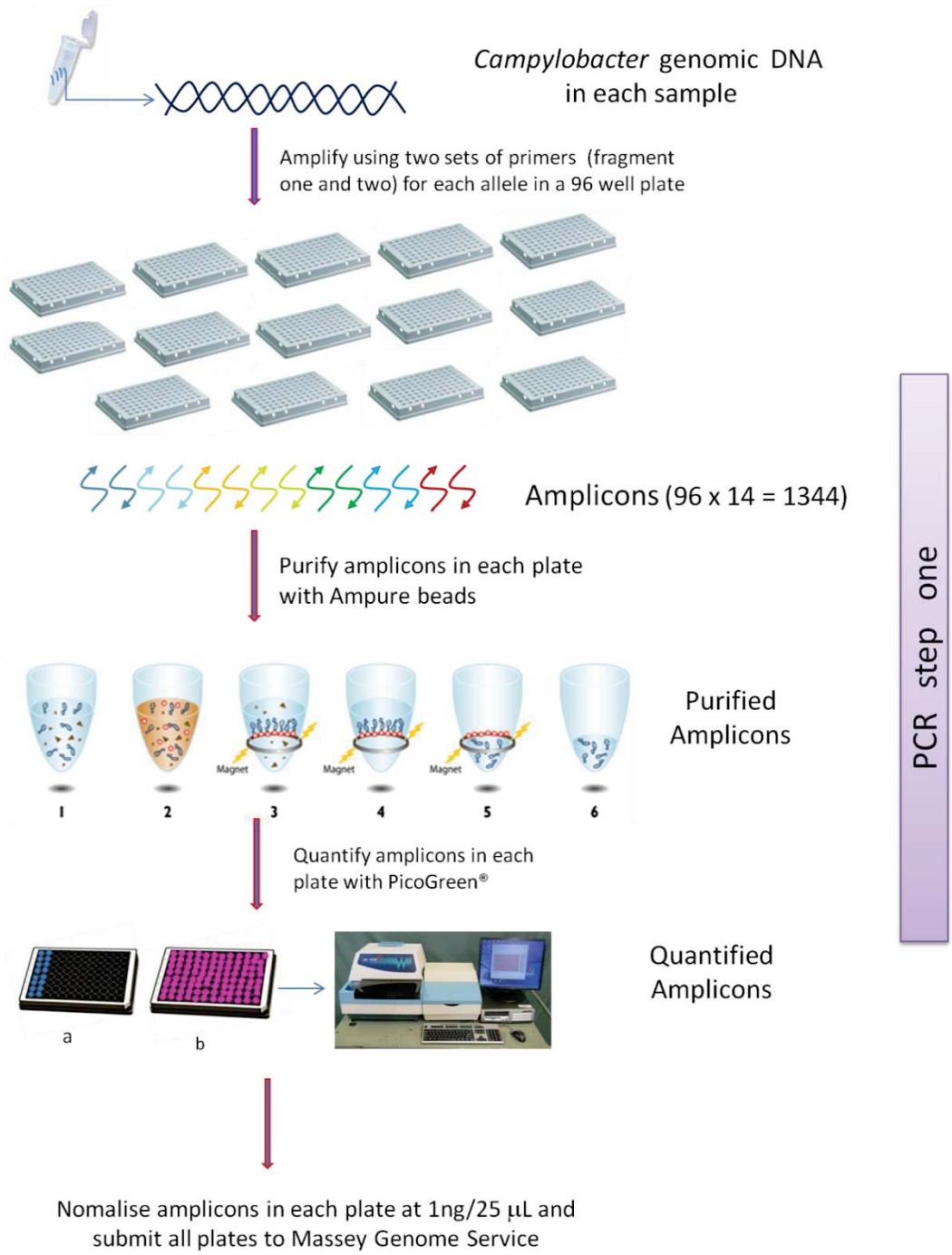


Figure 5.3: The amplification, purification, quantification, and normalisation steps involved in the first round of PCR (PCR 1) in the massMLST method. Purification steps 1 to 6 represent: PCR reaction (1); binding of amplicons to magnetic beads (2); separation of amplicons from contaminants using magnetic beads (3); ethanol washing of magnetic beads and amplicons (4); elution of amplicons from the magnetic beads (5); purified amplicons ready to transfer into the new plates (6) (adapted from http://www.beckmancoulter.com). The 'a' in quantifications step represents the standard and 'b' denotes the amplicons prepared for fluorometer.

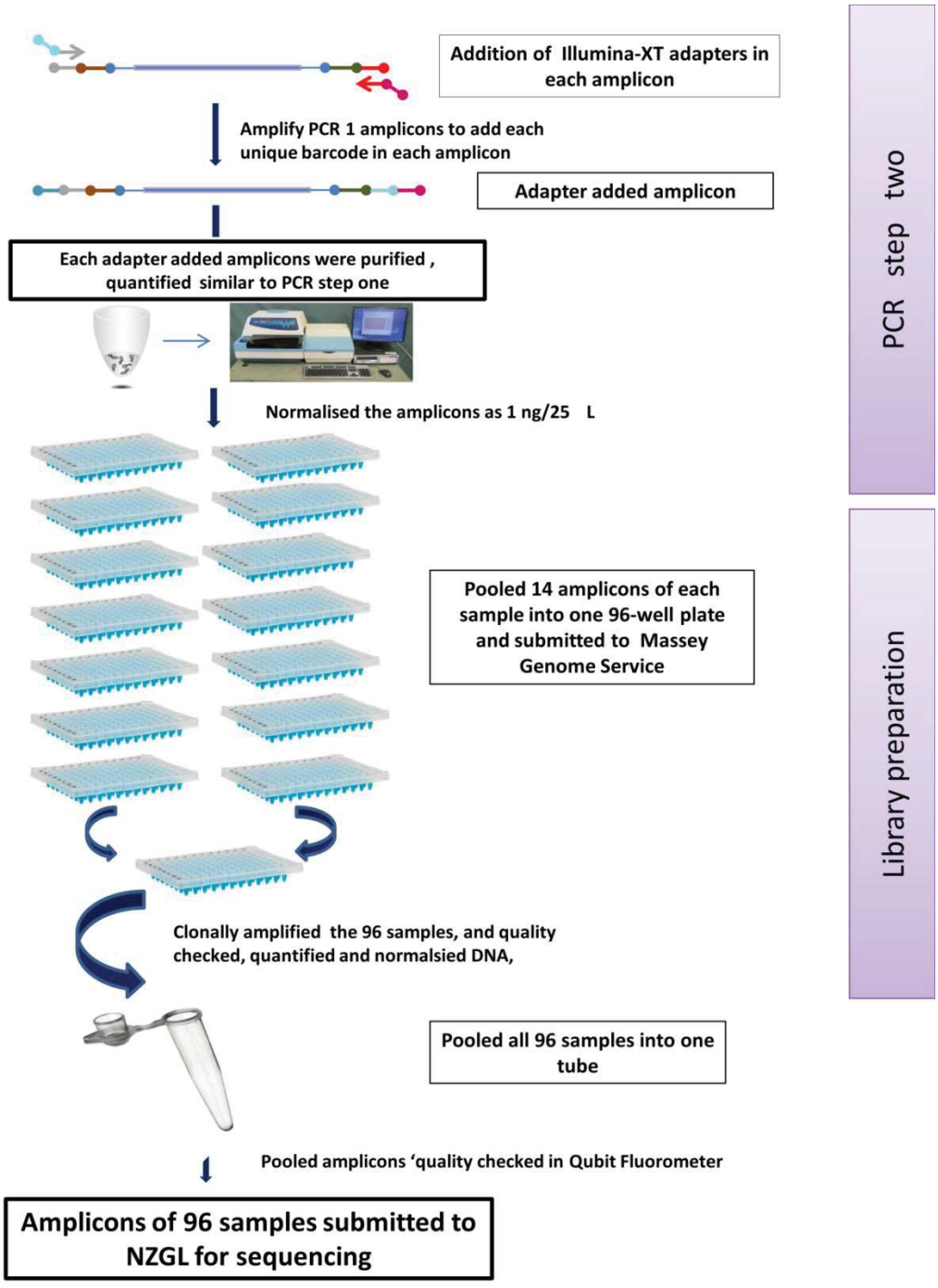


Figure 5.4: Amplification of PCR1 amplicons in order to add the adapters, and the purification, quantification, and normalisation steps involved in the second round of PCR (PCR2) in the massMLST method. The new amplicons are pooled into a single 96-well PCR plate, thereafter pooled into one PCR tube, and then quality checked before sending for sequencing.

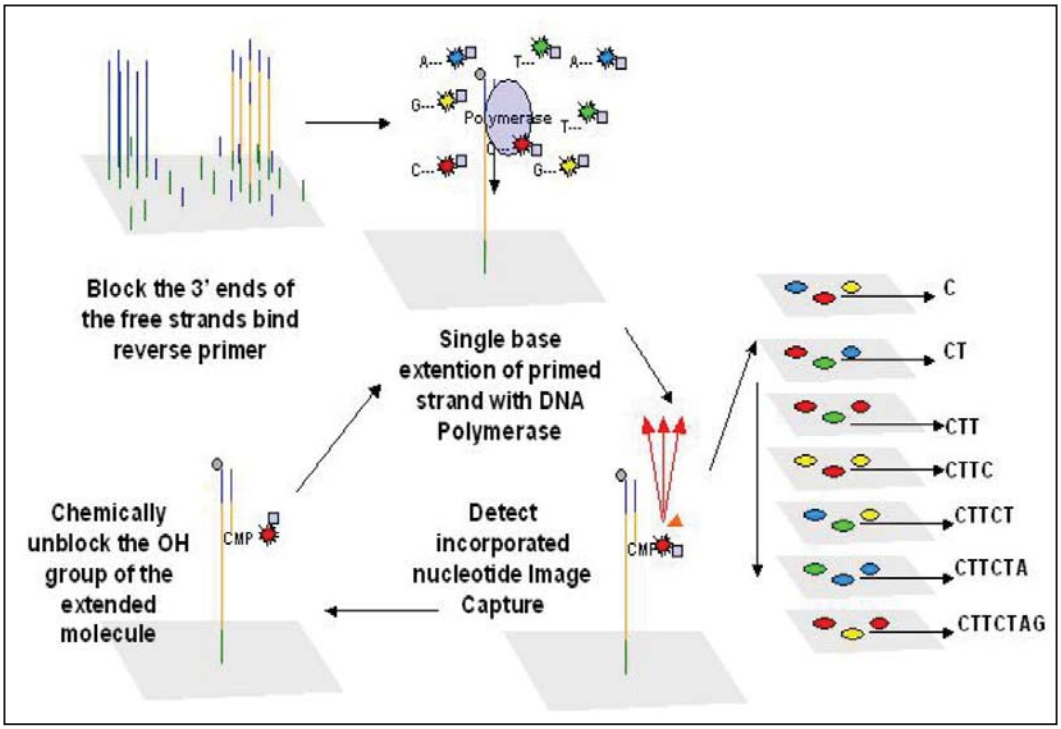


Figure 5.5: Steps involved in sequencing the amplicons from the prepared library of 96 samples using an Illumina MiSeq machine. This diagram is adapted from http://www.ipc.nxgenomics.org/newsletter/no11.htm.

Table 5.3: Preparation of reaction master mixes containing PCR buffer, Mg^{2+}, dNTPs, Taq polymerase enzyme, and water for amplification of the isolates.

Allele name	**PCR master mix (1400 µL)**					
	PCR buffer	**Mg^{2+} concentration**	**dNTPs**	**Taq polymerase enzyme**	**Water (depending on DNA concentration)**	
Concentration	10 x	50 mM/20 µL	2 mM/10 µL	5 U/ µL	If DNA volume is	water
mAspAF1/ GlnAF1/ mGltAF1/ mGlyAF1/ mPgmF1/ AspAF2/ mGlnAF2/ mGltAF2/ mGlyAF2/ mTktF2/ mAtpAF2	200 µL	60 µL	100 µL	20 µL	3 µL 2 µL 1.5 µL 1 µL	920 µL 1020 µL 1070 µL 1120 µL
mTktF1	200 µL	60 µL	100 µL	20 µL	4 µL 3 µL 2.5 µL 2 µL	820 µL 920 µL 970 µL 1020 µL
mAtpAF1/mPgmF2	200 µL	60 µL	100 µL	20 µL	4 µL 3 µL 2.5 µL 2 µL	775 µL 875 µL 925 µL 975 µL

Table 5.4: PCR conditions used in this study to amplify the two fragments of seven housekeeping multi-locus sequence typing alleles of *Campylobacter* species. Temperatures (Temp.) stated are in Celsius (^{o}C) and times are in minutes (min) or seconds (s).

Description	**Fragment 1**													
Allele name	mAspAF1		mGlnAF1		mGltAF1		mGlyAF1		mPgmF1		mTktF1		mAtpAF1	
Final product size	522		482		429		434		489		387		354	
PCR cycles	35		38		38		38		35		38		35	
PCR conditions	Temp.	Time	Temp.	Time	Temp.	Time	Temp.	Time	Temp.	Time	Temp.	Time	Temp.	Time
Initial denaturation	95	15min	95	5min	95	15min	95	5min	95	15min	95	5min	95	5min
Denaturation	94	30s	94	40s	94	30s	94	45s	94	30s	94	45s	94	45s
Annealing	50	30s	48	45s	50	30s	48	75s	50	30s	48	90s	48	75s
Extension	72	90s	72	90s	72	90s	72	90s	72	90s	72	90s	72	90s
Final extension	72	7min	72	3min	72	7min	72	7min	72	7min	72	7min	72	7min
Description	**Fragment 2**													
Allele	mAspAF2		mGlnAF2		mGltAF2		mGlyAF2		mPgmF2		mTktF2		mAtpAF2	
Final product size	324		436		435		442		389		501		496	
PCR Cycles	35		35		38		38		38		38		35	
PCR conditions	Temp.	Time	Temp.	Time	Temp.	Time	Temp.	Time	Temp.	Time	Temp.	Time	Temp.	Time
Initial denaturation	95	15min	95	15min	95	15min	95	5min	95	5min	95	5min	95	15min
Denaturation	94	30s	94	30s	94	30s	94	40s	94	45s	94	40s	94	30s
Annealing	50	30s	50	30s	50	30s	48	75s	48	75s	54	90s	50	30s
Extension	72	90s	72	90s	72	90s	72	90s	72	90s	72	90s	72	90s
Final extension	72	7min	72	7min	72	7min	72	7min	72	7min	72	7min	72	7min

5.4.4 Sequence data analysis

The next-generation sequencing library output consists of a set of "raw"[3] sequences, as well as a library generated from the internal control genome - PhiX. Therefore, the Burrows-Wheeler Aligner (BWA) software package was used to map the sequences against the PhiX genome and, thereby, identify and remove the control library and any other PhiX sequences (Li and Durbin, 2010). The remaining raw sequences were then converted to Fastq format using the SamToFastq.jar programme from Picard[4] suite and were subsequently named "processed" sequences. To get sequences of a better quality, they were trimmed at a quality cut-off of 0.01 to their longest contiguous segment using the DynamicTrim algorithm from the SolexaQA package (Cox et al., 2010), and named as "trimmed processed" sequences. Trimming may produce short reads of any length, and those ≤25 bp sequences were removed with LengthSort from the SolexaQA package (Cox et al., 2010).

The quality of sequences was further checked by detecting any single nucleotide polymorphisms (SNPs) or indels present in the PhiX genome and finally analysed through a set of quality checking tools – SolexaQA (Cox et al., 2010), fastQC[5] and fastQscreen[6]. SolexaQA determines the error probabilities by analysing 10,000 random sequences per tile and produces a heat map on a cycle by cycle basis, for each of these 28 tiles in the MiSeq ran. It also produces a histogram showing the longest contiguous sequence (cut-off = 0.05) and a cumulative graph of trimmed read lengths (Cox et al., 2010). FastQC gives a quick overview of any problems present in the datasets by providing summary graphs and tables in an HTML format. FastQscreen was run against a set of sequences to determine the level of potential contamination of PhiX, yeast, *E. coli*, Illumina adapters and cloning vectors.

The quality checked sequences were analysed using custom Perl scripts, and an overview of the workflow is shown in Figure 5.6. Firstly, the reads were separated into 96 samples. Each sample contains four reads, the two reads (forward and reverse) representing fragment one and the remaining two (forward and reverse) belonging to fragment two of that sample. Each of these reads was shorter than the

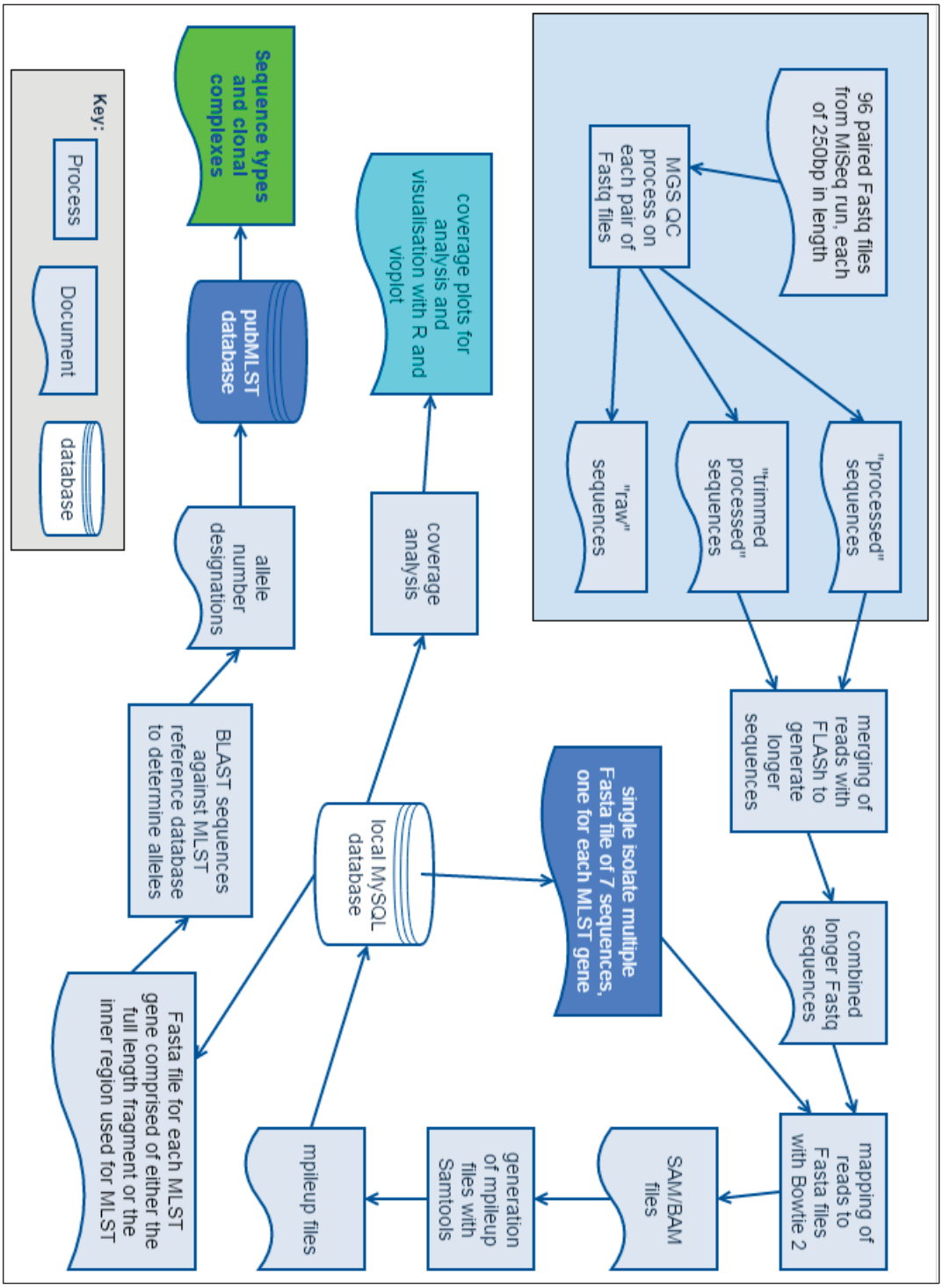
[3]These are Fastq sequence files that are generated by the automatic demultiplexing performed by the MiSeq machine when given a sample sheet to identify which samples are identified by which index.
[4] http://picard.sourceforge.net/
[5]http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ ;
[6]http://www.bioinformatics.babraham.ac.uk/projects/fastq_screen/

fragment length; therefore, two reads of each fragment were merged (such that each fragment is a single read of both the forward or reverse sequences) before using them for further assembly. For each sample, the paired-end reads were overlapped and merged with the overlapper software FLASh (Fast Length Adjustment of Short reads) to extend the reads by edging them together (Magoč and Salzberg, 2011). FLASh processes each paired-end read separately and searches for the correct overlap between the paired-end reads so that there is an ungapped alignment between two reads with at least min-olap[7] bases overlapping one another (Magoč and Salzberg, 2011). Once the correct overlap was found, the two reads were merged and produced an extended read that matched the length of the original DNA fragment from which the paired-end reads was generated (Magoč and Salzberg, 2011). This process was repeated when the overlap was longer than min-olap by calculating the overlap length and scoring the overlap as the ratio between the number of mismatches and the overlap length, ignoring N's if present.

The quality of generated paired-end reads should be assessed before mapping reads to a reference genome in order to avoid the introduction of chimeric reads and/or other sequencing artefacts in the alignment generated by some mapping or assembly tools (Li et al., 2008). One of the methods of improving the accuracy of the aligned reads was the trimming of processed sequencing reads. Trimming of the processed sequences was also performed in this study ("Processed trimmed data") to compare if there was any difference in the quality of trimmed and untrimmed (raw processed data) sequence data analysis. Trimming was performed as described using the DynamicTrim algorithm from the SolexaQA package (Cox et al., 2010). Both "processed" and "processed trimmed" overlapped reads were then aligned to each length of a MLST gene sequence using a memory-efficient short read aligner, Bowtie 2 (Langmead et al., 2009; Langmead and Salzberg, 2012). The aligned reads for each MLST gene were processed by SAMtools to generate a pileup file (SAM/BAM format) from which a sequence consensus file for two regions (either whole PCR products or the inner MLST region) was extracted, and a defined consensus sequence for each region was generated (Li et al., 2009).

[7] Minimum Overlap (min-olap) is set to be 10 bp. Magoč and Salzberg (2011) reported that lower values of min-olap resulted in many incorrectly extended reads as shorter overlaps often occur by chance in large WGS dataset, while, higher values of min-olap will miss too many true overlaps though it reduces bad merges further.

Figure 5.6: A flow diagram showing the steps involved from processing the resultant sequences to determining the sequence types and clonal complexes. The inset diagram in the light blue rectangle is the start of the sequence data analysis process, conducted by MGS, and green box is the final output of the analysis.



5.4.4.1 Descriptive analysis of the processed raw and consensus sequences

Initially, the nucleotide coverage (numbers of sequences) generated for each MLST gene per sample, each sample well on a plate, and for seven housekeeping genes of *Campylobacter* were visualised using R software (R Core Team, 2013). The nucleotide coverage data were highly skewed. Therefore, for visualisation purposes, the coverage data were log-transformed, and violin plots were produced using the "vioplot" package in R software (Hintze and Nelson, 1998; Addler, 2005; R Core Team, 2013). The violin plots were also produced from non-transformed nucleotide coverage data for each housekeeping MLST allele of *Campylobacter* to show the distribution of single nucleotide polymorphisms (SNPs) and to show consensus coverage from the full-length PCR products. The SNPs were determined using the VarScan[8] programme after mapping reads from SNP calls and adjustment of the quality score of a -10 log10 of VarScan's p-value from Fisher's Exact Test.

5.4.4.2 *Campylobacter* spp. sequence types and associated clonal complexes identification

The alleles for each MLST region were identified by choosing the best hit on the BLAST (Basic Alignment Search Tool) bit-score after a BLAST search of the consensus sequences against known MLST alleles. The bit-score describes the overall quality of an alignment, with higher numbers corresponding to a higher similarity. The best hit of the consensus sequences was chosen by sorting the bit-score, and choosing the highest bit-score and the longest alignment that matched with known MLST alleles. Where there was not a single hit of the best match, up to 10 top bit-scores that best matched with the known MLST alleles were chosen, and generated output as an MLST profile. The MLST profiles were generated from both "processed" and "processed trimmed" sequences.

Each assigned allele number in the generated MLST profiles were then entered into the search section of the *C. Jejuni* and *C. coli* PubMLST[9] database to determine the possible sequence types and clonal complexes from the allele combinations, where known. If the ST could not be determined for any isolate, potential STs were determined by mapping the sequence of the locus that has more than two different allele numbers against the reference database. For example, the possible MLST

[8] http://varscan.sourceforge.net/
[9] http://pubmlst.org/campylobacter/

profiles for one isolate indicated that it could be ST 61. However, this isolate had two possible allele numbers (say, 1 and 10) for the *asp*A allele and the allele number 10 was best hit (99%), but in fact it should have been allele number 1. Therefore, the reference sequences of those two *asp*A allele numbers (1 and 10) were mapped to each other and the number of variant nucleotides between those two allele numbers was determined, which may further help in identification of actual allele numbers. Possible clonal complexes were assigned if there were five or more allele matches to that complex.

5.5 Results

5.5.1 PCR optimisation and sequencing of isolates

Using the varied concentration of the PCR components described in Table 5.4, all 14 targets of seven housekeeping genes in 96 *Campylobacter* samples were amplified, with the amplicons processed for normalisation. Nearly 1299/1344 (96.7%) of amplicons fulfilled the minimum concentration requirements for amplicon sequencing. Therefore, the remaining 45 amplicons were sent for Sanger sequencing, which was 100% successful in determining the allele number[10] for that isolate (Table 5.5).

5.5.2 Sequence data analysis

Full and trimmed sequences reads from 96 isolates correctly identified 66/96 (68.75%) and 3/96 (3%) isolates with the full-length MLST alleles (~3309 *nt* across 7 loci) sequenced, respectively. When these sequences were analysed for each allele per sample, the majority of the *atpA* alleles of both processed (23/96) and processed trimmed (92/96) reads were not sequenced at full length (~489 bases).

[10] Each unique sequences of allele have been given a number, by combination of which sequence types of *Campylobacter* are determined.

Table 5.5: The number of samples (N=96) of *Campylobacter* that were successfully amplified for sequencing using NGS for each of the seven housekeeping genes, and the number of unsuccessful samples (N=45) that were amplified for traditional MLST using Sanger sequencing.

House-keeping genes	Markers or fragments	No. of samples successfully amplified for NGS	Samples that were sent for Sanger sequencing	
			Number	Sample names
aspA	Fragment 1	95	1	-
	Fragment 2	96		S716a
atpA/uncA	Fragment 1	96	2	-
	Fragment 2	94		S1123a, S869a
pgm	Fragment 1	96	1	-
	Fragment 2	95		S775a
gltA	Fragment 1	95	6	S843a
	Fragment 2	91		S1147a, S1123a, S771a, S1053a, S845a
glnA	Fragment 1	96	3	
	Fragment 2	93		S1036a, S1053a, S862a
tkt	Fragment 1	88	10	S1152, S994a, W725a, P1262a, S923a, S1053a, S1065a, S1123a
	Fragment 2	92		S862a, S869a
glyA	Fragment 1	78	22	S1152, S718a, S777a, S1150a, S1175a, S886a, S794a, S1032a, S1144a, S1147a, S1043a, S1036a, S1037b, S923a, S1053a, S1146a, S1189a, S1201a, S1205a
	Fragment 2	94		S994a, S715a, W725a

5.5.2.1 Descriptive analysis of the processed raw and consensus sequences

Violin plots in Figures 5.7 and 5.8 illustrate the overall distributions of single nucleotide polymorphisms (SNPs) and consensus sequences at the full length of each of the seven *Campylobacter* housekeeping MLST alleles, using the allelic regions from "processed" and "processed trimmed data", respectively. Both figures show that distribution of coverage at each nucleotide of the alleles *tkt* and *glyA* were the lowest in comparison to the other five alleles. In Figure 5.7, the median number of SNPs (upper graph) sequences were found to be the highest in the *gltA* allele, whereas a wide range of SNPs distribution (with the highest maximum number of sequences) was found in the *aspA* allele.

Conversely, both the median and range of nucleotide coverage in the consensus sequences (lower graph of Figure 5.7) were found to be the greatest in the *asp*A allele. Similar patterns were also observed in the processed trimmed sequence data (Figure 5.8). However, the range of sequence distributions of each allele was comparatively lower than the untrimmed (or processed only) data. In particular, distributions of *pgm* allele sequences were decreased greatly.

Figure 5.9 portrays violin plots of the log-transformed, processed sequence data for all alleles per sample by plate location. Overall, there was good evenness in the median distribution of sequence coverage between samples, although approximately half of the samples had a wide range of sequence coverage distribution. Conversely, when nucleotide coverage between samples for each allele was analysed using violin plots, wide variations in coverage between samples per allele were observed. An example of the nucleotide coverage (log scale) between samples for the *gltA* allele is shown in Figure 5.10. Unlike Figure 5.9, there is no evenness in the median distribution of sequences between samples for *gltA* allele. In addition, the range of sequences was partial in more than half of the samples.

Graphs for the remaining six alleles were shown Figures E.1-E.6 in the Appendix. These graphs illustrated that nucleotide coverage between samples among all the seven alleles varied greatly as some alleles had wide ranges of coverage while others had either curtailed or bi-peak sequence coverages for the same sample. For example, the sample in E01 plate had either a bi-peak or curtailed nucleotide coverage. Such coverage result was due to variation in the amount of 14 MLST gene pooled and their amplification for that sample (See Figures in 5.11 and E.1-E.6).

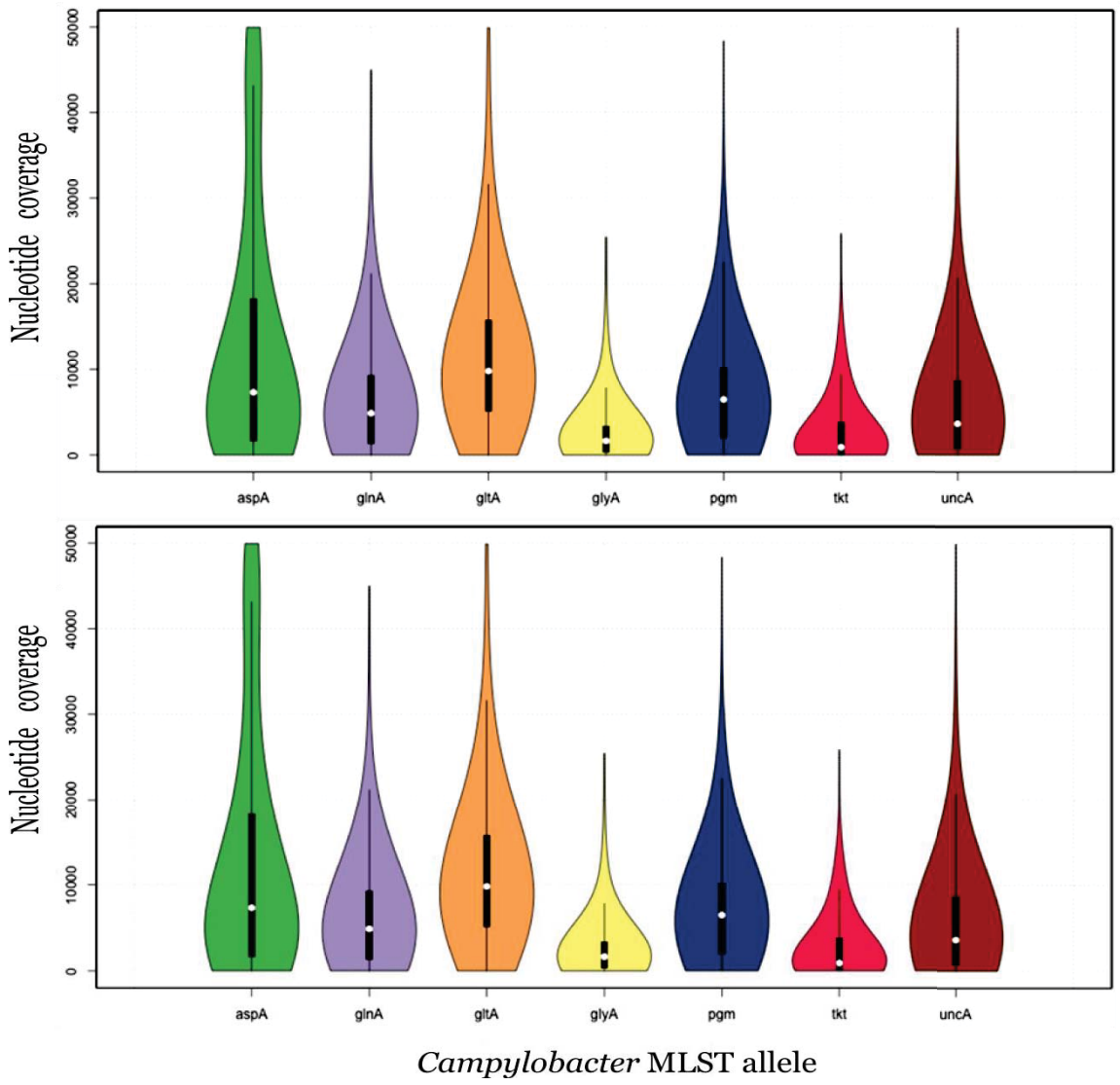


Figure 5.7: Violin plots of the overall coverage of nucleotide sequences for each of the seven *Campylobacter* MLST alleles using the data from the full-length PCR products to show SNP data. The upper and lower graph represents the "processed" and "processed trimmed data".

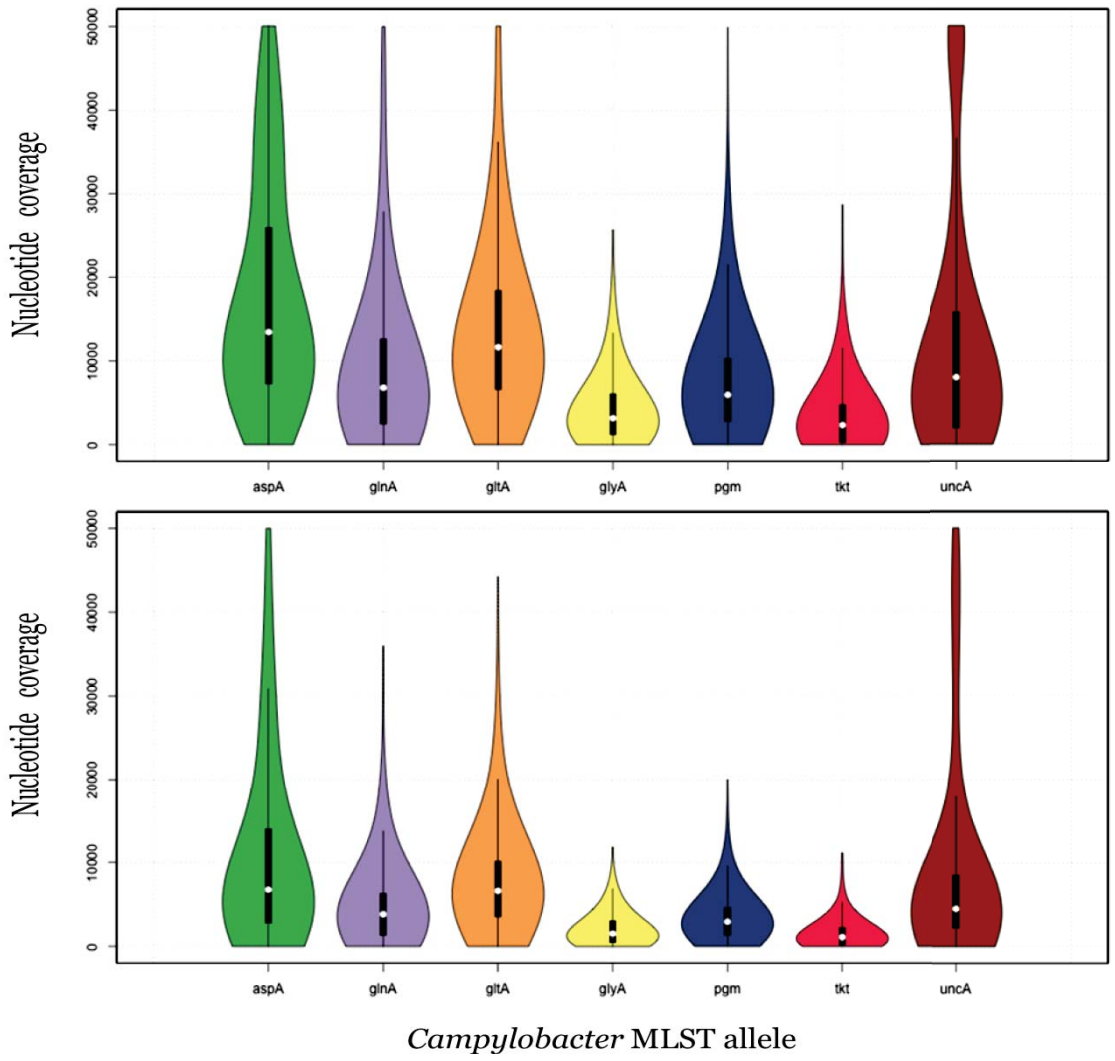


Figure 5.8: Violin plots of the overall coverage of nucleotide sequences for each of the seven *Campylobacter* MLST alleles using the data from full-length PCR products to show consensus coverage. The upper and lower graph represent the "processed" and "processed trimmed data".

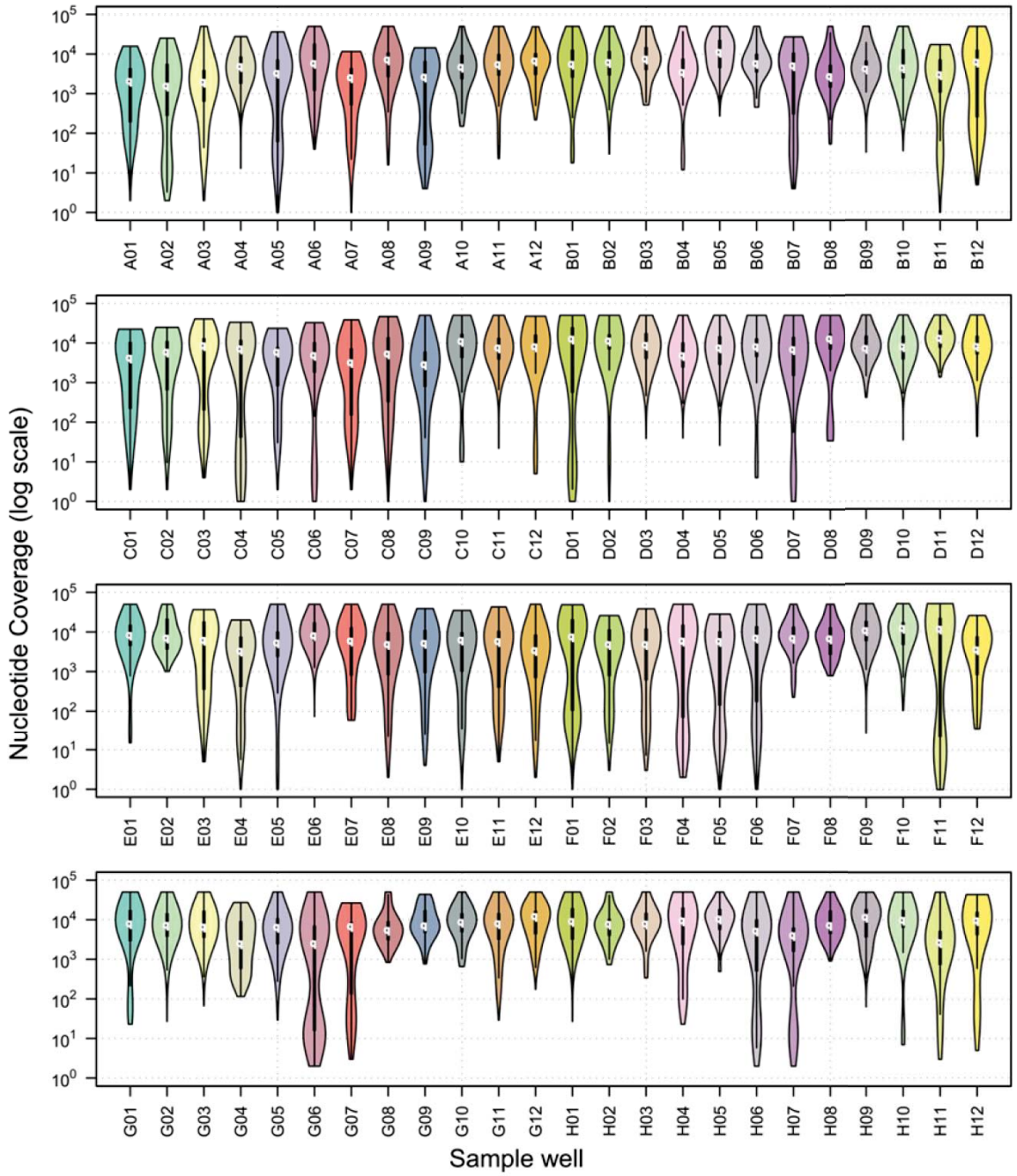


Figure 5.9: Violin plots of nucleotide sequence coverage in a log scale for each of 96 samples by plate location for all seven *Campylobacter* MLST alleles. The A01 to H12 represents wells in a 96-well plate, each of which contains a separate sample. The white dot is the median nucleotide sequence coverage and the dark black lines are the interquartile range along with whiskers.

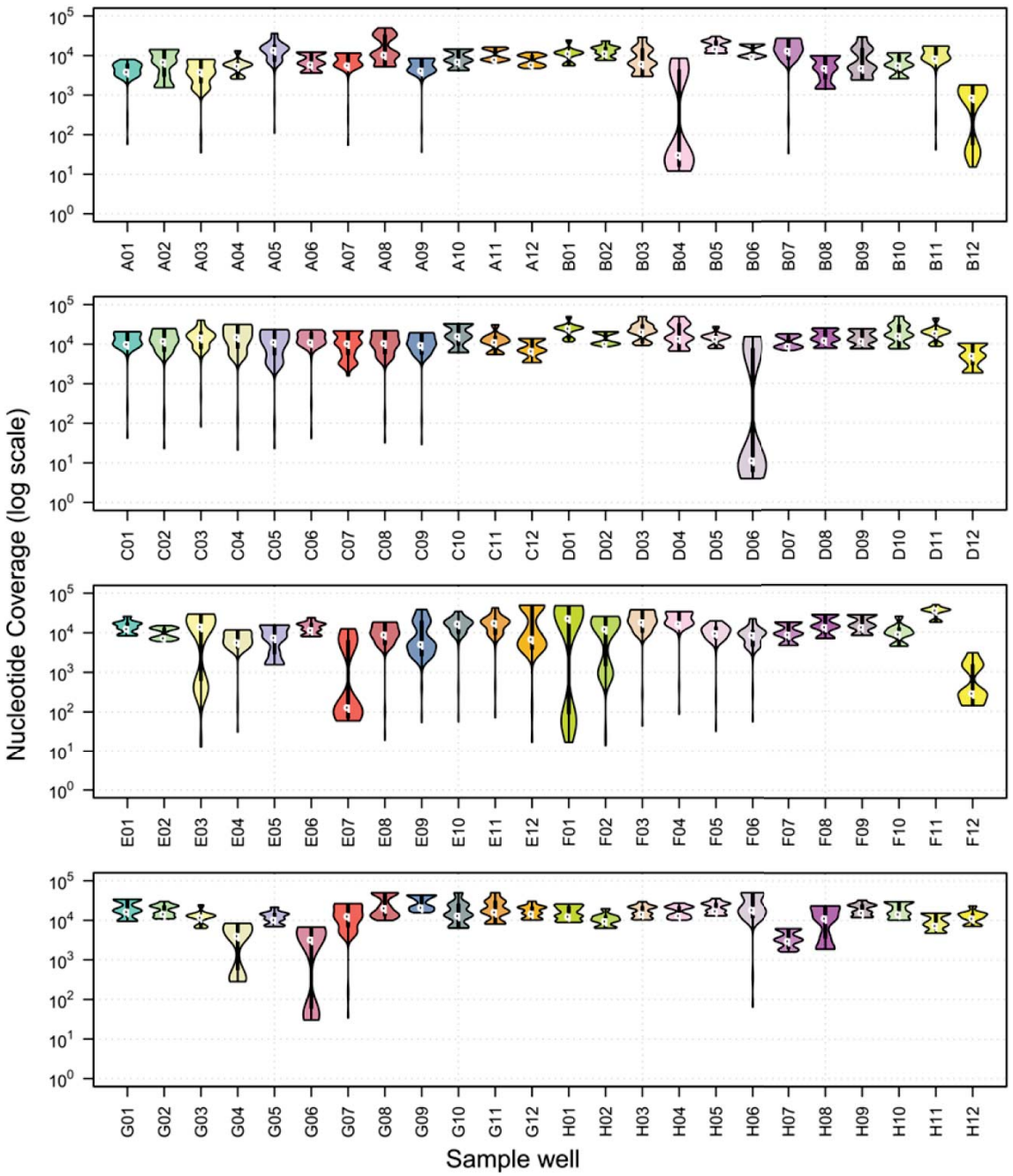


Figure 5.10: Violin plots of nucleotide sequence coverage in log scale for each of 96 samples by plate location for the *gltA Campylobacter* MLST allele. The A01 to H12 represents wells in a 96-well plate, each of which contains a separate sample. The white dot is the median nucleotide sequence coverage and the dark black lines are the interquartile range along with whiskers.

5.5.2.2 *Campylobacter* spp. sequence types and associated clonal complexes identification

The consensus sequences were searched with BLAST against a reference genome (NCTC_11168/NC_002163) of *Campylobacter jejuni/coli* MLST genes and the sequence-types (ST) for 10 of 96 samples annotated. Table 5.6 shows the allele numbers assigned and sequence types determined for each *Campylobacter* isolate from both "processed" (on the left side) and "processed trimmed" (on the right side) sequences. Full details of the table are shown in Table E.2 in the Appendix. The remaining 86 of 96 samples generated up to the six highest bitscores (i.e., have a combination of two to six different numbers for the same allele); of which many samples (66/86) had generated only two highest bitscores (see Tables in 5.6 and E.2). Those with the two highest bitscores were used to assign the possible sequence types by observing the best match scenario of allele numbers for potential STs. For some alleles, generated allele numbers from the massMLST were compared and replaced with MLST (Sanger sequencing) results and the possible STs were identified from those samples (see Table E.2).

For example, the human (H1579) and poultry (P1262a) control samples shown in Table 5.6 had assigned two different allele numbers for *unc*A/*atp*A (28 and 65) and *gly*A (247 and 4) alleles, respectively. The assigned allele numbers for these two isolates were searched in the PubMLST[11] database, and the PubMLST outputs were compared with the allelic profiles of known ST-61 (*aspA* 1, *glnA* 4, *gltA* 2, *glyA* 2, *pgm* 6, *tkt* 3, and *uncA* 17) and ST-45 (*aspA* 4, *glnA* 7, *gltA* 10, *glyA* 4, *pgm* 1, *tkt* 7, and *uncA* 1) for human and poultry control samples, respectively. This comparison confirmed that the poultry control sample was *C. jejuni* ST-45 because *glyA* allele number 4 determined in this massMLST result was matched with the PubMLST output. However, neither of the *uncA* allele numbers determined was matched with the PubMLST output. Therefore, the reference sequence for *uncA* allele 17 of ST-61 was BLASTed against *uncA* alleles 28 and 65 sequences to determine the variation of nucleotides between *uncA* alleles 17 and 28, and 17 and 65. The BLAST results showed one nucleotide difference between *uncA* 17 and *uncA* 28 alleles, whereas *uncA* 17 and *uncA* 65 had a difference of 51 nucleotides. Similarly, when the reference sequence glyA allele 247 sequence was compared to the glyA allele 4 by BLAST, a difference of only two nucleotides was found. These nucleotide variations were observed in the middle part of most of the aligned sequences. In the majority

[11] http://pubmlst.org/ campylobacter/

of the samples, variations in allele numbers were seen in the *gly*A, *gln*A, *tkt* and *unc*A alleles.

5.5.2.3 Occurrences of *C. jejuni* and *C. coli* strains in ruminant faeces and water

Based on the best match scenario, 30/96 samples were assigned to 16 different possible sequence types (STs) of which three were new sequence types, and 78/96 samples were assigned to nine possible clonal complexes (CC) including one unassigned CC (Table 5.7). Of these, two *C. jejuni* STs- 45 and 61 were associated with the control samples used in this study. Among the ten confirmed sequence types determined in this study: three were ST-4337, three were a new ST belonging to clonal-complex 61, two were ST-50, and one each were ST-53 and ST-2381 of *C. jejuni*.

In this study, 30/70 and 64/70 *C. jejuni* isolates originated from various sources were assigned to various possible sequence types and clonal complexes (Table 5.7), respectively. These sequence types and clonal complexes have been previously reported in the PubMLST database. Among the *C. jejuni* isolates from various sources, ST-42, ST-45, ST-2776, ST-4337 and one new sequence type identified originated from sheep; ST-19, ST-53, ST-4337, ST-61, ST-257 and two new sequence types detected were from dairy cattle; and ST-21, ST-50, ST-190, ST-1823, ST-61, ST-3425 were from beef cattle (Table 5.7). Many *C. jejuni* CC-21 that were determined had originated from dairy cattle (10/19), whereas the *C. jejuni* CC-61 determined had commonly originated from beef cattle (12/23) (Table 5.7). Clonal complexes 21 and 61 were determined in samples collected from all the seven farms, whereas ST-61 was reported in samples from 5/7 farms. In contrast, the *C. coli* isolates were less diverse with almost half (12/26) belonging to CC-828 (7/12) and ST-3222 (5/12). The majority of *C. coli* CC-828 and ST-3222 originated from sheep faecal samples (4/5 and 5/7, respectively) (Table 5.7).

Among the four *C. jejuni* isolates originating from water samples, ST-2381 was determined in one sample, and the remaining three were *C. jejuni* isolates possibly belonging to CC-3640.

Table 5.6: Examples showing the assigned allele numbers for each housekeeping gene of "processed" and "processed trimmed" sequences generated for each massMLST sample. The complete profile with defined sequence type (ST-best case based on non-ambiguous alleles) and clonal complex (CC) is highlighted in green. Red indicates the discrepant alleles between "processed" and "processed trimmed" sequences and yellow shows the extra allele determined based on the highest bit-scores for a given sample. (#) indicates the number of best bit-score determined for that sample, (*) denotes the possible ST and CC determined for some samples, and UD** represents the ST/CC that could not be determined.

	Processed sequences											**Trimmed sequences**									
Sample ID	**max Alleles**	**#**	***aspA***	***glnA***	***gltA***	***glyA***	***pgm***	***tkt***	***uncA***	**ST**	**CC**	**max Alleles**	**#**	***aspA***	***glnA***	***gltA***	***glyA***	***pgm***	***tkt***	***uncA***
H1579	2	1	1	4	2	15	6	3	28	61*	61*	2	1	1	4	2	15	6	3	28
		2	.	.	.	.	.	.	65				2	.	.	.	.	.	.	65
P1262a	2	1	4	7	10	247	1	7	1	45	45*	3	1	4	7	10	247	1	7	1
		2	.	.	.	4	.	.	.				2	.	.	.	4	.	.	.
S1155a	6	1	33	165	222	310	118	44	17	UD**	UD**	6	1	196	165	222	242	118	164	42
		2	.	264	44	82	.	.	28				2	33	264	**44**	.	209	44	.
		3	.	340	.	.	.	.	.				3	.	340	.	.	413	493	.
		4	.	50	.	.	.	.	.				4	.	50	.	.	.	77	.
		5	.	57	.	.	.	.	.				5	.	57	.	.	.	.	.
		6	.	61	.	.	.	.	.				6	.	61	.	.	.	.	.
S1025a	1	1	2	1	12	3	2	1	5	**50**	**CC21**	1	1	2	1	12	3	2	1	5

Table 5.7: Numbers of *Campylobacter jejuni* and *C. coli* clonal complexes (CC) and sequence types (ST) determined in the massMLST of the 96 samples from various origins. (-) denotes the absence of those CC/ST in the given isolates.

Clonal complexes	Sequence types	Sample Origins					Grand Total
		Beef	Dairy	Sheep	water	Control	
C. jejuni		**23**	**20**	**21**	**4**	**2**	**70**
21		**6**	**10**	**3**	**-**	**-**	**19**
	19	-	1	-	-	-	1
	21	2	-	-	-	-	2
	50	1	-	1	-	-	2
	53	-	2	-	-	-	2
	190	1	-	-	-	-	1
	1823	1	-	-	-	-	1
	4337	-	3	1	-	-	4
	UD**	1	4	1	-	-	6
42		**1**	-	**11**	**-**	**-**	**12**
	42	-	-	1	-	-	1
	UD**	1	-	10	-	-	11
45	45	**-**	**-**	**1**	**-**	**1**	**2**
48	48	**1**	**-**	**-**	**-**	**-**	**1**
61		**12**	**8**	**2**	**-**	**1**	**23**
	61	3	2	**-**	**-**	1	6
	2776	**-**	**-**	1	**-**	**-**	1
	3425	1	**-**	**-**	**-**	**-**	1
	-	8	4	**-**	**-**	**-**	12
	New	-	2	1	**-**	**-**	3
257	257	**-**	**1**	**-**	**-**	**-**	**1**
828	UD**	**1**	**-**	**2**	**-**	**-**	**3**
3640	UD**	**-**	**-**	**-**	**3**	**-**	**3**
UA*	2381	**-**	**-**	**-**	**1**	**-**	**1**
UD**	UD**	**2**	**1**	**2**	**-**	**-**	**5**
C. coli		**2**	**5**	**19**	**-**	**-**	**26**
UA**		**-**	**1**	**5**			**6**
	3222	-	1	4	-	-	5
	UD**	-	-	1	-	-	1
828	UD**	**-**	**2**	**5**	**-**	**-**	**7**
UD**	UD**	**2**	**2**	**9**	**-**	**-**	**13**
Grand Total		**25**	**25**	**40**	**4**	**2**	**96**

*Clonal complexes were not assigned yet in the PubMLST database

**Clonal complexes could not be determined for the *C. jejuni*/coli isolates analysed.

5.6 Discussion

This is the first study that attempts to sequence the full-length multi-locus sequence typing (MLST) of *Campylobacter jejuni* and *C. coli* amplicons using NGS technology, with the aim of improving the "traditional" MLST method. This massMLST method has simultaneously genotyped 96 isolates of *Campylobacter* species in comparison to the traditional MLST method.

5.6.1 Primer design, PCR optimisation and sequencing of isolates

The massMLST method was developed and run on a MiSeq machine (Illumina®). Although, at the time of writing the MiSeq machine can sequence a read of 300bp length, at the time of this study the MiSeq machine had the limitation of sequencing a read of up to 250bp. Therefore, the initial challenge for this study was to design a correct primer containing the Illumina® tags at one end, which could generate the accurate overlapped PCR products, and optimise the primers precisely. We successfully re-designed the primers with tags at one end, and the redesigned primers were verified by analysing each of the aligned primer sequences with the reference *Campylobacter* genome sequence (NCTC_11168/NC_002163), and testing them *in silico*. These primers were also optimised *in vitro* using known *C. jejuni and C. coli* samples, and the massMLST primers were successfully amplified.

Another important challenge was to amplify each of the two fragments per allele in the *Campylobacter* isolates using the redesigned primers because the massMLST protocol includes a master mixture preparation and the use of 96 samples at a time. Therefore, there is an increased possibility of: error in mixing the two isolates; or in having less master mixture preparation than required; or mixing the one primer with another; or in setting the mixture for a long time on the PCR bench before running it. In this study, two people were involved in running each PCR, and one person verified the other's work to avoid these errors. In this way, 14 primers and permissive PCR conditions allowed the detection of seven housekeeping genes of *Campylobacter* spp. in all the control samples and 96.7% of 94 isolates used for this study. Similar challenges were also present in the purification, quantification, and normalisation of amplicons, which were also performed by two people, verifying each other's work, with complete agreement in the results. Therefore, it was very important to consider the co-

ordination of two people during the running of the PCR, and during the purifying, quantifying and normalisation of the amplicons.

The design of a two–step PCR protocol with the tag at one end of the primers to incorporate the Illumina adapters has made this massMLST method flexible and economical as multiple samples from a variety of sources and species can be incorporated easily. This concept is similar to the 16S rRNA amplicon sequencing and has been effectively used successfully in high-throughput MLST (Boers et al., 2012; Singh et al., 2012).

5.6.2 Sequence data analysis

One of the objectives of this study was to sequence the full-length multi-locus sequence typing allele of *Campylobacter* spp. using NGS technology. Nearly 69% (66/96) of the samples resulted in full-length sequences for all alleles, whereas, the remaining samples did not amplify in either of one or a pair of primers used. In this study, the *atpA* alleles were successfully amplified for both pairs of *atpA* primers in the first attempt during PCR and they fulfilled the criteria for the minimum DNA amount required for sequencing. However, NGS sequenced only 73/96 samples to full-length for the *atp*A alleles, and the remaining samples had a variation of 1 - 8 bases in their sequences. On the contrary, many isolates that did not fulfil the minimum amount of DNA requirements for MiSeq sequencing were not sequenced in full-length for all seven alleles, as can be seen in the "processed trimmed data". These results showed that amplified DNA must meet the minimum DNA requirement criteria. If the concentration of DNA was high, however, we pipetted only small volumes, which led to the higher chances of volumetric error. However, the automation of the dilution and pipetting of the sample could reduce this error in the future.

5.6.2.1 Descriptive analysis of the processed raw and consensus sequences

Nucleotide coverage was not balanced across all the alleles (Figure 5.8), particularly for the *tkt* and *glyA* alleles' sequence of this study. It was reported that the amplicon size influences the number of reads obtained (Boers et al., 2012). Other factors that may play a role in the nucleotide coverage is the GC-content and intrinsic properties of the amplicon PCRs (Boers et al., 2012). In this study, amplicon size and GC-content might have played a minor role in getting unbalanced nucleotide coverage (Table 5.2). More balanced distributions of reads can be obtained by adapting the volumes of

amplicons during the pooling of the sample prior to PCR/sequencing. Therefore, the optimal volume required should be calculated empirically for each individual allele. Using this strategy, we were able to gain a more balanced distribution of reads by adapting the volumes of the *Campylobacter* spp. housekeeping genes during pooling of the sample. However, there is a variation in the individual allele nucleotide coverage per sample. As previously mentioned, automation of pipetting of samples for pooling could reduce such variation in the future.

Similarly, when processed sequences were trimmed, the median and range of nucleotide coverage for each allele were also reduced (Figure 5.7 and 5.8). This is expected as the numbers of reads covering each base were reduced with the increase in quality of sequences.

5.6.2.2 *Campylobacter* spp. sequence types and associated clonal complexes identification

Currently, MiSeq technology produces reads of either 250 or 300 bases, depending upon the run. The existing *Campylobacter* MLST protocols require 400-510 bp long sequences for profiling the *C. jejuni* and *C. coli* subtyping. Therefore, a bidirectional read of the amplicons is necessary to obtain a full sequence, but it cannot be performed with only one PCR product in this application. Because of this, correct read overlapping and sufficient nucleotide coverage are important issues to consider if you wish to confidently determine the correct nucleotides.

With massMLST, only 10/96 *C. jejuni* and *C. coli* isolates were profiled completely, and this determined the sequence types (ST) and associated clonal complexes (CC). This method identified three new sequence types among the 10 identified sequence types. In fact, all the isolates in this study were profiled to the ST or CC level. However, 66/96 isolates contained two different allele numbers in any of the seven housekeeping genes. With the absence of one allelic profile, and matching the best-case scenario, it is still possible to predict the possible clonal complexes and sequence types for that isolate. Using this strategy, we investigated the possible CC and ST for those 66 isolates, and no attempt was made to assign any ST or CC in the remaining 30 isolates as these isolates had more than two possible allele numbers (3 - 6) assigned.

One of the reasons for getting more than one hit or allele number could be due to the wrong annotation of the overlapped products in synthesising the one sequence of that gene. As a result, there were equal numbers of 2 - 6 different reads for that allele

resulting in similar bit-scores. Nevertheless, this problem may be negligible in the future if the MiSeq machine can sequence the read of ~550-600bp as it will sequence the full-length MLST gene in a single run, thus avoiding the preparation of overlapping sequence products.

The sequence types and clonal complexes determined in this study showed that cattle and sheep on farms of both study areas, Dannevirke and Shannon (Table 5.7), shed *Campylobacter* spp. with zoonotic potential. However, ST-2381 the ST that was frequently isolated from native Pukeko birds and from New Zealand water samples (Kvalsvig et al., 2014) was also determined in *Campylobacter* isolates from water used in this study.

5.6.2.3 *C. jejuni* and *C. coli* strains determined in ruminant and water samples and their public health implications

The massMLST *C. jejuni* and *C. coli* isolates demonstrated the presence of nine clonal complexes (CC), seven of these that originated from ruminants were circulating in the environment of Dannevirke and Shannon catchment areas of the Manawatu-Wanganui region (Table 5.7). The majority (83% of 64) of the *C. jejuni* strains belonged to CC-61 (34%) followed by CC-21 (30%) and CC-42 (19%), and these isolates were isolated from samples originating from ruminants, poultry and people (Table 5.7). This finding may have significant implications for disease control and prevention, because these strains are not only the most frequently isolated genotypes from samples provided by people in New Zealand and worldwide, but they also have the capacity to cause disease in ruminants and people (Dingle et al., 2001; 2002; Friedrich et al., 2015). Moreover, the remaining three clonal complexes (CC-257, CC-828 and CC-48) identified have also been associated with human infections, albeit on a smaller scale (Dingle et al., 2001; 2002; Friedrich et al., 2015). Since these isolates came from samples that had been collected from the ground on farms, the finding of these CCs supports the hypothesis that cattle, sheep, and their associated environment could act as important reservoirs for human campylobacteriosis, either through direct contact or indirectly through the contamination of water bodies.

The genotypic composition of the *C. jejuni* population in cattle was found to be consistent with longitudinal studies (Kwan et al., 2008a; 2008b), where CC-21 and CC-61 were highly prevalent and the ST-42 complex less common. These findings in our study further support the hypothesis that CC-61 and CC-21 isolates may be from a cattle-adapted *C. jejuni* lineage as suggested previously (Dingle et al., 2002; Manning et

al., 2003; Kwan et al., 2008). CC-828 determined in one cattle isolate in this study has also been found in other studies, suggesting cattle may be a potential source of CC-828 in the farm (Kwan et al., 2008a; 2008b). Further, CC-48 and CC-257 found in cattle, that are mainly associated with poultry, were also reported in cattle as a carrier during the examination of PubMLST database. Given the evidence that these clonal complexes have been isolated from New Zealand cattle and cattle overseas, along with their reported importance in human infections, the cattle may be an important source of human infections in these two catchment areas as well.

CC-42 was over-represented in sheep isolates, and this finding is in agreement with the previous studies in New Zealand and overseas (Dingle et al., 2002; Manning et al., 2003; Kwan et al., 2008). Other studies have also reported the other clonal complexes CC-21, CC-61, CC-45 and CC-828 from sheep sources, but these were less represented in this study. Similar to cattle CCs, when CCs in sheep sources were investigated in the PubMLST database, all these complexes have also been isolated from human sources suggesting that sheep should also be taken into consideration as a potential source of human disease via food and environmental routes.

The massMLST analysis determined two clonal complexes, CC-828 and an unassigned complex, and one ST (ST-3222) of *C. coli*. It has previously been reported that *C. coli* is less diverse than *C. jejuni* and is host-associated, which could be the reason why this study found that the *C. coli* isolates belonged to a low number of CC or ST (Dingle et al., 2005b). The ST-828 complex has been reported to be found in a variety of sources including cattle, sheep, poultry and environment samples, suggesting that many different transmission routes may play a role in the epidemiology of *C. coli (*Miller et al., 2005; PubMLST database). In addition, this complex is isolated from cases of human campylobacteriosis cases worldwide (Piccirillo et al., 2014). The ST-3222 of *C. coli* determined in this study was represented in the PubMLST database by an isolate that had originated from a sheep source in New Zealand and that had been determined by Penner typing in 2001. This finding suggests that ST-3222 could be restricted to certain geographic areas. Nevertheless, further studies should be conducted on *C. coli* subtypes in various animal and environmental sources in order to understand the epidemiology of *C. coli* in more detail.

This study demonstrated three new STs that belong to the ST-61complex, one from cattle and two from sheep sources, and a number of uncommon/undetermined sequence types, particularly subtypes of *C. coli* from sheep sources. This observation

may in fact suggest that nucleotides of the allele sequences undergo recombination or mutation generating new STs within the same clonal complexes, or that there is much nucleotide variation in overlapped reads during overlapping and merging of forward and reverse sequences of the allele. Further studies on a larger scale and use of better alignment and overlapping software, may be able to identify undetermined sequence types.

C. jejuni CC-3640 (3/4) and ST-2381 (1/4) were determined in four water isolates using massMLST. Our retrospective study on *C. jejuni* in recreational waters of New Zealand in Chapter 2 also determined CC-3640 and ST-2381, the latter being predominant in all the rivers sampled as part of that study. These findings further support the statement by Carter *et al.* (2009) that they are widespread, endemic, and unique to New Zealand. To date, ST-2381 has been isolated from the native wild birds (pukeko and takahe) and is possibly not pathogenic to humans (Kvalsvig et al., 2014). CC-3640 complex was shown to be related to ruminant-associated CC-42 (Carter et al., 2009), suggesting a possible link to cattle and sheep that needs to be further investigated. In addition, CC ST-3640 was found to share a number of rare alleles such as *aspA* 175, *gltA* 216, and *glyA* 282 with ST-2381. Of note is that the *glyA* 282 allele has only been identified in isolates sourced from rivers in New Zealand.

From the public health viewpoint, finding strains that are associated with ruminants and that have previously been implicated in human campylobacteriosis studies implies that there is a potential risk of zoonotic transmission of *Campylobacter* in the study areas. This highlights the need to recognise ruminants and their associated environment, including water bodies and the grounds of the farm as important potential reservoirs for human disease, particularly for clonal complex strains ST-61, ST-21, and ST-42 of *C. jejuni* and ST-828 complex of *C. coli*. Furthermore, there is potentially a lower risk of human infection through the river/stream water due to the findings that most of the *C. jejuni* strains identified from these sources were those associated with wild birds that are seemingly non-pathogenic for humans. Furthermore, the treatment of drinking water extracted from river/streams decreases the risk of transmission of campylobacteriosis by this route still further.

5.7 Conclusion

To conclude, this is the first study that uses NGS to characterise 96 *C. jejuni* and *C. coli* isolates in a single MiSeq run as a named method "massMLST". The massMLST

approach opens new perspectives for large-scale application of the robust MLST technique for *Campylobacter* species. The massMLST results in a substantial reduction of costs compared to traditional Sanger sequencing, and possibly, it will become an attractive and feasible technique for molecular typing, and high-throughput analysis of a large collection of *Campylobacter* strains in the near future.

This study added knowledge and understanding of *C. jejuni* and *C. coli* strains originating from ruminants and water in New Zealand, and further increases our ability to identify the potential zoonotic and waterborne sources of campylobacteriosis. However, these findings need to be further corroborated through studies involving a larger number of isolates from various sources and using the latest MiSeq machine that can sequence full-length in single PCR.

5.8 Highlights of massMLST study

- This is the first study that attempts to use NGS technology for multi-locus sequence typing of *Campylobacter* spp. isolates from samples collected from cattle, sheep, and water sources, as well as the control isolates from human and poultry origin.
- The massMLST method, in which 96 samples were sequenced in one NGS run, is an improvement of the current multi-locus sequence typing (MLST) method for *Campylobacter*. Amplification was successful for 96.7% (1299/1344) of alleles from 96 samples.
- Nucleotide coverage was distributed nearly evenly across all samples when all alleles in each sample were included. However, nucleotide sequence coverage for full-length data per allele varied widely, particularly for *tkt* and *glyA* alleles.
- Although only a few samples (n = 10) were completely profiled, it was also possible to profile other samples (n = 66) for possible ST or CC, indicating that this method has future scope if its shortcomings such as sequence data analysis are overcome.
- Based on the profiles generated from the sequenced samples, cattle and sheep in the Manawatu region of New Zealand are shedding *Campylobacter jejuni* and *C. coli* with zoonotic potential. The isolates from water, however, were associated with wild birds and have not previously been identified from cases of human disease.

5.9 Acknowledgements

This study would not be complete without acknowledging many people who have helped directly and indirectly for this project. Special thanks go to my co-supervisor, Dr Patrick Biggs, who generated the idea of combining MLST with NGS technology, and was involved in a major part of bioinformatics analysis. In addition, special thanks also go to Lynn Rogers for assisting me in laboratory procedures, Dr TanuShree Dutta for providing ELISA plates for quantification of DNA, Dr V.S.R Dukkipati for providing valuable suggestions, and Massey Genome Service for providing a magnet plate.

This project was funded by IVABS Post-graduate Research Fund and McGeorge Research Fund.

Chapter 6

Microbial communities present in the surface water destined for drinking purposes: use of next-generation sequencing technology

6.1 Preamble

In the earlier studies (chapters 2 and 3), potentially zoonotic pathogens *Campylobacter*, *Cryptosporidium* and *Giardia* species were identified in the surface water used for recreational purposes and/or as a drinking source. These pathogens were genotyped and/or subtyped using traditional typing methods: species-specific polymerase chain reaction (PCR) and Sanger sequencing. However, isolation and identification of these pathogens from water samples are not always successful. Many methods have been developed, including next-generation sequencing technology to improve the pathogen identification from water samples. Next-generation sequencing technology has also been used in understanding the microbial ecology of drinking source water, including identification of zoonotic pathogens. As a result, the work in this chapter was performed as a pilot study to get an overview of the microbial communities present in the surface water destined for drinking and to compare metagenomes found in different methods used for sample collections and preservation.

6.2 Abstract

Monitoring of drinking water is regularly practised in many countries, including New Zealand. However, there are limited studies on microbial communities present in the

surface water of New Zealand. Therefore, this study aims to provide a snapshot of the microbial ecology of surface water used in rural towns for drinking purposes and to determine if pathogens related to waterborne diseases using next-generation sequencing technology. Fresh samples collected in the years 2011 and 2012, samples from the same time that were frozen, and samples that were kept in a preservative called RNA*later* were sequenced using whole-genome shotgun sequencing on an Illumina MiSeq machine. *Proteobacteria* was detected in all the samples characterised, although there were differences in the genus and species level comparison between the samples. There were variations in the microbial diversity reported between the sampling of low (grab) and large volume (filtered) water samples, between samples collected in the year 2011 and 2012, and among fresh, frozen and RNA*later* preserved samples. In addition, further analysis of sequences suggested the presence of various waterborne pathogens such as *Cryptosporidium* spp. and *Campylobacter* spp. This is the first pilot study that attempted to analyse the microbiome through metagenomics approaches using next-generation sequencing technologies (NGS) in different volumes of surface water samples collected and in various sample preservation techniques used. Although this study has some limitations, NGS methods could possibly be used to monitor pathogens in drinking water sources, and thereby to formulate an effective water treatment plan for reducing the future waterborne zoonotic risk to the public.

6.3 Introduction

Microorganisms are omnipresent in the environment. However, approximately 1% of microorganisms in environmental samples are only able to grow in the laboratory due to our inability to find the right culture conditions with our current knowledge and technical limitations (Ghai et al., 2011). As a result, we rely only on organisms detected in the laboratory to track the source of water contamination, and on a few indicator bacteria (e.g., *E. coli*) to measure water quality for drinking purposes (Gronewold et al., 2008). Consequently, we underestimate the waterborne disease risk as many potentially pathogenic but undetermined organisms are circulating in the environment that could potentially risk human health. In addition, multiple studies have shown that the use of indicator organisms does not always reflect the actual water quality due to their unestablished correlation with other organisms such as *Cryptosporidium,* difference in survivability, and spatial and temporal variation of microorganisms in aquatic environments (Horman et al., 2004; Bonadonna et al.,

2002; Ishii and Sadowsky, 2008; Wilkes et al., 2009). Hence, the understanding of microbes present in drinking water sources is important for improvements in public health.

Recent improvements in molecular biology techniques have overcome the shortcomings of the traditional molecular techniques such as culture and PCR of each organism at one time without providing the information on those organisms' viability. In addition, the ability to analyse multiple samples in one run such as with next-generation sequencing (NGS) technology has dramatically reduced the cost of such analyses. Consequently, the application of NGS is becoming more widely used in environmental and public health studies. Despite the limitations of throughput, speed, scalability, resolution, and the incurred high cost to sequence a large number of samples, Sanger sequencing has been used in traditional techniques in the genotyping of pathogens (Metzker, 2005; [1]). However, with the advent of NGS, these limitations of Sanger sequencing have been overcome to a large degree. Therefore, NGS has also attracted people working in clinics, public health, and environmental health field (Metzker, 2010; Liu et al., 2013; [1]). For example, NGS was employed in the 2011 *E. coli* O104:H4 outbreak in Germany to determine the source of infection (Brzuszkiewicz et al., 2011; Mellmann et al., 2011); in surface water samples to identify the faecal pollution source and in drinking water to identify the microbial communities after the treatment process (Brzuszkiewicz et al., 2011; Shanks et al., 2011; Revetta et al., 2013; Baron et al., 2015).

The metagenome refers to the any DNA present in the ecosystem, for example, fungal, bacterial, and protozoal, algae and even human contamination (Handelsman et al., 1998; Singh et al., 2009). NGS has been employed as a culture-independent method for direct genetic analysis of metagenomes present in the ecosystem, and the technique is referred to as "metagenomics" (Handelsman et al., 1998; Thomas et al., 2012). Metagenomics also provides information on potentially novel biocatalysts, the metabolic activity of genomes and their function and structure, and the role of potentially novel genomes in evolution (Simon and Daniel, 2011). Metagenomics studies, like other studies, have challenges in getting pure DNA, avoiding contamination of DNA, pooling sequences and high processing costs. Nevertheless, a dramatic reduction of cost and fast turnaround of NGS have accelerated and widen the study of metagenomics using a whole genome shotgun sequencing (WGS)

[1] http://www.illumina.com

approach rather than analysis of a single gene such as the prokaryotic 16S ribosomal RNA (rRNA) gene study (Simon and Daniel, 2011; Thomas et al., 2012; Tringe and Rubin, 2005).

In metagenomics study, the whole genome shotgun sequencing has been introduced to replica entire library and to provide a broader view on microbial community rather than a single genome study (Tringe and Rubin, 2005). In whole-genome shotgun sequencing (WGS), the abundant species are expected to be deeply covered and well assembled, while, species of lower abundance may be represented by only a few sequences (Venter et al., 2004). Thus, WGS potentially provides a better picture of diverse microbial communities and their metabolic pathways and helps in assessing functional genes, intraspecies polymorphism and potentially discovering new genes. In fact, metagenome shotgun sequence datasets have been widely used in microbial ecology studies such as soil, water and in applied research, such as in animal production or in disease diagnosis (Morgavi et al., 2013; Holinger et al., 2014; Campana et al., 2014; Deusch et al., 2014; Howe et al., 2014; Lopes et al., 2015; Martin-Cuadrado et al., 2014; Patel et al., 2014; Tan et al., 2015).

In New Zealand, the use of NGS has also increased in the past few years, but there are very limited publicly available datasets to use. Consequently, there is very little information on microbial diversity in the drinking source water in New Zealand. Therefore, with the aim of understanding the microbial diversity of the surface water destined for drinking and detecting waterborne zoonotic pathogens present in water, shotgun sequencing on Illumina NGS platforms was performed in this study. Moreover, this study analysed differences in microbial diversity between the immediately processed samples and the samples processed after the three months of preservation in a chemical solution and in a -80^{o}C freezer.

6.4 Materials and methods

6.4.1 Water sampling and concentration by filtration

Two water samples were collected from Tamaki River, Dannevirke (n = 1) and Mangaore Stream, Shannon (n = 1) in November 2011 and another two water samples were collected from the same river/stream in November 2012. At each sampling, 3 L water was collected in two sterile bottles, one of 2 L and the other of 1 L. The samples were transported in a chilly-bin to mEpiLab within an hour, stored at 4^{o}C and processed within 48 hours. In the laboratory, the samples were filtered using

a 47mm magnetic vacuum filtration funnel pump through cellulose filters (0.45 and 0.22 μm pores size, Pall Corporation, USA) to collect genomic DNA. Each water sample was divided into four parts for filtration so that the collection of DNA could be maximised without clogging the filter (personal communication, Richard Fong, NZGL, Massey University). Therefore, 750 mL of water from each sample was filtered through both filters (keeping a 0.45 μm filter sitting on top of a 0.22 μm filter) at the same time. Each filter from the 2011 samples was placed in a sterile 50 mL tube such that the upper face of the filter faced the inner side of the tube. Then, these tubes were stored at -80°C until genomic DNA was extracted. Those filters from the 2012 samples were stored in three different ways: from each site, the four filters were stored at -20°C and DNA was extracted within 48 hours of collection, whereas the other two filters were stored in a tube containing the RNA stabilisation solution known as RNA*later*® (Sigma-Aldrich Co. LLC, Missouri, USA) at room temperature, and the remaining two filters were placed in a sterile tube and stored at -80°C for three months before extraction of the DNA. The objective of using different storage methods was to compare the effect of storage on the metagenomes obtained.

Grab samples[2] are useful to explore the microbial ecology of water, especially for bacteria, fungi and virus. However, protozoa such as *Cryptosporidium* and *Giardia* are difficult to capture and assess by grabbing a small volume of water. Therefore, 100 L of water was passed through the Filta-Max™ filter (IDEXX Laboratories, Inc., Westbrook, Maine, U.S.A) from each site, as described in Chapter 3. The captured organisms, including protozoa in the filter, were then concentrated in 500 mL PBS using a stomacher (Stomacher® 3500, Seward Limited, West Sussex, UK), and the solutions were stored in a sterile bottle at 4°C until processing. The samples using the Filta-Max™ filter were collected from each site in the year 2011 only. Similar to the grab samples, half of each stomacher solution (250 mL) was also filtered through a vacuum pump containing two filters (0.45 and 0.22 μm) and each labelled filter containing tubes were stored at -80°C until processing.

From here onwards, the stomacher solution analysed is referred as a "large volume" sample and each processing method used for the four samples collected were referred as an individual sample, for example, the results obtained from RNA*later* filter as an RNA*later* sample. Therefore, altogether nine samples were processed for

[2] A single sample collected at a particular time and place that represents the composition of the water, air, or soil only at that time and place. Here the sample is water.

metagenome analysis from the two surface water locations, Tamaki River at Dannevirke (n = 5) and Mangaore Stream at Shannon (n = 4).

6.4.2 Genomic DNA extraction

For high molecular weight metagenomic DNA extraction, the filters were cut into sections using sterile scissors, following the manufacturer's instructions (Metagenomic DNA Isolation Kit for Water, Epicentre®). In brief, the cut filter was washed with prepared filter wash buffer containing 0.2% Tween 20 to extract cells on the filter. The washed suspension tube was centrifuged to pellet the cells, which then lysed by lysozyme solution, RNase A and Proteinase K. MPC protein precipitating reagent was added to the lysed solution, and was vortexed and centrifuged to collect the metagenome DNA. Finally, the DNA pellet was purified using isopropanol and 70% alcohol, air dried and suspended in 50 μL of molecular grade water. Then, the isolated DNA was validated for size by visualisation on a 2% agarose gel after electrophoresis, and for concentration by the use of a Qubit® fluorometer (Invitrogen, MA, USA) and/or NanoDrop spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA) using the manufacturer's instructions (Thermo Scientific). All the DNA samples were submitted to the Massey Genome Service for NGS library preparation and sequencing.

6.4.3 NGS library preparation

All the DNA samples were assessed for both the quantity and quality of the DNA present using a Qubit® fluorometer (Invitrogen), along with RNA or protein contamination tests. Those samples that fulfilled the minimum requirements for library preparation were only processed for library preparation. The minimum requirement for WGS library preparation is 10 and 2.5 ng/uL for TruSeq and Nextera sequencing, respectively. The Massey Genome Service prepared whole genome shotgun libraries using Illumina Nextera XT and Illumina TruSeq DNA library preparation methods.

In brief, the genomic DNA was randomly fragmented using enzymatic shearing and barcoded Illumina adapters were added onto each end of the fragments during enrichment PCR to prepare the Illumina Nextera XT library. On the other hand, the Illumina TruSeq DNA library preparation method used mechanical shearing that randomly fragmented the DNA followed by ligating barcoded-Illumina adapters onto each end of the fragments. The adapter-ligated products were run on a 2% agarose

gel to visualise the fragments with an average insert size of 400bp. The fully ligated fragments were only enriched for the final products. All the prepared libraries were cleaned up and were checked for quality control using a Qubit® fluorometer (Invitrogen) to quantify and an Agilent 2100 Bioanalyzer to visualise the expected final size (650bp) of the library. Thereafter, the libraries were normalised, pooled and denatured before sequencing on the Illumina MiSeq™ machine. The Massey Genome Service then provided the sequencing results for data analysis.

6.4.4 Metagenome data analysis

For this study, the sequence data were analysed by a bioinformatician (Dr P. J. Biggs), whereas, I used the final converted sequence data from the bioinformatician to visualise the data and produce the final epidemiological output. Briefly, raw sequence data were checked for quality and those good quality data were analysed using the PAUDA (Protein Alignment Using a DNA Aligner) (Huson and Xie, 2013) algorithm for WGS to produce aligned DNA sequences when compared to a reference database. These sequences were then visualised in metagenomics visualisation software such as MEGAN (Huson et al., 2007) or sequence analysis software such as BioNumerics (Version 5.6, Applied Maths, Ghent, Belgium).

6.4.4.1 Data quality check

The provided data consists of folders of raw and partially processed sequences. The raw sequence files (which was given name "fQsequences") were mapped and hit against the PhiX[3] genome using Burrows-Wheeler Alignment (BWA)[4] (Li and Durbin, 2009) tool to mitigate sequencing challenges in unbalanced and low-diversity libraries and generated a fastq file using the SamToFastq.jar programme from Picard[5] suite that was named "processed".

These "processed" sequences were then trimmed at a quality cut-off of 0.01 to their longest contiguous segment using DynamicTrim software from the SolexaQA[6] package. The "processed" sequences were also analysed for mapping quality by detecting any SNPs or indels present when compared to the PhiX control, and

[3] The PhiX genome is a small, diverse and well-defined genome generated from the PhiX bacteriophage that is regularly used as a positive control in Illumina DNA sequencing. It enables quick alignment and estimation of error rates.
[4] BWA is a software package for mapping short read sequences against a reference genome. http://bio-bwa.sourceforge.net
[5] http://picard.sourceforge.net/
[6] http://solexaqa.sourceforge.net/

mapping the consensus file using the varscan[7] programme. Finally, SolexaQA, fastQC and fastQscreen were run for quality check (QC) analysis of the "fQsequences". SolexaQA determines the error probabilities by analysing 10,000 random sequences per tile and produces a heat map on a cycle by cycle basis for each of 28 tiles in the MiSeq, a histogram showing the longest contiguous sequence (cut-off = 0.05) and a cumulative graph of trimmed read lengths. Finally, fastQscreen were run against a set of sequences to determine the level of potential contamination of PhiX, yeast, *E. coli*, Illumina adapters and cloning vectors.

6.4.4.2 Sequence analysis

After checking the quality of the whole genome shotgun sequences data, the PAUDA[8] programme was used to analyse the DNA sequencing reads. The reads were first translated into protein sequences using a script called "pauda-run". Then, the dna2pna programme was run to convert the translated protein sequences into pseudoDNA (pDNA). The pDNA were produced by mapping the amino acid alphabet in the sequences onto a four-lettered alphabet of DNA sequences (Murphy et al., 2000). That is, the 20 amino acids were grouped into four classes [L, V, I, M, C], [A, G, S, T, P], [F, Y, W] and [E, D, N, Q, K, R, H] which were mapped as the letters A, C, G and T respectively. All other characters (besides amino acids) in the sequences were mapped as N (also known as *BLOSUM50* reduction). Any duplicates or contained pDNA were also removed after pDNA production and *BLOSUM50* reduction. Thereby, the database should contain approximately 25% fewer entries. This step is crucial, as it will help DNA aligner not to consider duplicates/contained pDNA sequences, accelerate the analysis, and ensure that conserved sequences are not deemed "repetitive" by the employed DNA aligner. This programme creates an auxiliary file to facilitate the expansion of the computed matches back to the full reference proteins, and also converts the pDNA reference sequences into FASTA file format. Then, the bowtie2-build tool was employed to construct a Bowtie2 reference index as per the user manual of Bowtie2 (Langmead and Salzberg, 2012); and the bowtie2-align programme was employed for aligning the pDNA sequences against the PNA index (Langmead and Salzberg, 2012). All the aligned pDNA sequences were then run with the pna2blastx programme to translate the pDNA back into protein-protein alignments along with their bitScores. Finally, the programme produced BLASTX-format alignments for all protein-protein alignments based on significant bit-score of

[7] (http://varscan.sourceforge.net/).
[8] http://ab.inf.uni-tuebingen.de/software/pauda/

more than 30. The reads were mapped to the non-redundant (nr) database using PAUDA, and the resulting files were mapped to the NCBI taxonomy to generate the RMA (read-match archive) file using the MEGAN (Huson and Xie, 2013).

6.4.4.3 Data visualisation

The RMA file that summarises the BLAST results taxonomically was uploaded into MEGAN. MEGAN assigned taxa based on the lowest common ancestor algorithm. The assigned reads were used to compare the taxonomic analysis of multiple sample datasets and were visualised by MEtaGenome Analyzer (MEGAN) version 6.2 written by D.H. Huson (2007). During visualisation, data values were transformed to square root values for better comparison and output visualisation. The data were compared at the taxonomic phyla to species levels. The dissimilarity between the microbial diversity present in different samples was also analysed in MEGAN using hierarchical clustering (UPGMA tree) and principal coordinate analysis (PCoA). The PCoA uses a Euclidean algorithm for a linear mapping of distances or dissimilarities between samples on ordination space (Ramette, 2007).

6.5 Results

6.5.1 Sample analysis

The concentration of DNA from fresh samples collected in the year 2011 (0.35 ng/µL) and 2012 (1.1 ng/µL) and RNAlater preserved (2.15 ng/µL) samples from Shannon were less than the minimum DNA concentration requirement for sequencing (10 ng/µL), and thus were not analysed by shotgun sequencing (Table 6.1). Among the samples, eight of them were successfully sequenced using Illumina Nextera XT and TruSeq techniques for shotgun sequencing except the three samples from Shannon (Table 6.2). Of the eight samples, only one sample from Dannevirke collected in 2011 was run for TruSeq. The remaining number of samples run in the NexteraXT were the frozen samples from Dannevirke (n = 1) and Shannon (n = 1), the large volume samples from Dannevirke (n = 1) and Shannon (n = 1), the samples collected in 2011 (n = 1) and 2012 (n = 1) from Dannevirke, and the RNAlater (n = 1) preserved samples from Dannevirke.

Table 6.1: The minimum DNA concentration (ng/µL) required for whole-genome shotgun metagenomics library preparation method, and the DNA concentration present in the water samples collected from the two sites in the year 2011 and 2012. The libraries were run on an Illumina MiSeq machine at the Massey Genome Service.

Sequencing methods	Sampling		Samples types	DNA concentration (ng/µL)	
	sites	year		minimum required	present in the samples
TruSeq	Dannevirke	2011	Fresh[9]	10	59.9
Nextera XT	Dannevirke	2011	Fresh[9]	2.5	2.22
	Dannevirke	2012	Fresh[9]		17.5
	Dannevirke	2012	Large volume[10]		67.4
	Dannevirke	2012	Frozen[9]		14
	Dannevirke	2012	RNA*later*[9]		14.4
	Shannon	2011	Fresh		-
	Shannon	2012	Fresh		-
	Shannon	2012	Large volume[10]		14
	Shannon	2012	Frozen		1.4
	Shannon	2012	RNA*later*		2.15

Table 6.2: The sequences generated from different samples used in the whole-genome shotgun metagenomics study. (-) indicate sequencing was not conducted for that sample.

Samples location	Sampling year	Samples type	No. of raw sequences generated
Dannevirke	2011	Fresh	155429
Dannevirke	2012	Fresh	584340
Dannevirke	2012	Large volume	2180536
Dannevirke	2012	Frozen	942064
Dannevirke	2012	RNA*later*	686994
Dannevirke	2011	Fresh/TruSeq	3091315
Shannon	2011	Fresh	-
Shannon	2012	Fresh	-
Shannon	2012	Large volume	1957514
Shannon	2012	Frozen	2013
Shannon	2012	RNA*later*	-

[9] is a grab sample
[10]is a 100L water passed through Filta-Max™ filter

6.5.2 Whole genome shotgun sequencing output

The Illumina whole genome shotgun sequencing returned a range of 2013 to 3091315 reads of raw DNA sequence containing between 60 bp and 251 bp fragment size DNA in the eight samples (Table 6.2). Among the generated sequences, 85% were >100 bp and > 45% of the sequences were between 200 - 250 bp except the sequences from the Shannon frozen sample (≥ 20%). The Shannon frozen sample generated a very low number of sequences (n = 2013), therefore, it was not included in the comparative taxonomy analysis in MEGAN. As a number of reads varied among the seven samples, the reads were normalised to 155,429 reads per sample in software MEGAN. MEGAN assigned 79% of the normalised hit reads to different taxa for comparison of the diversity between the samples. Figure 6.1 depicted the comparative histogram of various types of microbes at the domain level and a number of sequence reads in each of seven samples created in software MEGAN.

The assigned sequence reads of all the seven samples were compared to the phyla, class, family, genus levels, and the microbes' abundance between the samples was explored and visualised using a word cloud within MEGAN (Figures 6.2 to 6.5). A total of 37 diverse microbial phyla, including unclassified and environmental categories were depicted in the metagenome. Of these, 24 phyla were affiliated with the bacterial domain, of which majority sequences were affiliated to *Proteobacteria,* whereas the *Opisthokonta* phylum was abundant among eight eukaryotic phyla and the most prominent of the five archaeal phyla *was Euryarchaeota*. The metagenomes analysed comprised a total of 715 families, including those that belong to viruses, with the abundant presence of *Pseudomonadaceae* followed by *Caulobacteriaceae*.

Among the assigned sequence reads, the phylum *Proteobacteria* predominated in all the samples followed by the phylum *Bacteroidetes* (Figure 6.2). At the class level, *Betaproteobacteria* predominated in all the grab samples followed by *Alphaproteobacteria* except in RNA*later* samples, where *Alphaproteobacteria* and *Gammaproteobacteria* abundances were nearly equal (Figure 6.3). On the other hand, the large volume samples had a dominance of *Gammaproteobacteria* followed by *Alphaproteobacteria* (Figure 6.3). The family *Comamonadaceae* was predominant in all the grab samples followed by *Burkholderiaceae* except in the RNA*later* samples, where *Pseudomonadaceae* was the second-most predominant families (Figure 6.4). In contrast, family *Pseudomonadaceae* was predominant in both the large volume

samples followed by the family *Caulobacteriaceae* (Figure 6.4). At the genus level *Pseudomonas* was found in greater amount in the large volume and RNA*later* samples, whereas the remaining samples were dominated by the *Leptothrix* and *Curvibacter* organisms (Figure 6.5).

These word cloud figures (Figures 6.3 to 6.5) also depict that there was more microbial' richness in the grab samples (all five samples) than in the large volume samples (Dannevirke and Shannon) and in the Dannevirke sample processed with TruSeq than Nextera XT (Figure 6.6). The fresh samples from Dannevirke collected in the year 2012 had an increased abundance of microbes when compared to those collected in the year 2011 (Figure 6.4). When comparing the sequence reads from the fresh, the 3-months frozen and the 3-months RNA*later* preserved Dannevirke samples, the microbial' richness observed was similar in the fresh and frozen samples, but this richness was reduced in the RNA*later* samples (Figure 6.7).

The two-dimensional principal co-ordinate analyses (PCoA) of seven samples showed that the microbial communities from large volume samples clustered separately to the other samples (Figure 6.8). There are also differences between the fresh samples processed in the years 2011 and 2012. However, there were not significant differences in the microbial diversity between the grab samples during the year 2012 for the different processing methods (frozen, fresh, RNAlater), and between the sequencing methods (Nextera XT vs. TruSeq) and the Dannevirke 2011 sample. In addition, the UPGMA tree in Figure 6.9 also shows similar patterns of clustering to the PCoA. For example, the large volume samples from two different locations had a similar diversity pattern and were grouped together in the UPGMA tree.

At the species level, DNA sequences of possibly belonging to various waterborne organisms such as *Campylobacter jejuni,* and *Cryptosporidium spp.* have been identified in the samples that were shotgun sequenced. However, the species level identification was not further investigated in detail.

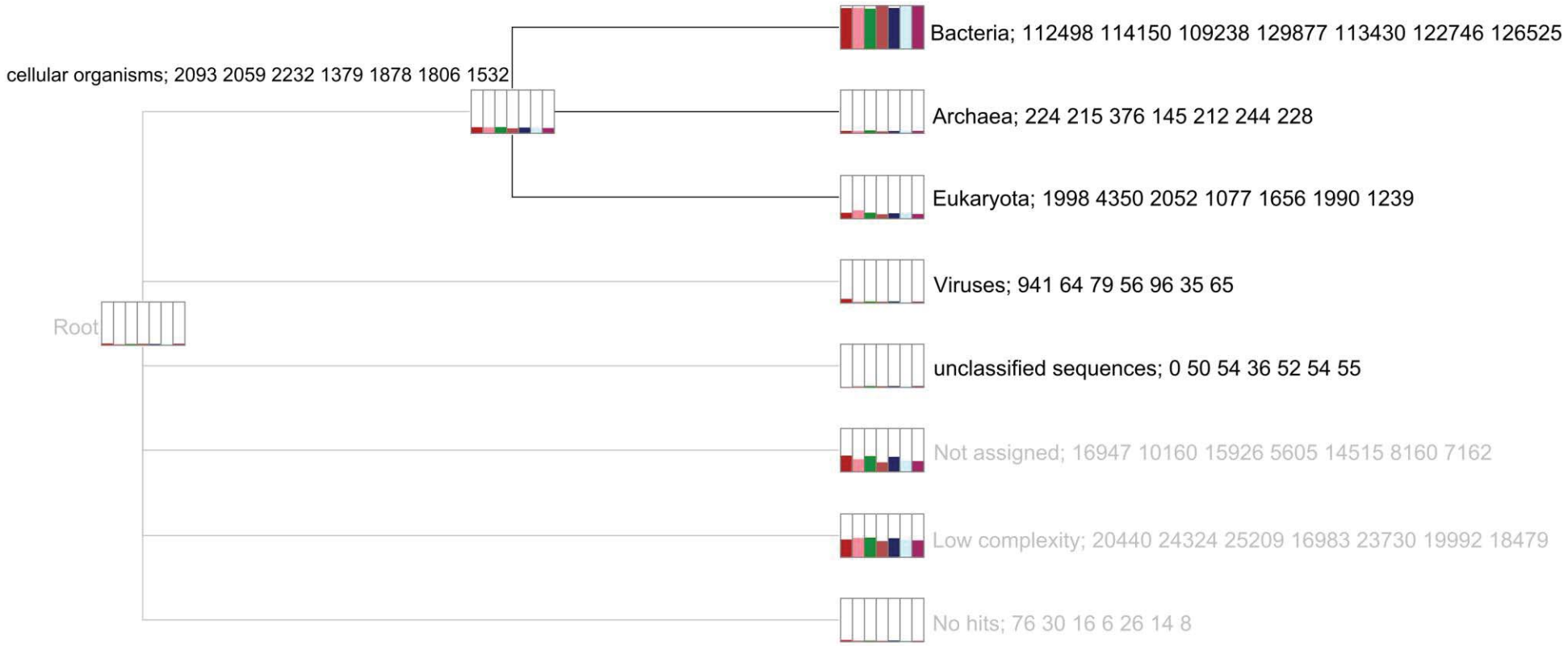


Figure 6.1: This bar-chart is a MEGAN output on whole genome sequencing data. The bar-chart shows various organisms that were obtained in NGS of the seven samples processed by different methods that are compared to their taxonomic kingdom classification. The colours in each bar represent the samples: Red-Dannevirke 2011, Pink-Dannevirke 2012, Green-Dannevirke frozen, Brown-Dannevirke large volume, Dark blue-Dannevirke RNAlater, Sky blue-Dannevirke TruSeq, and Purple-Shannon large volume.

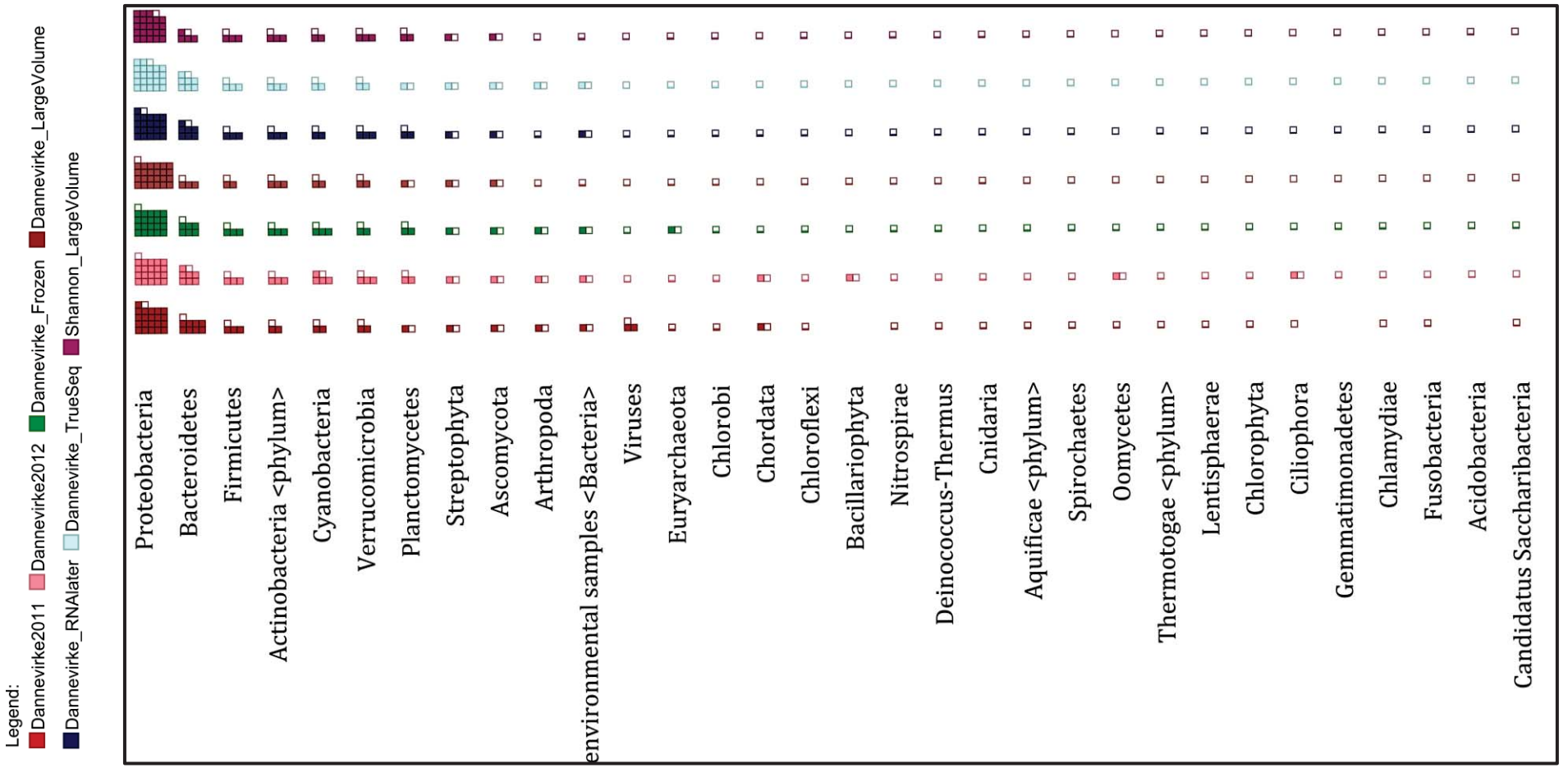


Figure 6.2a: This brick chart is a MEGAN output on whole genome sequencing data that shows predominant phyla (x-axis) in descending order for the seven samples (y-axis) processed by different methods. Each tile represents 40 sequence reads (square-root transformation). The x-axis represents different phyla. Dannevirke 2011 and 2012 are the fresh samples collected in the years 2011 and 2012. The remaining samples were samples collected in the year 2012.

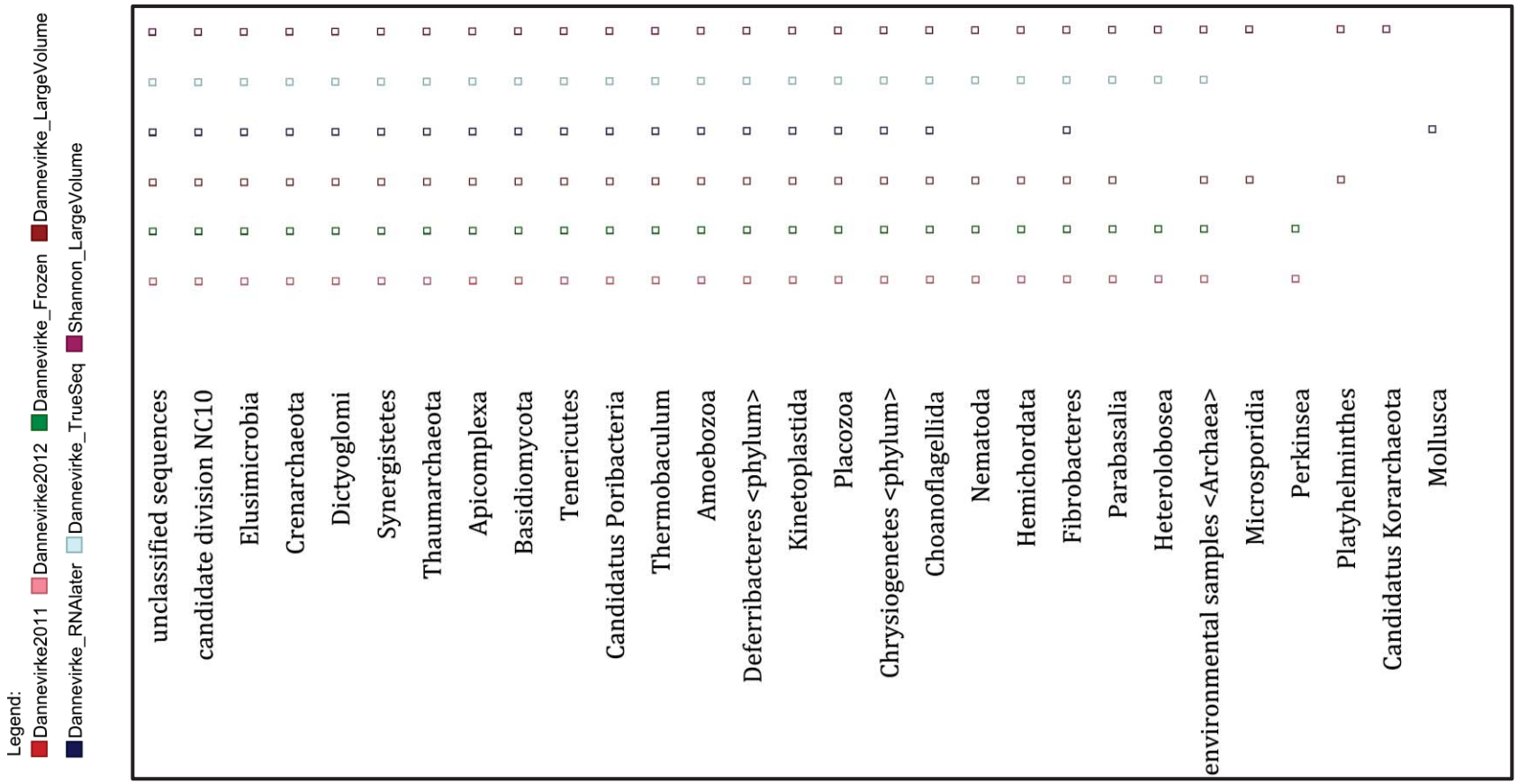


Figure 6.2b: This bricks chart is a continuance of MEGAN output of figure 6.2a that shows predominant phyla (x-axis) in descending order for the seven samples (y-axis) processed by different methods. Each tile represents 40 sequence reads (square-root transformation). The x-axis represents different phyla. Dannevirke 2011 and 2012 are the fresh samples collected in the years 2011 and 2012. The remaining samples were samples collected in the year 2012.

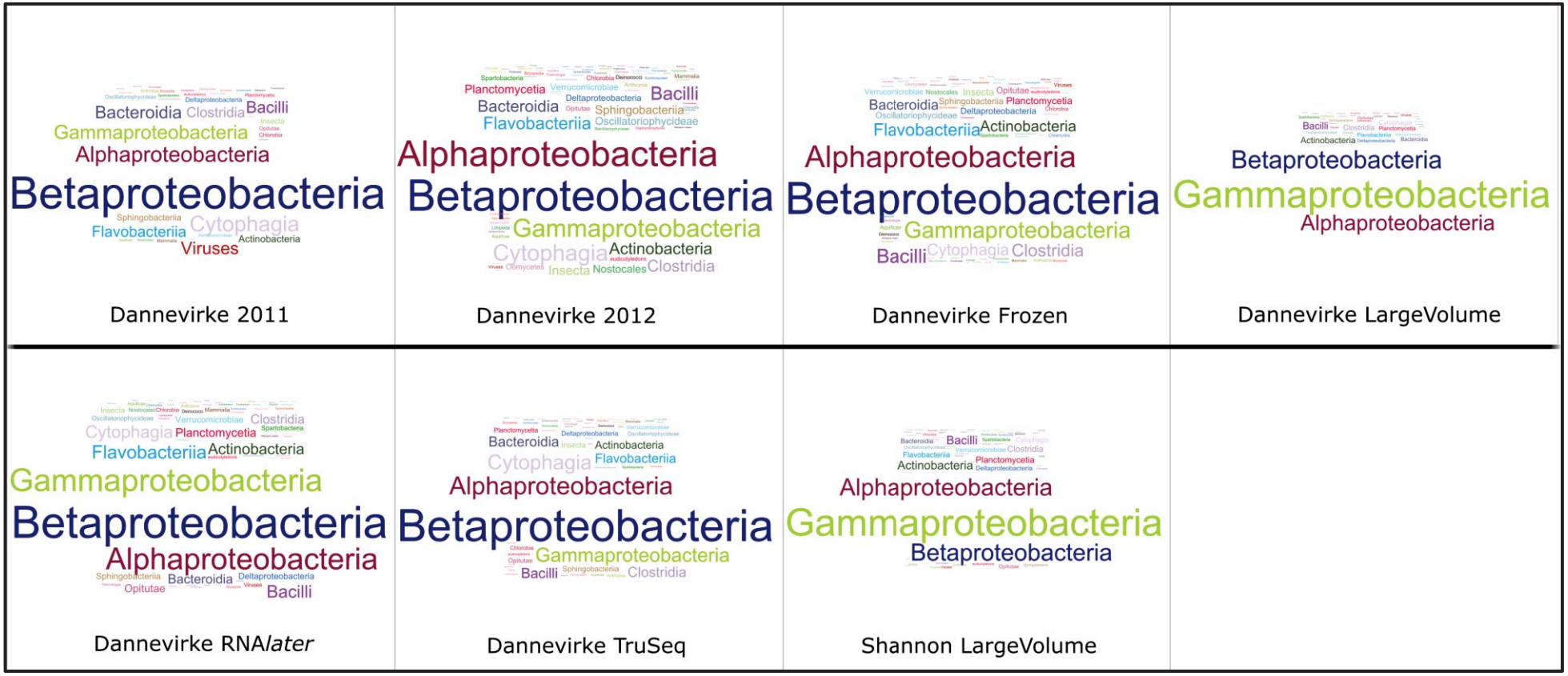


Figure 6.3: This word cloud is a MEGAN output on whole genome sequencing data. The word cloud shows predominant class with a larger size text for those classes and the seven samples processed by different methods are compared to their taxonomic class classification. The methods used are shown on the x-axis of each square. Dannevirke 2011 and 2012 are the fresh samples collected in the years 2011 and 2012. The remaining samples were samples collected in the year 2012.

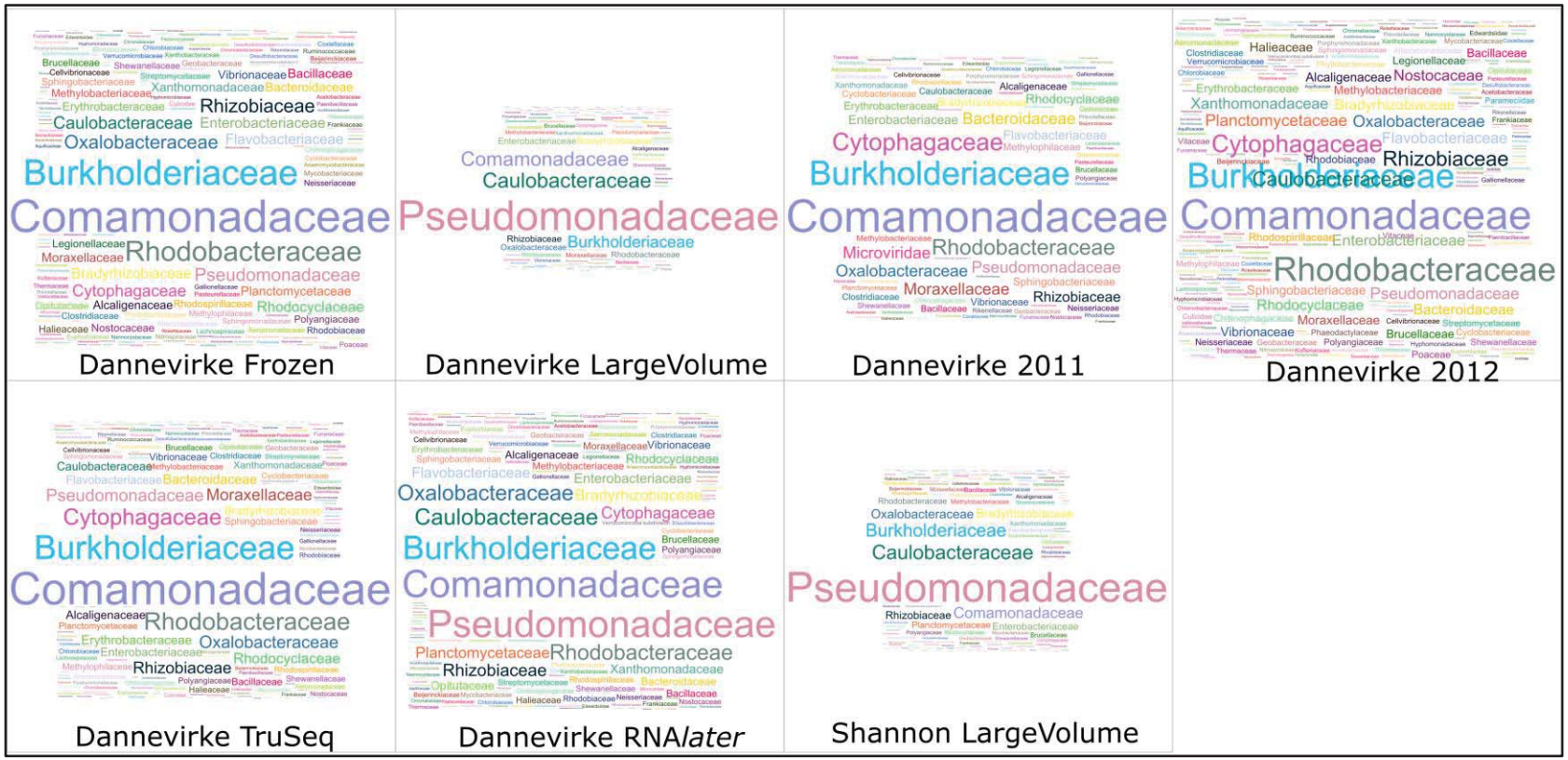


Figure 6.4: This word cloud is a MEGAN output on whole genome sequencing data. The word cloud shows the predominant family with a larger size text for those families and the seven samples processed by different methods are compared to their taxonomic family classification. The methods used are shown on the x-axis of each square. Dannevirke 2011 and 2012 are the fresh samples collected in the years 2011 and 2012. The remaining samples were samples collected in 2012.

Figure 6.5: This word cloud is a MEGAN output on whole genome sequencing data. The word cloud shows predominant genera with a larger size text for that genus, and the seven samples processed by different methods are compared to their taxonomic genus classification. The methods used are shown on the x-axis of each square. Dannevirke 2011 and 2012 are the fresh samples collected in the years 2011 and 2012. The remaining samples were samples collected in 2012.

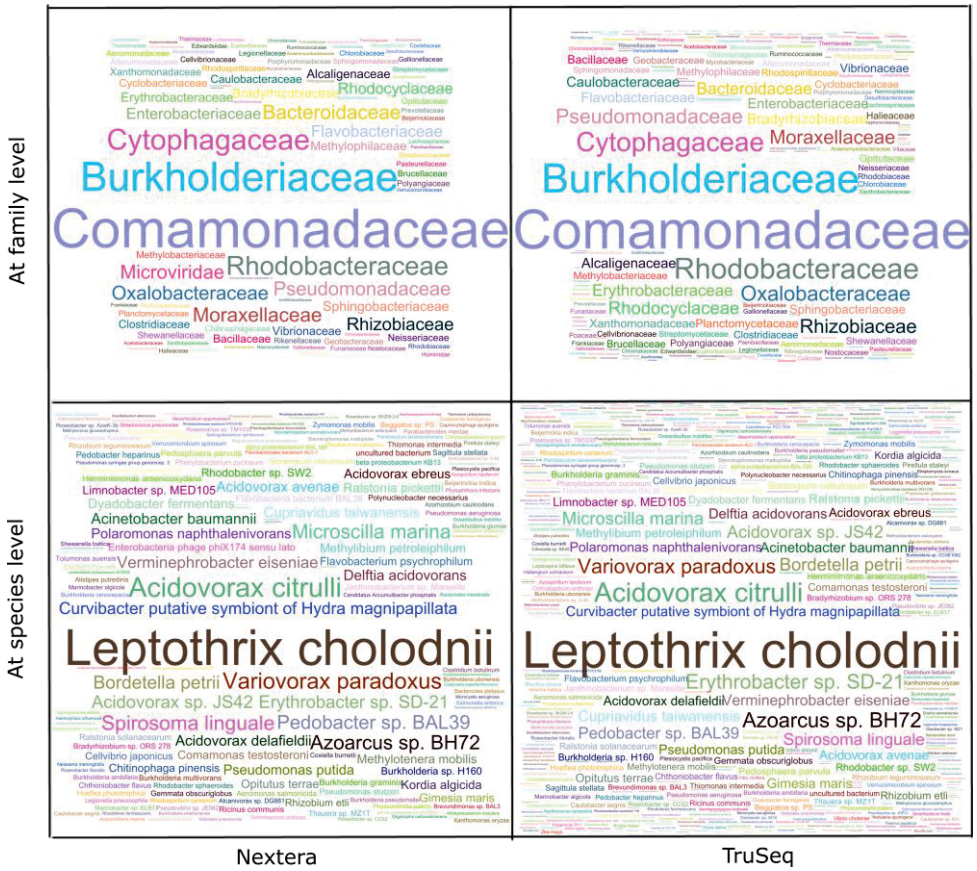


Figure 6.6: This word cloud is a MEGAN output on whole genome sequencing data. The word cloud shows microbial richness between Dannevirke samples processed with Nextera and TruSeq sequencing methods at the taxonomic family (upper) and species (lower) levels of classification. The methods used are shown in the x-axis and taxonomic classification in the y-axis.

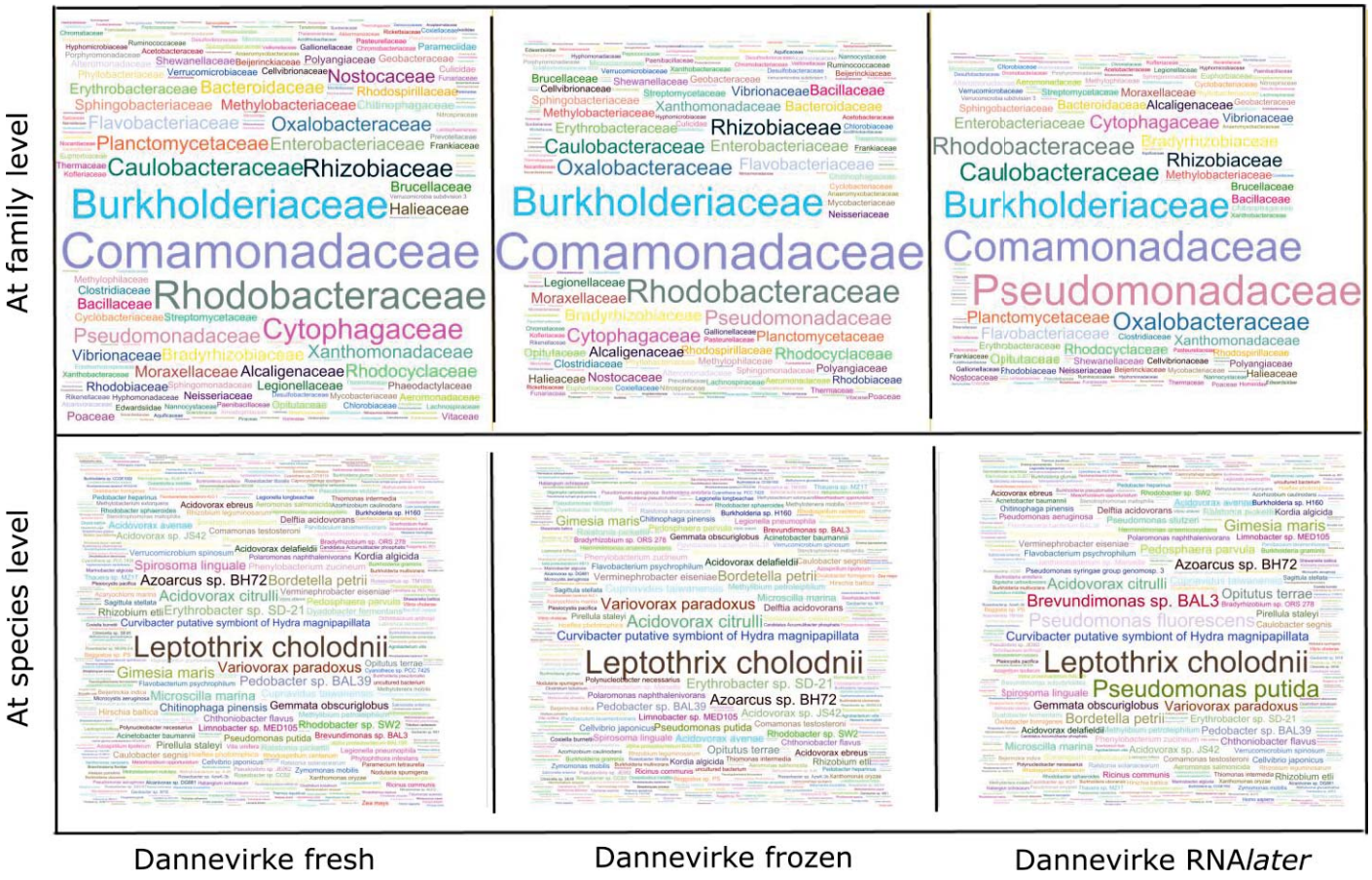


Figure 6.7: This word cloud is a MEGAN output on whole genome sequencing data. The word cloud shows microbial richness between Dannevirke samples that were processed fresh, after 3 months of freezing and of preservation in RNA*later* at the taxonomic family (upper) and species (lower) levels of classification. The methods used are shown on the x-axis and taxonomic classification in the y-axis.

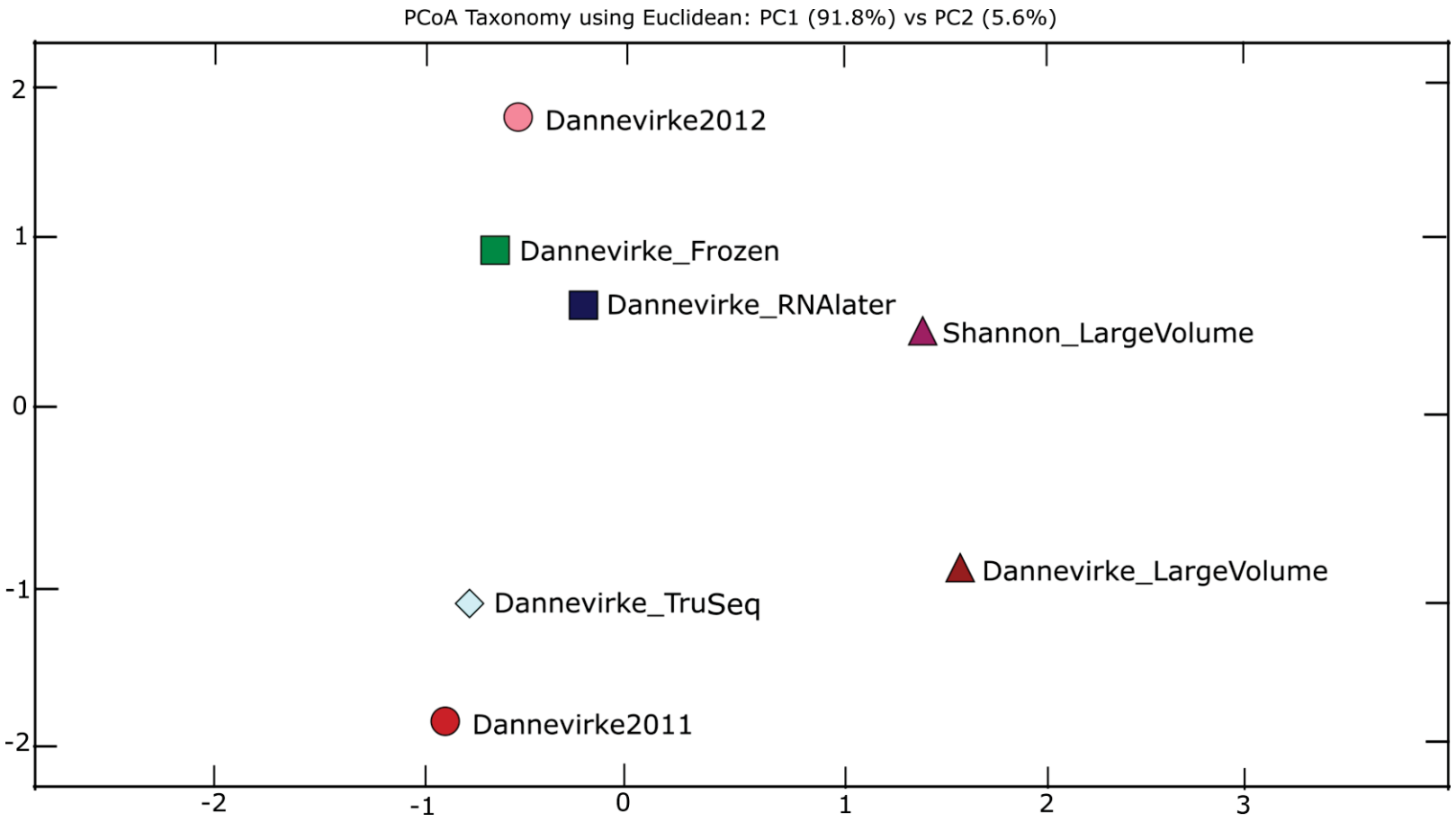


Figure 6.8: This two-dimensional principal co-ordinate analyse (PCoA) is a MEGAN output on whole genome sequencing data. The two-dimensional principal co-ordinate analyses microbial diversities between the seven different samples. The x-axis and y-axis denote principal component one and two, respectively.

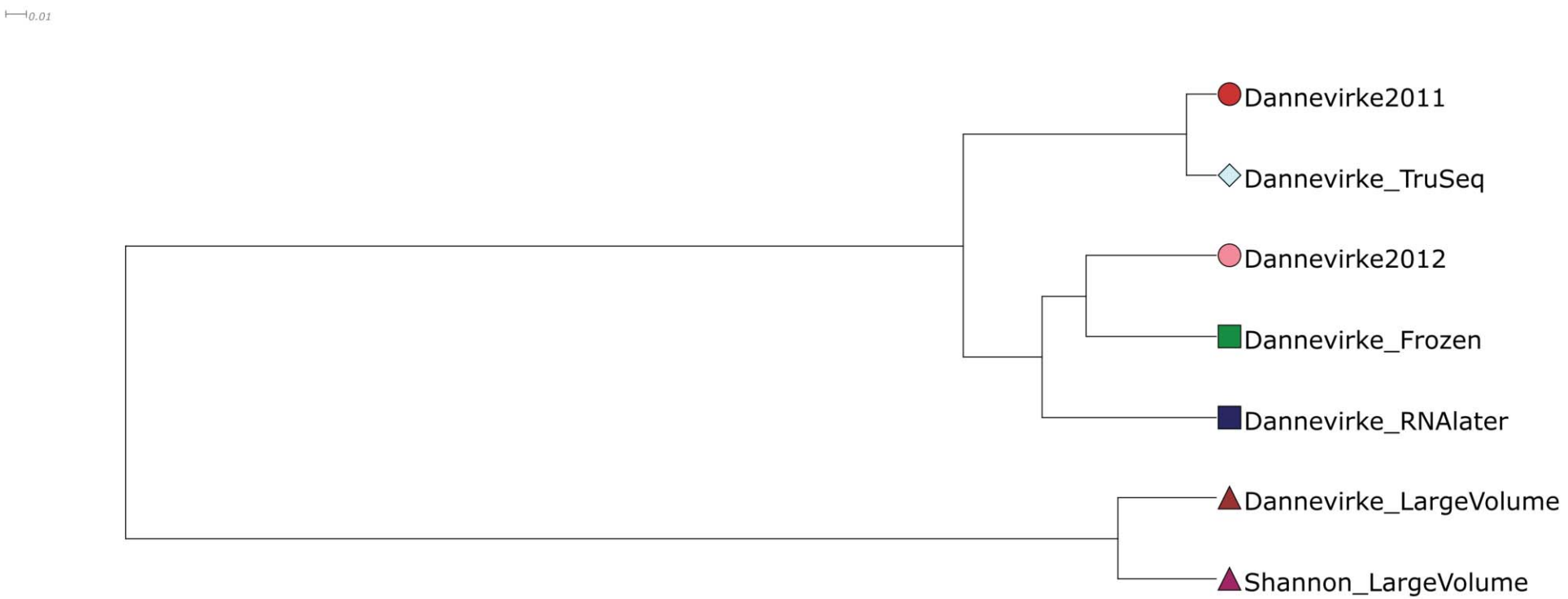


Figure 6.9: This dendogram tree is a MEGAN output on whole genome sequencing data, and shows microbial diversities between the seven different samples. Fresh samples from Dannevirke were collected in the years 2011 and 2012.

6.6 Discussion

In this study, I attempted to examine the microbial communities present in two surface water locations destined for drinking purposes by a metagenomic approach using Illumina NGS technology. To our knowledge, this is the first study of this kind performed in a rural town's drinking water source of the Manawatu-Wanganui region of New Zealand. This study endeavours to determine the feasibility of using a metagenomic approach to monitoring pathogens in drinking water sources so that the associated public health risk could be prevented. Although this study has not used control samples or analysed species sequences in-depth, it has provided basic information on the effect of different types of sampling, processing, and preservation methods used, which could lead to further detailed studies on the source water microbiome.

6.6.1 Sample analysis

It is important to use only genomic DNA that fulfils the minimum concentration of DNA requirements for NGS so that a better quality of sequences can be obtained (Illumina®). In this study, only seven samples were processed for a shotgun sequencing because the remaining one sample did not fulfil the minimum DNA concentration requirements input for sequencing. This is due to either lower quality or quantity of DNA, which are influenced by factors such as the volume of a water sample collected, the use of different pore size filters and DNA extraction kits.

According to Staley *et al.,* (2013), although both 0.45 and 0.22 µm filters are acceptable to characterise bacteria, the 0.45 µm filter may fail to capture some potentially important bacteria when compared to the 0.22 µm filter. Our study used both 0.45 and 0.22 µm filters. Therefore, it is likely that many bacteria were captured in water samples. The authors have also filtered a large volume of samples (40 L) from each site and indicated that the volume of samples will definitely influence the bacterial community (Staley et al., 2013). A study conducted at Massey University (data not shown; personal communication, Richard Fong, NZGL) suggested that a sample (say, 10 L) that is filtered through a single filter will produce a smaller amount of DNA in comparison to the divided volume of that sample (say, two 5 L) filtered into two different filters. Therefore, I divided each sample into four parts and filtered them individually to retrieve more microbial DNA. A phenol-chloroform method and/or commercial kits such as those from epicentre® (e.g. Metagenomic DNA Isolation Kit

for Water) are used to isolate high-quality DNA. With the advancement of molecular biology techniques, the commercial DNA extraction kits are more popular due to their ease of use. Nevertheless, thorough studies are required to develop a standard protocol for water sampling, retaining the microbes in water, and obtaining a good quality of DNA to sequence so that the microbial diversity in water is better understood.

6.6.2 Whole genome shotgun sequencing output

The majority of the annotated sequences were related to the bacterial domain (76.6 %), some of them were low complexity reads (13.7 %), unassigned (7.2 %), unclassified sequences (0.03 %), and the remaining annotated sequences (1.6 %) were related to eukaryotes, archaea, and viruses (Figure 6.1). Similar to other metagenomics studies such as those of the sea and hot springs, bacteria were the predominance domain in our water samples (Wu et al., 2013; Mangrola et al., 2015). The dominance of bacteria could be due to the presence of a lot of bacterial DNA in the sample and sampling environment. With improvements in the DNA extraction and sequencing technologies, there will be a continuous update in the reference genomes database for bacteria as well as other domains such as eukaryotes and virus. Therefore, with the latest reference databases, it could possibly change the percentage of microbes' presence in a sample from the same place or same sample.

A taxonomic comparison at the level of phyla showed no differences among the samples, the predominant phylum being *Proteobacteria* followed by *Bacteroidetes* and *Firmicutes*. Higher levels of *Proteobacteria* have been identified in fresh water, sewage and soil, suggesting that *Proteobacteria* has been a frequently available phylum in the water environment (Sanapareddy et al., 2009). However, the findings of abundant members of the *Bacteroidetes* and *Firmicutes* phyla indicate faecal contamination, as they have been found in mammalian faeces (Jeong et al., 2011). Therefore, faecal contamination might have occurred in the Dannevirke River and Mangaore stream during the sampling period.

In this study, a large volume of water samples was collected to determine protozoa that might have been missed in the grab samples and to compare if the large volume samples can be used for microbial diversity. Microbial richness observed was found to be greater in grab samples than in the large volume water samples. In addition, a distinct difference in microbial diversity between the large volume and grab samples was also observed when compared at the taxonomic levels of class, family and genus (Figures 6.3 to 6.5). The PCoA and UPGMA analyses also showed that grab samples

clustered separately to the large volume samples (Figure 6.8 and 6.9), implying that method of sample collection may affect in the microbial diversity in the samples. These differences might be due to the differences in the volume of water collected for grab and large volume samples (3L vs 100L). A large volume of water sampling may allow the recovering of pathogens that were missed in samples collected in smaller volumes (USEPA, 2005). Other possible reasons could be the contamination of the pump and the pipe used for collection of the large volume samples, and the multiplication of certain bacteria in water and PBS-eluted solution such as *Pseudomonas* spp. Therefore, the effect of the collection method used to determine the microbial diversity in water environment should be studied in greater detail.

Microbial diversity and richness between the large volume samples from Dannevirke and Shannon sites were compared at the taxonomic levels of class, family and genus, and similar microbes were dominantly found in the both sites. These two large volume samples were also clustered together in both PCoA and UPGMA analyses, implying that microbial diversity may not be affected by the sample's origin or sites (Figure 6.8 and 6.9). It may be possible that samples from these two water bodies are surrounded by similar environments. Nevertheless, there is a need for further investigations of microbes in the environment and water before making concrete conclusions.

Samples collected in the years 2011 and 2012 from Tamaki River, Dannevirke had similar microbial diversity in the taxonomic levels of class and family (Figures 6.3 to 6.4). Although these samples are collected from the same source and from the same sampling point, the microbial richness varied widely (Figures 6.3 to 6.4). Therefore, these two samples were also clustered separately in the PCoA and UPGMA analyses (Figures 6.8 and 6.9). The seasonal or temporal variation in finding different genera has been observed in tap water (Powell et al., 2000). The temporal variation in the microbial richness of this study may be because of variations in the water environment parameters such as pH, temperature, rainfall, runoff, water stagnation at the time of sampling (Lindstrom et al., 2005). These observations warrant further investigation before drawing conclusions on the temporal variation of microbes in those water environments.

In our study, one water sample collected from Dannevirke in the year 2011 was processed with two different shotgun sequencing methods - NexteraXT (Dannevirke 2011) and TruSeq (Dannevirke TruSeq). When PCoA and UPGMA analyses were performed, reads from these two methods were clustered together in both analyses (Figures 6.8 and 6.9), implying that there is no differences in microbial diversity

between the sequencing methods used for the same sample. However, there is an increase in microbial richness in the sample processed with TruSeq compared to NexteraXT (Figures 6.3 to 6.6), which could be due to the use of a higher amount of DNA in the TruSeq sequencing method.

Similarly, when three different samples based on processing methods (fresh, 3-months after freezing, and 3-months after preservation in RNA*later*) were compared, microbial diversity among them were found to be similar at the taxonomic level of up to family (Figures 6.3 to 6.5). At the genus level, *Pseudomonas* spp. were predominant in RNA*later* samples, unlike the other samples. In addition, microbial richness was also lower in RNA*later samples* than in fresh or frozen samples (Figure 6.7). *Pseudomonas* spp. are ubiquitous in the water environment, produce biofilms, and have potential to multiply at room temperature, and some of them are potentially zoonotic (Tortotra et al., 2012). In our study, a huge volume of water was passed through the filtration unit and the filter was eluted in PBS solution for large volume samples, whereas, samples preserved in RNA*later* is kept at room temperature for three months. These could be the reason that the large volume and RNA*later* samples may have a predominance of *Pseudomonas* spp.

This shotgun sequencing of source water for drinking purposes is a pilot study. Here we used only one of each sample for comparison of sequencing and processing methods. This study provides a snapshot of metagenomes in the water environment, but lacks some degree of validity by not including the control samples in the study. Despite useful applications of metagenomics, there are potential false positive or negative results, sampling bias or PCR-generated errors, the reliability of reference database for certain genomes, and over-interpretation of data (Thomsen and Willerslev, 2015). This study also compared microbial richness in different methods used at a taxonomic level of species and found some potentially zoonotic pathogens that are of public health risk in some of our samples. However, this study had many challenges, including rapid changes in sequence analysis software, availability of reference databases, and budget allocation for further study/analysis during the study period. As a result, an in-depth analysis of sequences of those pathogens was not performed to avoid false interpretation of zoonotic pathogens related to waterborne public health risk. It is expected that this data would be utilised by other investigators using the current sequence analysis technology for associated public health risks in the rural towns of New Zealand. Further studies with a sufficient number of water samples

and control samples are also anticipated in future for comparing the sequencing methods and preservation methods before making any concrete conclusions.

6.7 Conclusion

This NGS study provides potentially valuable baseline data on the microbial diversity present in the drinking source water for two rural towns in New Zealand. In addition, the majority of the annotated sequences obtained from water samples were assigned to the domain bacteria. Overall, *Proteobacteria* spp. were the predominant microorganisms found in the surface water samples. *Bacteroidetes* and *Firmicutes* were the second and third most dominant bacterial phyla in these water samples, implying the presence of faecal contamination in the Tamaki River and the Mangaore Stream during sample collection. There is also a significant difference in bacterial population between the different sample collection methods (grab vs. large volume), different processing methods of the samples (fresh vs. frozen vs. RNA*later* preservation), and different sequencing methods (Nextera vs. TruSeq), reinforcing the diagnostic value of sample collection, processing and sequencing methods. Microbiological diversity variation was also observed in the samples collected from the same river in different years, however, further studies are warranted before making definitive conclusions. Potentially zoonotic pathogens could also be identified through NGS. Further analysis of virulence factors and the source of the pathogens could enable us to understand the viability and faecal contamination source of the pathogens. Nevertheless, this study is a pilot study that has added the useful baseline information on the use of the NGS technology to get a snapshot of the water environment in New Zealand. The data generated in this study could be used in future to determine the potential zoonotic pathogens in those water samples.

6.8 Highlights of this study

- Of the eight samples, seven samples were successfully used for whole genome shotgun sequencing.
- Immediately (fresh) processed samples have shown a different and greater variety of microbes present than those in preserved samples (frozen and RNA*later)*.
- Samples collected from the same water environment varied between the time of year collected.
- *Proteobacteria* was the predominant microorganism found in all the samples.

- The data generated in this study could be used to study in details about the different organisms determined, and virulence genes for those organisms.

6.9 Acknowledgements

I would like to thank my co-supervisor Dr Patrick Biggs for bioinformatics analysis of WGS data, Trish McLenachan providing help in genomic DNA extraction from water samples, and Richard Fong for providing MGS information.

This project was funded by IVABS Post-graduate Research Fund and Ministry of Health Fund for Aquatic Protozoa Research Unit.

Chapter 7

General Discussion

7.1 Introduction

Globally, water-borne zoonosis poses challenges to ensuring the safety of drinking or recreational water It is, therefore, a matter of great concern for human health. New Zealand is not an exception to this. There are several emerging and re-emerging water-borne pathogens recognised worldwide. However, this thesis focused only on *Campylobacter*, *Cryptosporidium* and *Giardia* spp. because the three most commonly notified water-borne human illnesses in New Zealand are attributed to these pathogens. The studies presented in this thesis were designed with the aims of 1) examining the prevalence of and factors associated with the presence of *Campylobacter* and ruminant-associated *Campylobacter* in surface water for recreational purposes; 2) detecting the three pathogens in cattle, sheep and surface water within the drinking water catchment areas; and 3) analysing the possible use of next-generation sequencing technology for detection of these pathogens in faecal and water samples. By combining the molecular analysis, epidemiology and novel approach to detect the type or subtype of the organisms, this work has thrown additional light both on the epidemiology of New Zealand's water-borne zoonoses and on the importance of the detection methods employed.

This thesis began by reviewing the role of ruminants in waterway contamination (Chapter 1) highlighting the description of three main ruminant-associated waterborne pathogens, their survivability in the environment and modes of transmission to waterways. The pros and cons of the diagnostic methods available for water/faecal samples and possible approaches to reduce the waterway contamination were also discussed. Ruminants are considered as one of the leading causes of water contamination, but other factors such as environmental and farm

management are equally important to consider along with the diagnostic techniques used to determine the risk of pathogens.

The sources, dynamics and population structure of *Campylobacter* in high-use recreational rivers in the Manawatu-Wanganui region were studied in detail in Chapter 2. The analysis revealed differences in retrieving *Campylobacter* spp. from winter versus summer months, among the six recreational river sites and from water samples when the river flow rate is either high or low. Like in other studies (Carter et al., 2009; Jokinen et al., 2011; Bailey et al., 2015), wild bird-associated *Campylobacter* (e.g., ST-2381) were also dominant in these river water samples. Yet, public health risk cannot be ignored due to the findings of some ruminant-associated *Campylobacter* (e.g., ST-61) that were also found in human campylobacteriosis cases in New Zealand.

Regular monitoring of all the recreational or drinking source water for protozoa has not been implemented yet due to the fact that sentinel data used in the previous chapter did not cover the protozoa data. In reality, the protozoa *Cryptosporidium* and *Giardia* spp. are important water-borne pathogens to be monitored. Therefore, Chapter 3 complements the previous chapter by including the two protozoa and *Campylobacter* detection in a repeated cross-sectional study carried out during lambing and calving seasons and within the water catchment areas of two rural towns in the Manawatu region. Surprisingly, *Campylobacter* spp. were not detected in one farm from the Shannon catchment area and *Cryptosporidium* spp. were detected only in dairy calves during the first sampling. There are differences in detecting these pathogens between cattle and sheep, dairy and beef cattle, young and adult animals, and between two locations. Differences were also found in the detection of three pathogens between water samples collected from Dannevirke and Shannon catchment areas. This study revealed that ruminants on the farms could shed zoonotic potential subtypes of *Cryptosporidium* and *Giardia* spp. In this chapter, the protozoan detection in water samples was challenging because several factors, including the turbidity of water, the volume of water collected and losses of (oo)cysts during processing of samples influence the (oo)cysts detection. Therefore, the recovery rate during each sample collection from surface water should be performed to evaluate the techniques used within that laboratory.

The genotyping of *Cryptosporidium* spp. in the previous chapter showed an unexpected result of mixed infection in dairy calves. This finding was new in the New Zealand context, and it may be necessary to cross-examine the PCR-sequencing result

to see if it is due to an error, or if it, in fact, real. Thus, chapter 4 includes the study to determine the mixed *Cryptosporidium* infection using iterative multiple locus PCR-sequencing. To date, many studies reported mixed *Cryptosporidium* infection in ruminants. Nonetheless, the mixed *Cryptosporidium* species infection was confirmed successfully and is novel in terms of finding three endemic cattle *Cryptosporidium* species in one faecal sample and new genetic variants of *Cryptosporidium* in calves in New Zealand. With the advent of new technology, the molecular biology field has changed dramatically in the last five years. Next-Generation sequencing platforms are one of them, and many studies have been employing this technology to identify microbial presence in the environment and characterise them to species levels. In chapter 5 and 6 of this thesis, a modified MLST (named "massMLST") and a metagenomic study were performed using Next-Generation Sequencing platforms, respectively. *Campylobacter* isolates obtained from the chapter 3 study were utilised for massMLST. This novel approach, massMLST, could be a potential subtyping method in future because of the possibility to type many isolates at a time. A metagenomic study was performed using whole genome shotgun sequencing of the metagenomes from the water samples used in chapter 6. In this study, phylum *Proteobacteria* was determined in all the samples. Many pathogens, including *Campylobacter*, *Cryptosporidium*, and *Giardia* spp. were determined, suggesting that there was potential faecal contamination of water bodies during the sampling time. Microbial diversity was also observed between various preservation methods. However, we used only two water samples as a preliminary study. Therefore, further research work is warranted to compare between the use of fresh samples or of different preservation methods on the microbial diversity and on the detailed analysis of sequences of faecal pathogens identified. Nevertheless, in future, a decrease in NGS cost and an improvement on metagenomes extraction methods will possibly help the water industry to use shotgun sequencing for the analysis of pathogens in water sources for drinking or recreational activity

This thesis uses retrospective data from surveillance, as well as repeated cross-sectional studies, and employed novel molecular methods to understand the transmission dynamics of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. in New Zealand. During my PhD studies, a number of challenges were encountered during the employment of study methodologies. In addition, the development of new sophisticated technology led to new molecular analysis methods. It is, therefore, relevant to discuss data collection, molecular epidemiology and future perspectives in this concluding chapter.

7.2 Molecular epidemiology and disease surveillance

Nowadays, molecular epidemiology is a thriving field of research due to the recent advancement in molecular biology techniques, and also due to the possibility of integrating these techniques with the traditional epidemiological approach as well as in disease surveillance. This has led to a dramatic increase in the number of studies of the molecular epidemiology of infectious diseases, particularly zoonotic, during the last five to seven years. Some researchers have also concluded that the increasing availability of molecular tools, new diagnostic tools, and their application beyond the organisms' species level have changed the way epidemiological studies are done, in the field of-infectious diseases (Loomis and Wing, 1990; Foxman and Riley, 2001; Mullner et al., 2009; 2013). Although conventional techniques cannot be completely substituted by molecular techniques, these researchers recommend the practical application of new molecular tools in epidemiological studies, particularly concerning the interface between clinicians, epidemiologists, microbiologists, molecular biologists, statisticians, bioinformaticians, and computational biologists. This integration of researchers and technicians from various fields is reflected in the chapters of this thesis. In future, such interfaces between researchers from different disciplines are more likely because of the introduction of novel molecular technology that may produce huge but more complex information on diseases beyond the subspecies level.

According to the WHO, surveillance is "the ongoing systematic collection, analysis, interpretation, and dissemination of the data". Surveillance studies provide crucial information on trends in pathogen incidence and identify the emergence of pathogens at national and global levels. Therefore, routine surveillance is critical to enable the development of control approaches. In New Zealand, a sentinel surveillance site (the Manawatu region) was established in 2005 to monitor human campylobacteriosis cases and to identify potential sources of *Campylobacter* infection. The samples from cases and from potential sources were collected and typed using multilocus sequence typing (MLST) simultaneously over ten years. The sentinel data have been used to examine the epidemiology of human campylobacteriosis, and to develop the source attribution model in order to estimate the contribution of different sources to human campylobacteriosis cases *(*Mullner et al., 2008; 2009; 2010; Marshall et al., 2013; French et al., 2014).

In addition, six high-use recreational surface waters within the Manawatu-Wanganui region have been monitored for more than three years as part of the Manawatu sentinel surveillance study. I am fortunate to be able to use this l data in chapter 2 of this thesis to provide better insights into the sources of *Campylobacter* spp. in recreational water. This sentinel data contained additional information on sampling dates, sites, river flow rates and laboratory analysis data (culture, PCR and MLST) that had led to a finding that *Campylobacter* spp. were identified more frequently in the winter months and during periods of increased river flow rates. , It also revealed that that the rate differed with sampling sites, and that it originated from wild birds and ruminants. Such information is highly useful to public health officers and policy makers in determining the role played by recreational water in spreading human campylobacteriosis. It is also useful in increasing public awareness of these pathogens, and in planning effectively in order to reduce the campylobacteriosis burden in New Zealand. This recreational water surveillance was discontinued, however, due to a number of reasons including the high cost of water analysis. Consequently, there is limited data available on *Campylobacter* spp. in water in the current Manawatu sentinel surveillance database. Despite these limitations, the data has, however, meant that is it is possible to understand the occurrence of *Campylobacter* spp. in various hosts, as well as its sources and its diversity. It has also enabled us to apprehend risk factors and source attribution in human campylobacteriosis, as well as mutation and recombination within *Campylobacter* spp. Finally, it has facilitated a comprehension of the evolution of *Campylobacter* spp. in New Zealand (French et al., 2009; Mullner et al., 2009; 2010; Biggs et al., 2011; Mohan et al., 2013; Binney et al., 2014; Marshall et al., 2014; this thesis). These authors also showed a shift in the sources of campylobacteriosis, a decrease in poultry-associated cases, and an increase in ruminant-associated cases. Moreover, they demonstrated the dominance of the two poultry-associated sequence types of *C. jejuni* (ST-474 and ST-48) and the dominance of ST-2381 that were particularly originated from surface water and wild birds in New Zealand. Such changes could be expected in water environments as well due to climate change and change in land use. Therefore, in future, it is expected that recreational water surveillance will include again using the latest molecular tools such as whole genome sequencing to unravel the complex transmission dynamics of *Campylobacter* spp. in New Zealand.

In New Zealand, the trend of molecular epidemiological studies of *Cryptosporidium* and *Giardia* species in various sources has been increasing for more than a decade (Al Mawly et al., 2015b; Grinberg et al., 2005; 2008; Learmonth et al., 2004; Houque et al.,

2002; Winkworth et al., 2010; Simmons et al., 2001; Chilvers et al., 1998; Chapter 3 study). These studies have examined on occurrences, concentrations, types and subtypes of these two protozoa in the faeces of humans and animals, on pasture, and in bodies of water. The Ministry of Health has implemented aquatic protozoa research projects in collaboration with a molecular epidemiology laboratory at IVABS, Massey University. These projects have been monitoring, screening and characterising aquatic protozoa *Cryptosporidium* and *Giardia,* and bacteria *Campylobacter* species in the water supplies of 20 cities of New Zealand, in animal faeces, and in human stools for more than five years for the Ministry of Health. This research unit also has a focus on developing a database archive and genotyping library of *Cryptosporidium* and *Giardia* strains, and on assessing the differences in infectivity between *Cryptosporidium* sub-genotypes found in New Zealand. The annual reports from the Aquatic Protozoa Research Unit have shown that for *Campylobacter, Cryptosporidium* and *Giardia* species there is a prevalence of between 25% and 60% for those 20 water supplies that have advanced water treatment plants (such as UV treatment). Water supplies for small towns and rural areas have not, however, been monitored in New Zealand, and it is known that water treatment plants in small towns and rural areas are not advanced and that this can potentially lead to an outbreak of water-borne diseases in those areas. Therefore, two surface water supplies, which supply drinking water to the rural towns of Dannevirke and Shannon in the Manawatu-Wanganui region, were monitored for four months to identify and characterise microbes, including *Campylobacter, Cryptosporidium* and *Giardia* species, using a range of molecular biological techniques, conventional (PCR-sequencing) to advance (next-generation sequencing) molecular biological techniques (Chapters 3-6).

The molecular epidemiological survey in this thesis is a preliminary study on environmental epidemiology that has provided an large amount of information about the zoonotic potential of *Campylobacter* spp., as well as mixed and new *Cryptosporidium* spp. It has also provided information on *Giardia duodenalis* assemblages circulating in the rural farm and water environment. These data were unique and had not been reported previously in New Zealand. In addition, a novel method "massMLST" and a metagenomic study showed potential for analysing a large number of isolates at one time and for revolutionising testing of water quality. This study, therefore, provides a basis for which it would be possible for New Zealand to develop surface-water surveillance programmes at a national level, in order to

cultivate a better understanding of the transmission dynamics of waterborne diseases, and to develop effective control strategies.

7.3 Lessons learnt and future direction

A number of challenges were encountered during the development and implementation of this thesis study. For example, there were missing data about river flow rate for the sites of two sources- Tokomaru River and Kaikokopu Stream (Chapter 2). The river flow data were generated from an automated river flow recording system present only in the 4/6 rivers monitored. The availability of such data could provide more information about those two areas of surface water as well. During the development and planning of chapter 3, a detailed epidemiological approach was implemented to collect water and faecal samples from Dannevirke and Shannon catchment areas in the Manawatu-Wanganui region of New Zealand. However, it was realised that access to streams, sampling near paddocks and sampling from young animals was not always possible. Access to these sources was frequently obstructed owing to flooding and to non-availability of the animals on the paddock. Consequently, it was necessary to collect water samples from the tap that was linked to untreated water in the Shannon catchment area and to collect faecal samples from paddocks near waterways, regardless of the age of the animal. These methods were used so that the main objective of the study could be achieved.

During molecular analysis, it was realised that repeated analysis of the samples was required in order to extract the DNA of relevant pathogens. Repetition of this type is possible for protozoa, as these pathogens can survive for a long time in faeces. However, for bacteria like *Campylobacter,* culture from fresh samples kept for more than few hours could give negative results because this bacterium survives only for few hours in the in-vitro environment. In addition, there is a possibility of slow or poor growth of *Campylobacter* on mCCDA agar kept for 48 hours. Therefore, samples should be processed immediately for *Campylobacter* detection and, if no growth appears, the plate should be cultured for another 24 hrs. If the volume of faeces available is very low, faecal swabs in a transport medium such as Amies charcoal transport medium could be an alternate choice for identification of *Campylobacter* in faeces, although the medium is more expensive than a container (OIE, 2004).

Cryptosporidium and *Giardia* (oo)cysts have been identified in both animal faeces and human stools using direct microscopy, direct immunofluorescence assay (IFA),

enzyme immunoassay, immunochromatographic lateral-flow 'dipstick' test, PCR, RT-PCR, or PCR-RFLP. However, there is not much success in the development of a high sensitivity method for the detection of *Cryptosporidium* and *Giardia* (oo)cysts in water samples. The presence of (oo)cysts in raw or treated water could be low, thus requiring the collection of large volumes of samples, especially if the aim of the collection is to assess water for treatment. Therefore, the standard United States Environmental Protection Agency (USEPA) method has been recommended for water industries. Currently, the USEPA method 1623 is widely used in water industries worldwide, but the recovery rate of *Cryptosporidium* oocysts and *Giardia* cysts in surface water is comparatively low and varies between different laboratories.

Initially, reagent water was used for matrix spiking to determine the recovery rate of (oo) cysts in a study carried out for the Chapter 3 study. This resulted in a recovery rate of 26% and 36 % of *Cryptosporidium* and *Giardia*, respectively. The recovery rate for raw or source water has not been performed as a matter of course in the molecular epidemiology laboratory (mEpiLab) at Massey University but it was conducted in this study after receiving expert advice. As only four samples were collected, Bayesian analysis was performed to determine the recovery rate of *Cryptosporidium* and Giardia. This resulted in a recovery rate of 7.6 and 11.7 % for *Cryptosporidium* and between 13 and 17.3 % *Giardia* (Chapter 3 study). Recently, the USEPA method 1623 has been modified and is referred to as the USEPA method 1623.1. The USEPA 1623.1 method recommends determining the recovery rate each time that the sample is collected so that the actual recovery rate can be estimated (USEPA, 2013). Therefore, it is expected that projects or studies that may involve protozoa detection and surveillance in water of New Zealand will also implement determining the recovery rate at each sampling time in future, although the cost for water sampling and recovery rate determination is still too expensive for many laboratories, and requires huge funds for research purpose.

Although next-generation sequencing has revolutionised disease identification and mechanisms, it was realised that there is a need for improvement in the metagenomic DNA extraction method and data management. In my study, an Epicentre® kit was used for the extraction of DNA, which often results in low DNA quantity. There have also been reports that NZGL is ineffective in extracting high quality DNA, and this requires further research. The NGS resulted in a huge number of sequences, and this was challenging for data management and analysis. During my study period, limited data management and analysis software were available in a situation that required

advanced software and computer skills. As data analysis was in its infancy, my study in Chapter 6 did not include a detailed analysis of the various zoonotic pathogens and their virulence. These data contain valuable information about microbes in a water environment, and I expect that researchers will utilise this data with the latest data analysis techniques in future.

Multilocus sequence typing (MLST) of seven housekeeping genes has been a commonly used subtyping method for *Campylobacter* spp. for more than a decade (Dingle et al., 2005). MLST uses the Sanger sequencing method to sequence ~450 bp of the internal fragments of each gene. These are unambiguously characterised by a series of seven integers to determine the allele numbers for each gene and to sequence types for each isolate *(*http://pubmlst.org/general.shtml*)*. However, this Sanger sequencing MLST provides limited information based on that sample, and ignores the number of nucleotide differences between the alleles. In addition, this method is time-consuming due to the need to sequence a large number of isolates to assign sequence types. Therefore, MLST using Illumina technology of the NGS method was developed (massMLST), and is presented in chapter 5 of this thesis. Although the massMLST method is successful in characterising clonal complexes and/or sequence types, there were still many challenges to consider during this study. The MiSeq was only able to sequence up to 250 bp, for example, so it was challenging to develop nested primers for each fragment of each gene. Consequently, there were an increasing number of errors during 1) the amplification of these primers resulted from the need to pipette a small volume of samples, 2) the alignment of the resultant four short reads for each gene to a single read using appropriate software, and 3) the assignation of the allele numbers for each gene. Nevertheless, these challenges might be overcome if pipetting is automated and the capacity of the MiSeq machine is increased to sequence ~ 550 bp. These developments would be enhanced by the development of good alignment software.

Recently, whole genome sequencing of *Campylobacter*, *Cryptosporidium* and *Giardia* spp. has increased rapidly (Biggs et. al., 2011; Cody et. al., 2013; Paparini et. al., 2015; Hadfield et. al., 2015; Kovanen et. al., 2014a; 2014b; Hanevik et. al., 2015; Prystajecky et. al., 2015a; 2015b). These studies have been conducted by sequencing an individual organism using next generation and/or third generation sequencing technologies to gain insights into the epidemiology of the three pathogens. Using this technology, researchers have been able to interpret the extensive data, and this provided a high resolution for determining *Campylobacter*, *Cryptosporidium* and

Giardia spp. diversity at low cost. However, despite many opportunities, a number of challenges remain. These include, for example, cleaning of data, correct assembly of short reads, and other Bioinformatics analysis challenges, particularly for the metagenomic data where there are huge numbers of reads. The abovementioned shortcomings of new technologies explain why Sanger sequencing has been the preferred method for many laboratories for the last 30 years. Nevertheless, in future, next/third generation sequencing may replace Sanger sequencing even in routine clinical diagnosis due to its increasing affordability and throughput, as well as its fast turnaround. In addition, it offers the possibility of obtaining similar results to Sanger sequencing, and it is capable of providing information on the genome with high resolution (Schuster, 2008; Claustres, 2015; Yamamota et. al., 2015).

It is hoped that the information examined in this thesis will provide a basis for further exploration, especially following consideration of the challenges encountered during the study. This, in turn, could form the basis of a surface-water surveillance programme at a national level, which could be developed as part of an effective control strategy for reducing the burden of water-borne diseases in New Zealand.

7.4 Conclusion

The molecular epidemiological studies conducted for this study have provided additional knowledge on the transmission dynamics of the three major water-borne zoonotic diseases: campylobacteriosis, cryptosporidiosis and giardiasis in New Zealand. It is known that ruminants can shed *Campylobacter*, *Cryptosporidium*, and *Giardia* in the environment, thereby possibly contaminating the surface water. This study has provided new data on the zoonotic potential of *Giardia duodenalis* and its assemblages, and strains of *Campylobacter* and *Cryptosporidium* circulating on the farm and the water environment within catchment areas, which could be used for future source attribution and genetic diversity studies. A novel, possibly cost-effective molecular characterisation method "massMLST" has been developed and has assessed the use of next generation sequencing technology in determining the ecology of surface water used for drinking purposes. Considering the dramatic reduction in the cost of next-generation sequencing technologies, it is expected that massMLST will be useful in characterising a huge number of isolates. In addition, this technology could be further used by the water industry to monitor the ecology of the water environment, and for understanding mixed infection (Chapter 4) and genetic diversity within isolates of *Cryptosporidium* and *Giardia*.

Although many challenges were encountered during this PhD project, the findings of this thesis have provided scientific evidence of the existence of three potentially zoonotic pathogens on farms or in water environments, which will be the basis for developing control strategies for water -borne diseases in New Zealand, particularly in rural areas.

Appendix A

Supplementary materials for Chapter 1

Table A.1: Lists of major *Cryptosporidium* species and genotypes that are public health importance, and their major host (Source: Fayer and Xiao, 2008; Chalmers and Davies, 2010; Slapeta, 2013).

Species name	Genotype designation	Public health importance	Hosts	Identified in Cattle
C. hominis	Human (I) genotype	Major (sporadic, outbreaks)	Humans	Yes
C. parvum	Bovine (II) genotype	Major (sporadic, outbreaks)	Humans, mammals	Yes
C. meleagridis	-	Moderate (sporadic)	Homoeo-thermic birds; mammals	(Yes)
C. cuniculus	Rabbit genotype	Moderate (sporadic, outbreaks)	Rabbit and Humans	-
C. felis	Cat genotype	Moderate (sporadic)	Cat	(Yes)
C. viatorum	-	Moderate (sporadic)	Various mammals	-
C. muris	C. muris B genotype	Minor (rarely)	Rodents	-
C. tyzzeri	Mouse I genotype	Minor (rarely)	Mice	-
C. andersoni	C. muris A genotype	Minor (rarely)	Cattle	Yes
C. suis	Pig genotype II	Minor (rarely)	Pig	Yes
C. fayeri	Marsupial genotype I	Minor (rarely)	Marsupial	-
C. scrofarum	Pig genotype II	Minor (rarely)	Pig	(Yes)
C. canis	Dog genotype	Minor (sporadic)	Dog	(Yes)
C. ubiquitum	Deer genotype	Minor (sporadic)	Deer	Yes
C. viatorum	-	Minor (sporadic)	Humans	Yes
C. bovis	Bovine B genotype	Minor (sporadic)	Cattle	Yes

Table A.2: Prevalence of *Campylobacter* species found in different age groups of cattle and sheep from selected studies across the world.

Species/Location	Age	Identification method used	Animal prevalence	Farm/herd prevalence	References
Dairy cattle					
Austria	Preweaned calves	enrichment culture (Bolton broth, mCCDA)	14.9% (57/382)	33% (33/100)	Klein *et al.*, 2013[1]
Lithunia	1-3 months 3-12 months 13-24 months	enrichment culture (Bolton broth, mCCDA)	86.5% (51/59) 86.2% (70/80) 60.6% (36/60)	100% (3/3)	Ramonaite *et al.*, 2013
USA	Adult cows	Campy-Cefex plates	51.2% (735/1435)	97.9% (94/95)	Englen *et al.*, 2007
New Zealand	<21 days	enrichment culture (Bolton broth, mCCDA)	36% (58/161)	75% (18/24)	Grinberg *et al.*, 2005
Spain	-	enrichment culture (Preston broth, Campylosel plates)	66.7% (64/96) [2]	67.1% (55/82)	Oporto *et al.*, 2007
France	< 6 months > 6 years old	direct plating and culture (Karmali agar)	39.1% (292/747) 4.6% (35/754)	- -	Chatre *et al.*, 2009
Beef cattle					
Spain	NS[2]	enrichment culture (Preston broth, Campylosel plates)	5.4% (3/56) [2]	58.9% (73/124)	Oporto *et al.*, 2007
UK		direct plating and culture (Nutrient broth, mCCDA)	21.9% (74/338)	1-farm	Sproston *et al.*, 2011
Egypt	Healthy cattle	enrichment culture (Bolton broth, mCCDA)	34% (17/50)	-	Khalifa *et al.*, 2013
France	6-18 months	direct plating and culture (Karmali agar)	6.0% (45/754)	-	Chatre *et al.*, 2009
Sheep					
Spain	-	enrichment culture (Preston broth, Campylosel plates)	8.8% (17/194) [2]	55% (66/120)	Oporto *et al.*, 2007
UK	-	direct plating and culture (Nutrient broth, mCCDA)	14.0% (30/214)	1-farm	Sproston *et al.*, 2011
Egypt	-	direct plating and culture (mCCDA)	10.6% (33/310)	100% (3/3)	Chanyalew *et al.*, 2013
Egypt	-	enrichment culture (Bolton broth, mCCDA)	38% (27/71)	-	Kassa *et al.*, 2007
UK	-	direct plating and culture (Nutrient broth, mCCDA)	29.3% (123/420)	-	Stanley et al., 1998
US	lambs	enrichment culture (Bolton broth, mCCDA)	12.4% (44/356)	-	Pao *et al.*, 2014
New Zealand	<12 months >12 months	enrichment culture (m-Exeter agar)	80.9% (85/105) 30% (66/220)	gut content 1-farm	Moriarty *et al.*, 2011

Table A.3: Lists of recognised *Giardia* species, assemblages and sub-assemblages (Source: Feng and Xiao, 2011;Luzan and Svard, 2011; Plutzer et al., 2013)

Species name	Assemblages	Subassemblages	Hosts	Public health importance	Discovered by/ proposed name
G. agilis	-	-	Amphibians	No	Kunstler, 1882
G. ardeae	-	-	Birds	No	Noller, 1920
G. microti	-	-	Muskrats and voles	No	Benson, 1908
G. muris	-	-	Rodents	No	Benson, 1908
G. psittaci	-	-	Birds	No	Erlandsen and Bemrick, 1987
G. varani	-	-	Lizards	No	Lavier, 1923[a]
G. duodenalis (intestinalis, lamblia)			Mammals	Yes	Davaine, 1875
	A	AI, AII, AIII	Humans, nonhuman primates, domestic and wild ruminants, alpacas, pigs, horses, domestic and wild canines, cats, ferrets, rodents, marsupials, other mammals	Yes	*G. duodenalis* sensu stricto?[b]
	B	BI, BII, BIII, BIV	Humans, nonhuman primates, cattle, dogs, horses, rabbits, beavers, muskrats	Yes	*G. enterica*?[b]
	C		Domestic and wild canines	No	*G. canis*?[b]
	D		Domestic and wild canines	No	*G. canis*?[b]
	E		Domestic ruminants, pigs	No	*G. bovis*?[b]
	F		Cats	No	*G. cati*?[b]
	G		Mice, rats	No	*G. simondi*?[b]
	H		Seals	No	

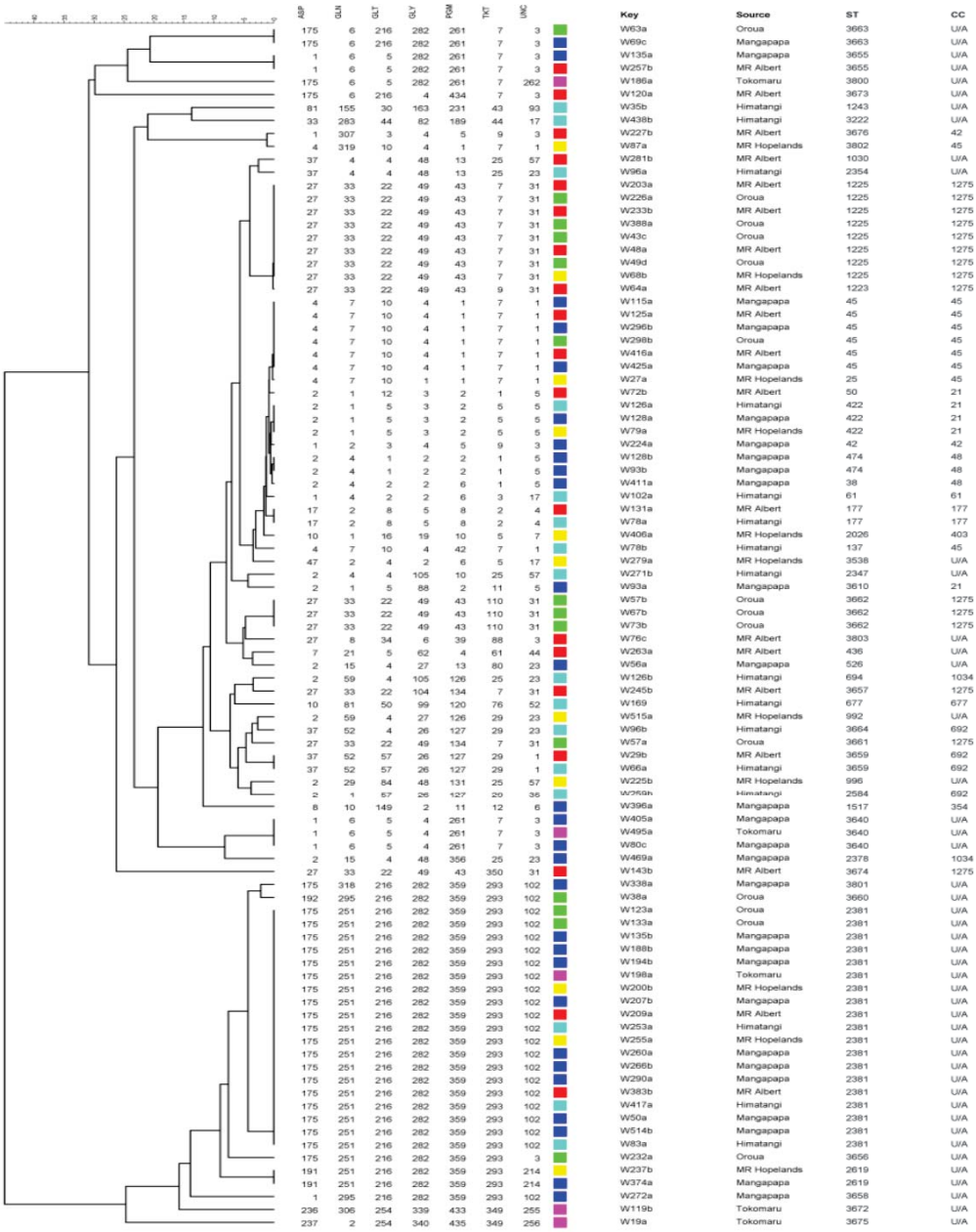
[a]Molecular characterisation required

[b]Species name proposed

Appendix B

Supplementary materials for Chapter 2

Table B.1: UPGMA tree of the *Campylobacter* sequence types detected in four recreational waters (represented by colours) of New Zealand, created using a BioNumerics software version 5.6.



ASP	GLN	GLT	GLY	PGM	TKT	UNC	Key	Source	ST	CC
175	6	216	282	261	7	3	W63a	Oroua	3663	U/A
175	6	216	282	261	7	3	W69c	Mangapapa	3663	U/A
1	6	5	282	261	7	3	W135a	Mangapapa	3655	U/A
1	6	5	282	261	7	3	W257b	MR Albert	3655	U/A
175	6	5	282	261	7	262	W186a	Tokomaru	3800	U/A
175	6	216	4	434	7	3	W120a	MR Albert	3673	U/A
81	155	30	163	231	43	93	W35b	Himatangi	1243	U/A
33	283	44	82	189	44	17	W438b	Himatangi	3222	U/A
1	307	3	4	5	9	3	W227b	MR Albert	3676	42
4	319	10	4	1	7	1	W87a	MR Hopelands	3802	45
37	4	4	48	13	25	57	W281b	MR Albert	1030	U/A
37	4	4	48	13	25	23	W96a	Himatangi	2354	U/A
27	33	22	49	43	7	31	W203a	MR Albert	1225	1275
27	33	22	49	43	7	31	W226a	Oroua	1225	1275
27	33	22	49	43	7	31	W233b	MR Albert	1225	1275
27	33	22	49	43	7	31	W388a	Oroua	1225	1275
27	33	22	49	43	7	31	W43c	Oroua	1225	1275
27	33	22	49	43	7	31	W48a	MR Albert	1225	1275
27	33	22	49	43	7	31	W49d	Oroua	1225	1275
27	33	22	49	43	7	31	W68b	MR Hopelands	1225	1275
27	33	22	49	43	9	31	W64a	MR Albert	1223	1275
4	7	10	4	1	7	1	W115a	Mangapapa	45	45
4	7	10	4	1	7	1	W125a	MR Albert	45	45
4	7	10	4	1	7	1	W296b	Mangapapa	45	45
4	7	10	4	1	7	1	W298b	Oroua	45	45
4	7	10	4	1	7	1	W416a	MR Albert	45	45
4	7	10	4	1	7	1	W425a	Mangapapa	45	45
4	7	10	1	1	7	1	W27a	MR Hopelands	25	45
2	1	12	3	2	1	5	W72b	MR Albert	50	21
2	1	5	3	2	5	5	W126a	Himatangi	422	21
2	1	5	3	2	5	5	W128a	Mangapapa	422	21
2	1	5	3	2	5	5	W79a	MR Hopelands	422	21
1	2	3	4	5	9	3	W224a	Mangapapa	42	42
2	4	1	2	2	1	5	W128b	Mangapapa	474	48
2	4	1	2	2	1	5	W93b	Mangapapa	474	48
2	4	2	2	6	1	5	W411a	Mangapapa	38	48
1	4	2	2	6	3	17	W102a	Himatangi	61	61
17	2	8	5	8	2	4	W131a	MR Albert	177	177
17	2	8	5	8	2	4	W78a	Himatangi	177	177
10	1	16	19	10	5	7	W406a	MR Hopelands	2026	403
4	7	10	4	42	7	1	W78b	Himatangi	137	45
47	2	4	2	6	5	17	W279a	MR Hopelands	3538	U/A
2	4	4	105	10	25	57	W271b	Himatangi	2347	U/A
2	1	5	88	2	11	5	W93a	Mangapapa	3610	21
27	33	22	49	43	110	31	W57b	Oroua	3662	1275
27	33	22	49	43	110	31	W67b	Oroua	3662	1275
27	33	22	49	43	110	31	W73b	Oroua	3662	1275
27	8	34	6	39	88	3	W76c	MR Albert	3803	U/A
7	21	5	62	4	61	44	W263a	MR Albert	436	U/A
2	15	4	27	13	80	23	W56a	Mangapapa	526	U/A
2	59	4	105	126	25	23	W126b	Himatangi	694	1034
27	33	22	104	134	7	31	W245b	MR Albert	3657	1275
10	81	50	99	120	76	52	W169	Himatangi	677	677
2	59	4	27	126	29	23	W515a	MR Hopelands	992	U/A
37	52	4	26	127	29	23	W96b	Himatangi	3664	692
27	33	22	49	134	7	31	W57a	Oroua	3661	1275
37	52	57	26	127	29	1	W29b	MR Albert	3659	692
37	52	57	26	127	29	1	W66a	Himatangi	3659	692
2	29	84	48	131	25	57	W225b	MR Hopelands	996	U/A
2	1	57	26	127	29	36	W259b	Himatangi	2584	692
8	10	149	2	11	12	6	W396a	Mangapapa	1517	354
1	6	5	4	261	7	3	W405a	Mangapapa	3640	U/A
1	6	5	4	261	7	3	W495a	Tokomaru	3640	U/A
1	6	5	4	261	7	3	W80c	Mangapapa	3640	U/A
2	15	4	48	356	25	23	W469a	Mangapapa	2378	1034
27	33	22	49	43	350	31	W143b	MR Albert	3674	1275
175	318	216	282	359	293	102	W338a	Mangapapa	3801	U/A
192	295	216	282	359	293	102	W38a	Oroua	3660	U/A
175	251	216	282	359	293	102	W123a	Oroua	2381	U/A
175	251	216	282	359	293	102	W133a	Oroua	2381	U/A
175	251	216	282	359	293	102	W135b	Mangapapa	2381	U/A
175	251	216	282	359	293	102	W188b	Mangapapa	2381	U/A
175	251	216	282	359	293	102	W194b	Mangapapa	2381	U/A
175	251	216	282	359	293	102	W198a	Tokomaru	2381	U/A
175	251	216	282	359	293	102	W200b	MR Hopelands	2381	U/A
175	251	216	282	359	293	102	W207b	Mangapapa	2381	U/A
175	251	216	282	359	293	102	W209a	MR Albert	2381	U/A
175	251	216	282	359	293	102	W253a	Himatangi	2381	U/A
175	251	216	282	359	293	102	W255a	MR Hopelands	2381	U/A
175	251	216	282	359	293	102	W260a	Mangapapa	2381	U/A
175	251	216	282	359	293	102	W266b	Mangapapa	2381	U/A
175	251	216	282	359	293	102	W290a	Mangapapa	2381	U/A
175	251	216	282	359	293	102	W383b	MR Albert	2381	U/A
175	251	216	282	359	293	102	W417a	Himatangi	2381	U/A
175	251	216	282	359	293	102	W50a	Mangapapa	2381	U/A
175	251	216	282	359	293	102	W514b	Mangapapa	2381	U/A
175	251	216	282	359	293	102	W83a	Himatangi	2381	U/A
175	251	216	282	359	293	3	W232a	Oroua	3656	U/A
191	251	216	282	359	293	214	W237b	MR Hopelands	2619	U/A
191	251	216	282	359	293	214	W374a	Mangapapa	2619	U/A
1	295	216	282	359	293	102	W272a	Mangapapa	3658	U/A
236	306	254	339	433	349	255	W119b	Tokomaru	3672	U/A
237	2	254	340	435	349	256	W19a	Tokomaru	3675	U/A

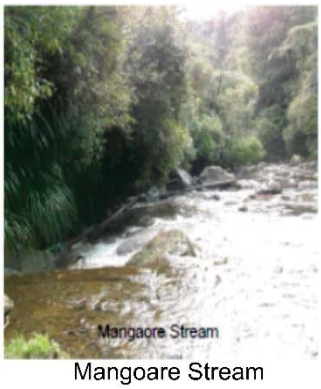
Appendix C

Supplementary materials for Chapter 3

1. Shannon and Dannevirke drinking-water sources

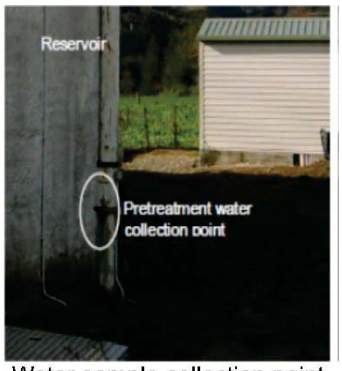
Shannon

In Shannon, water is obtained via direct intake from the Mangaore stream. The Mangaore stream lies approximately 800 metres upstream from the weir at the power station. The reservoir is located in Mangaore Road approximately two km from Shannon. The water gravity feeds from the stream through a 225 mm pipe into the Shannon Township and is pumped up from the reservoir to the tanks servicing Mangaore village with 1500 people. At this site, during heavy rain turbidity in excess of 100 NTU has been recorded (http://www.horowhenua. govt.nz/). To remove turbidity, water is settled and then chlorinated before storage and distribution.



Mangoare Stream


Stream water pumped



Water sample collection point

Dannevirke

Water is supplied from an infiltration gallery adjacent to the Tamaki River to the 6000 residents of Dannevirke. Nine manholes (MH1 to MH9) are an integral part of water intake, in which flow recording system was also installed to monitor river flow. Water is treated with lime to adjust pH before gravity fed prior enter into two reservoirs and chlorination is performed prior to the second reservoir Supply of water to the receivers are controlled by a float level operated control valve. At this site, turbidity of the river and town water has reached up to 1600 and 32 NTU respectively, during heavy rain (http://www.tararua.govt.nz/).

Tamaki River (Upstream from sampling points)

Tamaki River (Abstraction point)

2. Questionnaire used to collect the information about farms and farm management

Farm Questionnaire

Date:___/___/2011 Town:_____ Questionnaire ID___

1. General Information

1.1. What is the size of your farm in hectares (ha)?			
1.2. What is the topography of your farm?	% of farm		
• Hill country			
• Rolling			
• Flat			
• Other			
1.3. What are the total numbers of livestock on your farm?	Adults	Calves/lambs	Total
• Dairy			
• Beef			
• Sheep			
• Other (please specify)			
1.4. What ages of animals are present on your farm? Please specify the numbers present in each category.	Months		
	<3	3-12	>12
• Dairy			
• Beef			
• Sheep			
• Other (please specify)			
1.5. When are the calving and/or lambing periods for your farm? Please specify the start and end dates.	Date		
	Start	End	
• Calving			
• Lambing			

1.6. What types of housing are used for calves or lambs (for newborn up to 3 months)?

Calves	☐ Closed barn ☐ No housing	☐ Open barn ☐ Other (please specify)_____	☐ Partial open area
Lambs	☐ Closed barn ☐ No housing	☐ Open barn ☐ Other (please specify)_____	☐ Partial open area

2. Management

2.1. What are the main sources of drinking water for your livestock? Please tick all applicable.

☐ Town supply ☐ Ground water ☐ Irrigation ditches

☐ Rain water ☐ Stream or pond ☐ Bore hole

☐ Other (please specify):_____

2.2. Is drinking water disinfected or sanitised before it is provided to the livestock? If yes, please specify the type of treatment you use.

☐ Yes ☐ No Treatment type:_____

2.3. Is there any wastewater/effluent-treatment facilities present on your farm? Please specify the number and total capacity of systems present. ☐ Yes ☐ No

	Effluent Pond	Wetland	Other (please specify)
Number			
Capacity			

2.4. If you have answered yes to Q2.3, how do you use the treated effluent or wastewater? Please tick all applicable.

☐ Irrigate onto grassed paddock ☐ Irrigate onto crop

☐ Discharge into a river or stream ☐ Other (please specify)_____

2.5. If you irrigate effluent onto paddocks or crop fields, what kind of irrigation do you use?

☐ Spray irrigation ☐ Border strip irrigation ☐ Both spray and border strip irrigation ☐ Basin irrigation

2.6. How often do you irrigate effluent onto paddocks or crop fields in different seasons? Please specify the no. of days in each season.

Autumn	Winter	Spring	Summer

2.7. When do you irrigate onto the paddock or crop field?

☐ Immediately following the grazing of stock ☐ 2-6 days after grazing

☐ 7-15 days after grazing ☐ More than 15 days after grazing

2.8. What kinds of pasture/crops are planted in your paddock **this year (July 2010 to July 2011)?** Please tick or specify all applicable.

Grass	☐ Brome grass ☐ Rye Grass ☐ Kikuyu ☐ Other (please specify)_____
Legumes	☐ Lucerne ☐ Clovers ☐ Other (please specify)_____

Crops (please specify)	
Other (please specify)	

2.9. Do streams or rivers run through the farm?

☐ Yes ☐ No

If yes, how many branches of streams or rivers run through your farm?

☐ One ☐ Two ☐ Three ☐ Four ☐ More than four

2.10. If you have answered yes to Q2.9, are the streams/rivers fenced off?

☐ Yes ☐ No

If yes, please tick the type of fence/s, and specify the distance (in meters) of fence/s from the streams.

Stream/ River branch	Types of fence				Nearest distance of fence from river/stream (m)
	Hedge	Wire	Hedge and wire	Other (please specify)	

2.11. What types of drainage system are present on your farm?

☐ Trench/ditch ☐ Sub-surface drain ☐ Open drain ☐ Tile drain

Other(please specify)_____

2.12. Where is the outlet for your drainage system?

☐ No outlet ☐ Into a stream or river

☐ Other (please specify)_____

_____

☐ Into a ditch, that is not part of the farm. Where does the ditch drain?

_____

2.13. Do you have experience of flooding or failure in your drainage system?

☐ Yes ☐ No

If yes, what was the frequency of flooding in each season between **July 2010 and June 2011**?

☐ Winter	☐ 1-3 times	☐ 4-10 times	☐ >10 times
☐ Spring	☐ 1-3 times	☐ 4-10 times	☐ >10 times
☐ Summer	☐ 1-3 times	☐ 4-10 times	☐ >10 times
☐ Autumn	☐ 1-3 times	☐ 4-10 times	☐ >10 times

2.14. Do animals graze on the paddocks next to waterways?
☐ Yes ☐ No If yes, how long have they been on the paddocks adjacent to the waterways **in the last 30 days**? ☐ 1-5 days ☐ 6-10 days ☐ 10-20 days ☐ > 20 days

3. Disease

3.1. How many animals would you estimate have had scours or diarrhoea **in the last 30 days**?	Numbers
Calves (up to 45 days old)	a. Beef b. Dairy
Dairy cows (including dry stock and milking cows)	
Beef cattle	
Lambs (up to 45 days old)	
Sheep	
Other (please specify)	

3.2. Do you know the causes of scours or diarrhoea in your animals?
☐ Yes ☐ No If these were diagnosed by a laboratory, please tick the cause/s: ☐ Rotavirus ☐ *Giardia* ☐ *Cryptosporidium* ☐ *E. coli* ☐ Coccidia ☐ Other illness (please specify)_____

3.3. Did any animals die **in the last 30 days** that had had diarrhoea or scours?	
☐ Yes ☐ No If yes, please specify the number and probable causes of death.	
Number of deaths	Causes of death e.g *Giardia*

3.4. Has any member of your family or workers on your farm had diarrhoea **in the last 30 days**?	☐ Yes ☐ No

3.5. Has any member of your family or workers on your farm been diagnosed by a doctor as infected with the following agents **in the last 30 days**? Please tick.
☐ *Campylobacter* ☐ *Cryptosporidium* ☐ *Giardia* ☐ Do not know

Only answer these questions if calves/lambs are in pens	
3.6. If calves/lambs get scours, do you separate the diseased animals from the other stock?	☐ Yes ☐ No
3.7. What types of bedding do you use in pen?	
3.8. Do you change the bedding of the pen where all newborns are kept?	☐ Yes ☐ No
3.9. How often do you change the bedding?	☐ Weekly ☐ Fortnightly ☐ Used only once ☐ Monthly ☐ Not at all ☐ Others
3.10. How do you dispose of the used bedding?	☐ Burn it ☐ Bury it ☐ Other (please specify) _____

For the purpose of this study, we would like to request a map of your farm. This will help us to identify the class of stock in different paddocks when collecting the faecal samples from the ground. We would be grateful for your help and again assure you that neither you nor your farm will be identifiable in any results that are released from this work.

Please indicate the paddock, the number of animals and class of stock on your farm.

Code for paddock used for grazing (according to map)	Numbers	Class of stock

Thank you for your help with this work and for giving us your valuable time!

****

3. Follow up questions to collect the information about farms and farm management changes in each month

4.1. What are the reasons for any changes in the number of the livestock in your farm **in the last 30 days**?		
Reasons	Numbers	Class of stock
☐ Newborn		☐Beef ☐Dairy ☐ Sheep
☐ Bought		☐Beef ☐Dairy ☐ Sheep ☐Calves ☐Lambs
☐ Died		☐Beef ☐Dairy ☐ Sheep ☐Calves ☐Lambs
☐ Sold		☐Beef ☐Dairy ☐ Sheep ☐Calves ☐Lambs
☐ Others e.g. grazing others' livestock (please specify)		☐Beef ☐Dairy ☐ Sheep

4.2. How many animals would you estimate have had scours or diarrhoea **in the last 30 days**?	Numbers
Calves (up to 45 days old)	
Lambs (up to 45 days old)	
Dairy cows (including dry stock and milking cows)	
Beef cattle	
Sheep	
Others	

4.3. Do you know the causes of scours or diarrhoea in your animals?

☐ Yes ☐ No

If these were diagnosed by a laboratory, please tick the cause/s:

☐ Rotavirus ☐ *Giardia* ☐ *Cryptosporidium*

☐ *E. coli* ☐ Coccidian ☐ Other illness (please specify)_____

4.4. Was there any flooding/failure in your drainage system **in the last 30 days**?

☐ Yes ☐ No

4.5. Did any animals die **in the last 30 days** that had had diarrhoea or scours?

☐ Yes ☐ No

If yes, please specify the number and probable cause of death.

Number of deaths	Causes of death		
4.6. Has any member of your family or workers on your farm has had undiagnosed diarrhoea **in the 30 days**?	☐ Yes ☐ No		
4.7. Has any member of your family or workers on your farm been infected with the following agents in **the last 30 days**? Please tick.	*Campylobacter*	*Cryptosporidium*	*Giardia*
	☐ Yes ☐ No	☐ Yes ☐ No	☐ Yes ☐ No

Thank you for your co-operation and valuable time!

****

4. Letter of invitation to the farmers in the Shannon and Dannevirke catchment areas to participate in this study

Molecular Epidemiology Laboratory
Hopkirk Research Institute
Institute of Veterinary, Animal and Biomedical Sciences
Private Bag 11 222
Palmerston North
New Zealand
Telephone: +64 ({Fayer) 350 5799
Facsimile: +64 ({Fayer) 350 5716
www.massey.ac.nz

Study of Waterborne Diseases

Dear Farmer

I am a PhD student at Massey University and I am investigating three waterborne diseases that are common in New Zealand. We are planning a study investigating the presence of *Campylobacter*, *Cryptosporidium* and *Giardia* in two catchment areas local to Massey- Shannon and Dannevirke. The overall aim of this study is to increase our understanding on how these waterborne diseases behave in the environment.

In order to do that, we need to understand when these disease agents occur in water and when they occur on land adjacent to water sources. Therefore, we would like to collect faecal samples from the ground (and from calf/lamb pens, if available) on farms in your area. To do this, we would like permission to visit your farm once a month from end of July 2011 to November 2011. On the initial visit, we would ask you to answer a short questionnaire about your farm to gather details such as the numbers of animals present, the sources of their drinking water and recent health events within your stock. We guarantee that the information gathered will be treated with strict confidentiality. No information collected in this study will be used in any way that would allow people external to the study to identify any individual farm or farmer.

We would very much appreciate your involvement with this work and we would like to **contact you by phone within the next fortnight** to talk further about it and to answer any questions you may have. In the meantime, if you have any queries, you are welcome to contact us at the following

Rima Shrestha

PhD Student

Hopkirk Research Institute

Private Bag 11222, IVABS

Massey University

Palmerston North

06 350 5799 extn 81208

Debbie Prattley

Lecturer in Veterinary Public Health
Hop irk Research Institute
Private Bag 11222, IVABS
Massey University
Palmerstone North
06 350 5799 extn 81186
D.J.Prattley@massey.ac.nz

5. Information sheet on research study for the farmers of Shannon and Dannevirke catchment areas

Molecular Epidemiology Laboratory
Hopkirk Research Institute
Institute of Veterinary, Animal and Biomedical Sciences
Private Bag 11 222
Palmerston North
New Zealand
Telephone: +64 ({Fayer) 350 5799
Facsimile: +64 ({Fayer) 350 5716
www.massey.ac.nz

Study title: Study of waterborne diseases in two catchment areas in the North Island of New Zealand

We are a veterinary research team at Massey University and we are writing to invite you to participate in a study of waterborne diseases. The principal researcher is Rima Devi Shrestha, a veterinarian from Nepal, who is doing this study as part of her postgraduate research programme at Massey. Her supervisory research team is all vets working at Massey, namely Dr. Debbie Prattley (team leader), Dr Alex Grinberg, Dr Eve Pleydell and Prof. Nigel French.

Background to this research

Waterborne diseases occur all over the world. In New Zealand, *Campylobacter*, *Cryptosporidium* and *Giardia* are the main agents that cause human enteric diseases. This study aims to increase our understanding of how these three common microbes are behaving in the environment. In particular, we want to look at the distribution of different genetic types of the pathogens (disease causing bugs) and to try to understand the factors that help them to spread in the environment.

Why have you been invited to take part?

We are approaching water treatment plants and farms in the two catchment areas within a reasonable distance of Massey: Shannon and Dannevirke. From water treatment plants, we will take pre-treatment water samples to assess the presence of these three bugs in the water at different dates during the study period. We would also like to collect faecal samples from the ground (and from calf/lamb pens, if available) on farms in those catchment areas to help us to identify whether these bugs are present on farms in the catchment area and whether they are causing disease problems on these farms. This information will help us understand the distribution and movement of the three pathogens in the catchment environments, and will allow us to assess the best ways to help prevent these human and animal diseases

What will happen if you agree to take part?

1. We will visit your farm four times between August 2011 and November 2011 in order to collect faecal samples of your young stock.

2. On the first visit, we would like to spend some time (30-60 mins) with you asking some general questions about your farm, such as numbers and classes of stock, water sources, effluent disposal and recent health events in the stock.
3. During the subsequent three visits, we will ask (either in person or by telephone) a small number of follow-up questions to ascertain if there have been any important changes in stock numbers, grazing areas, and health events since the last visit.
4. After every visit, we will send you a report detailing the results of the diagnostic tests that will be run on the faecal samples we collect.

Do you have to take part?

No, not at all. Your participation in this study is voluntary; and you may decline to participate at any time before or during the study.

Confidentiality

We will be sending the results of the water samples to the District Council. However, neither the information collected from your farm during conversations or interviews nor the faecal test results will be disclosed to any person outside our group other than yourself. The individual farm results will only be seen by the researcher, and her supervisors, and will be remain strictly confidential. Data will be stored securely and will be made available only to persons conducting the study unless the participating farmers specifically give permission, in writing, to do otherwise. Furthermore, neither the individual farmers nor their farms will be identifiable in any study results that may be published or presented without their explicit consent.

Many thanks for your interest in the study.

Rima Shrestha will visit your farm and will collect the samples. However, if you have any questions or wish to discuss the study further, please feel free to contact either of us:

Rima Shrestha

PhD Student

Hopkirk Research Institute

Private Bag 11222, IVABS

Massey University

Palmerston North

06 350 5799 extn 81208

Debbie Prattley

Lecturer in Veterinary Public Health
Hopkirk Research Institute
Private Bag 11222, IVABS
Massey University
Palmerston North
06 350 5799 extn 81186
D.J.Prattley@massey.ac.nz

6. Letter of consent to participate in the research

Molecular Epidemiology Laboratory
Hopkirk Research Institute
Institute of Veterinary, Animal and Biomedical Sciences
Private Bag 11 222
Palmerston North
New Zealand
Telephone: +64 ({Fayer) 350 5799
Facsimile: +64 ({Fayer) 350 5716
www.massey.ac.nz

MASSEY UNIVERSITY
TE KUNENGA KI PŪREHUROA

Confidentiality Agreement between Participant and Researchers for study of waterborne Diseases

The research team appreciates your involvement in this study of waterborne diseases and is committed to privacy of all personal information.

All information included in the questionnaire and all test results will be treated in confidence and will not be published or disclosed to third parties (for example your council) by the research team. Data gathered will be presented and analysed in a way that will not disclose personal details and will not be traceable to you as an individual or your properties.

All information will be stored in a computer database that is accessible only by the researchers using a security code. Paper documents will be stored up to 3 years and then destroyed.

Date _____ Signed _____

Researchers
Rima Shrestha, Debbie Prattley, Alex Grinberg, Eve Pleydell and Prof. Nigel French.
Molecular Epidemiology Laboratory, Hopkirk Research Institute, Institute of Veterinary, Animal and Biomedical Sciences, Private Bag 11 222, Palmerston North, Telephone: 350 5799

I understand that the researchers will make every attempt to ensure my privacy as outlined above.
Date _____ Signed _____

Participant:
Name _____

Address_____

_____

"This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.
If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher (s), please contact Professor John O'Neill, Director (Research Ethics), telephone 06

7. Filtration equipment used in the field to collect water samples from the Tamaki River in Dannevirke

8. Matrix spiking process

Widely varying recovery rates have been reported worldwide at different sampling sites, and for differing water qualities, using the USEPA 1623 technique. The recovery rate of (oo)cysts from river water depends on the amount of (oo)cysts present and the physical (e.g. turbidity) and chemical (e.g. pH) properties of water, and it has not yet been estimated for New Zealand Rivers or raw water. Therefore, to determine the recovery rate of (oo)cysts in New Zealand raw surface water, additional water samples were collected, (oo)cysts were spiked into the water samples and a number of (oo)cysts on the slides were counted.

Briefly, the water samples were collected from both sampling sites once every three months during 2012, using a filtration technique as described in section Chapter 3. Approximately 97 L water was pumped through the Filta-Max® filter (IDEXX, Westbrook, Maine, USA), while an additional three litres water were collected in a sterile bottle. In the laboratory, two vials of ColorSeed™ (BTF Pty Ltd. Sydney,

Australia) containing (oo)cysts were spiked into the bottled water and were filtered through the same filta-max filter. Each ColorSeed™ vial contains exactly 100 flow sorted *Cryptosporidium parvum* (Iowa Strain) oocysts and 100 flow sorted *Giardia lamblia* cysts labelled with a permanent red dye. The (oo)cysts are inactivated with gamma-irradiation that display typical morphology under the microscope. Then the sample was processed as described in Chapter 3. Finally, the slide was examined microscopically for determining the presence of ColourSeed™ (oo)cysts on the slide. For each slide, the number of the ColourSeed™ (oo)cysts were recorded.

Appendix D

Supplementary materials for Chapter 4

1. **Sequences of *Cryptosporidium* spp. isolates amplified at 18S SSU RNA gene. The number denotes isolate from each specimen.**

>4079

TGACTCATAATAACTTTACGGATCACATTATGTGACATATCATTCAAGTTTCTGACCTAT
CAGCTTTAGACGGTAGGGTATTGGCCTACCGTGGCTATGACGGGTAACGGGGAATTAGGG
TTCGATTCCGGAGAGGGAGCCTGAGAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGC
GCAAATTACCCAATCCTAATACAGGGAGGTAGTGACAAGAAATAACAATACAGAACCTTA
CGGTTTTGTAATTGGAATGAGTTAAGTATAAACCCCTTAACAAGTATCAATTGGAGGGCA
AGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTTGCA
GTTAAAAAGCTCGTAGTTAATCTTCTGTTAATTTTTATATATAATATCACGATATTTATA
TAATATTAACATAATTCATATTACTTTTTAGTATATGAAACTTTACTTTGAGAAAATTAG
AGTGCTTAAAGCAGGCTATTGCCTTGAATACTCCAGCATGGAATAATATTAAGGATTTTT
ATTCTTCTTATTGGTTCTAGAATAAAAATGATGATTAATAGGGACAGTTGGGGGCATTTG
TATTTAACAGTCAGAGGTGAAATTCTTAGATTTGTTAAAGACAAACTACTGCGAAAGCAT
TTGCCAAGGATGTTTTCATTAATCAAGAACGAAAG

>4080

TGACTCATAATAACTTTACGGATCACATTATGTGACATATCATTCAAGTTTCTGACCTAT
CAGCTTTAGACGGTAGGGTATTGGCCTACCGTGGCTATGACGGGTAACGGGGAATTAGGG
TTCGATTCCGGAGAGGGAGCCTGAGAAACGGCTACCACATCTAAGGAAGGCAGCAGGCGC
GCAAATTACCCAATCCTAATACAGGGAGGTAGTGACAAGAAATAACAATACAGAACCTTA
CGGTTTTGTAATTGGAATGAGTTAAGTATAAACCCCTTAACAAGTATCAATTGGAGGGCA
AGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTTGCA
GTTAAAAAGCTCGTAGTTAATCTTCTGTTAATTTTTATATATAATATCACGATATTTATA
TAATATTAACATAATTCATATTACTTTTTAGTATATGAAACTTTACTTTGAGAAAATTAG
AGTGCTTAAAGCAGGCTATTGCCTTGAATACTCCAGCATGGAATAATATTAAGGATTTTT
ATTCTTCTTATTGGTTCTAGAATAAAAATGATGATTAATAGGGACAGTTGGGGGCATTTG

TATTTAACAGTCAGAGGTGAAATTCTTAGATTTGTTAAAGACAAACTACTGCGAAAGCAT
TTGCCAAGGATGTTTTCATTAATCAAGAACGAA

>4083

GTGACTCATAATAACTTTACGGATCACATTAAATGTGACATATCATTCAAGTTTCTGACC
TATCAGCTTTAGACGGTAGGGTATTGGCCTACCGTGGCAATGACGGGTAACGGGGAATTA
GGGTTCGATTCCGGAGAGGGAGCCTGAGAAACGGCTACCACATCTAAGGAAGGCAGCAGG
CGCGCAAATTACCCAATCCTAATACAGGGAGGTAGTGACAAGAAATAACAATACAGGACT
TTTTGGTTTTGTAATTGGAATGAGTTAAGTATAAACCCCTTTACAAGTATCAATTGGAGG
GCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTT
GCAGTTAAAAAGCTCGTAGTTGGATTTCTGTTAATAATTTATATAATAAAATATTTTGAA
TATTTATATAACATTAACATAATATTTCATATTACTATATTTTTTAGTATATGAAATTTT
ACTTTGAGAAAATTAGAGTGCTTAAAGCAGGCATATGCCTTGAATACTCCAGCATGGAAT
AATATTAAAGATTTTTATCTTTCTTATTGGTTCTAAGATAAGAATAATGATTAATAGGGA
CAGTTGGGGGCATTTGTATTTAACAGTCAGAGGTGAAATTCTTAGATTTGTTAAAGACAA
ACTAATGCGAAAGCATTTGCCAAGGATGTTTTCATTAATC

>4085

ATTTGGTGACTCATAATAACTTTACGGATCACATTAAATGTGACATATCATTCAAGTTTC
TGACCTATCAGCTTTAGACGGTAGGGTATTGGCCTACCGTGGCAATGACGGGTAACGGGG
AATTAGGGTTCGATTCCGGAGAGGGAGCCTGAGAAACGGCTACCACATCTAAGGAAGGCA
GCAGGCGCGCAAATTACCCAATCCTAATACAGGGAGGTAGTGACAAGAAATAACAATACA
GGACTTTTTGGTTTTGTAATTGGAATGAGTTAAGTATAAACCCCTTTACAAGTATCAATT
GGAGGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAG
TTGTTGCAGTTAAAAAGCTCGTAGTTGGATTTCTGTTAATAATTTATATAAAATATTTTG
ATGAATATTTATATAATATTAACATAATTCATATTACTATATATTTTAGTATATGAAATT
TTACTTTGAGAAAATTAGAGTGCTTAAAGCAGGCATATGCCTTGAATACTCCAGCATGGA
ATAATATTAAAGATTTTTATCTTTCTTATTGGTTCTAAGATAAGAATAATGATTAATAGG
GACAGTTGGGGGCATTTGTATTTAACAGTCAGAGGTGAAATTCTTAGATTTGTTAAAGAC
AAACTAATGCGAAAGCATTTGCCAAGGATGTTTTCATTAATCAAG

>4086
TGACTCATAATAACTTTACGGATCACATTAAATGTGACATATCATTCAAGTTTCTGACCT
ATCAGCTTTAGACGGTAGGGTATTGGCCTACCGTGGCAATGACGGGTAACGGGGAATTAG
GGTTCGATTCCGGAGAGGGAGCCTGAGAAACGGCTACCACATCTAAGGAAGGCAGCAGGC
GCGCAAATTACCCAATCCTAATACAGGGAGGTAGTGACAAGAAATAACAATACAGGACTT

TTTGGTTTTGTAATTGGAATGAGTTAAGTATAAACCCCTTTACAAGTATCAATTGGAGGG
CAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTTG
CAGTTAAAAAGCTCGTAGTTGGATTTCTGTTAATAATTTATATTATAAAATATTTTGAAT
ATTTATATAATATTAACATAATTCATATTACTAGTATATATTTTAGTATATGAAATTTTA
CTTTGAGAAAATTAGAGTGCTTAAAGCAGGCATATGCCTTGAATACTCCAGCATGGAATA
ATATTAAAGATTTTTATCTTTCTTATTGGTTCTAAGATAAGAATAATGATTAATAGGGAC
AGTTGGGGGCATTTGTATTTAACAGTCAGAGGTGAAATTCTTAGATTTGTTAAAGACAAA
CTAATGCGAAAGCATTTGCCAAGGATGTTTTCATTAATCAAGAA

>4087

TGGTGACTCATAATAACTTTACGGATCACATTATGTGACATATCATTCAAGTTTCTGACC
TATCAGCTTTAGACGGTAGGGTATTGGCCTACCGTGGCTATGACGGGTAACGGGGAATTA
GGGTTCGATTCCGGAGAGGGAGCCTGAGAAACGGCTACCACATCTAAGGAAGGCAGCAGG
CGCGCAAATTACCCAATCCTAATACAGGGAGGTAGTGACAAGAAATAACAATACAGAACC
TTACGGTTTTGTAATTGGAATGAGTTAAGTATAAACCCCTTAACAAGTATCAATTGGAGG
GCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTGTT
GCAGTTAAAAAGCTCGTAGTTAATCTTCTGTTAATTTTTGTATATAATATCACGATATTT
ATATAATATTAACATAATTCATATTACTTTTTAGTATATGAAACTTTACTTTGAGAAAAT
TAGAGTGCTTAAAGCAGGCTATTGCCTTGAATACTCCAGCATGGAATAATATTAAGGATT
TTTATTCTTCTTATTGGTTCTAGAATAAAAATGATGATTAATAGGGACAGTTGGGGGCAT
TTGTATTTAACAGTCAGAGGTGAAATTCTTAGATTTGTTAAAGACAAACTACTGCGAAAG
CATTTGCCAAGGATGTTTTCATTAATCAAGAACGAAAGTTAGGGGATCGAAGACGA

>4089

TGGTGACTCATAATAACTTTACGGATCACATTAAATGTGACATATCATTCAAGTTTCTGA
CCTATCAGCTTTAGACGGTAGGGTATTGGCCTACCGTGGCAATGACGGGTAACGGGGAAT
TAGGGTTCGATTCCGGAGAGGGAGCCTGAGAAACGGCTACCACATCTAAGGAAGGCAGCA
GGCGCGCAAATTACCCAATCCTAATACAGGGAGGTAGTGACAAGAAATAACAATACAGGA
CTTTTTGGTTTTGTAATTGGAATGAGTTAAGTATAAACCCCTTTACAAGTATCAATTGGA
GGGCAAGTCTGGTGCCAGCAGCCGCGGTAATTCCAGCTCCAATAGCGTATATTAAAGTTG
TTGCAGTTAAAAAGCTCGTAGTTGGATTTCTGTTAATAATTTATATAAAATATTTTGATG
AATATTTATATAATATTAACATAATTCATATTACTATATATTTTAGTATATGAAATTTTA
CTTTGAGAAAATTAGAGTGCTTAAAGCAGGCATATGCCTTGAATACTCCAGCATGGAATA
ATATTAAAGATTTTTATCTTTCTTATTGGTTCTAAGATAAGAATAATGATTAATAGGGAC
AGTTGGGGGCATTTGTATTTAACAGTCAGAGGTGAAATTCTTAGATTTGTTAAAGACAAA
CTAATGCGAAAGCATTTGCCAAGGATGTTTTCATTAATCAAGAACGAAAGTTAGGGGATC
GAAGA

2. Sequences of *Cryptosporidium* spp. Isolates amplified at gp60 gene. The number denotes isolate from each specimen.

>3795

TGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCATCATCATCATCATCATCATCATC
ATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACTGGAGAAGACGC
AGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGTTCTGAAGAGGA
AGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACTCCAGCTCAAAG
TGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGCGGCACTTCATT
TGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGTGCCTACACTAT
CGTCTATGCACCTATAAAAGACCAAACASATCCCGCACCAAGATATATCTCTGGTGAAGT
TACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTTAACGGTCAGGA
TTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGATCTGCGGGTCA
GGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCTGCTGCAACCGT
CGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCTGTTCCAAACGT
CGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAACCTTTCTATAC
CGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAACGGAGACTTGGT
GATA

>4009

TCATCATCATCATCATCATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCA
AATAAGGCAAGAACTGGAGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCT
TCTGGTAGCCAGGGTTCTGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGTT
TCCCAACCCACTACTCCAGCTCATAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACT
CCAAAAGAAGAATGCGGCACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCG
ACATTGAAGTGTGGTGCCTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCC
GCACCAAGATATATCTCTGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACA
GTTAAAATCAAGGTTAACGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGCAGTCCA
ACTGAAAATGGCGGATCTGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAG
GAAACCAGTGAAGCTGCTGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAA
AGAATTGAAGTGGCTGTACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTG
GTTGCAGACGATAAACCTTTCTATACCGGCGCAGACAGCGGCACTACCAATGGTGTCTAC
AGGTTGAATGAGAACGGAGA

>4075

TCGTCATCATCATCATCATCATCATCATCATCATCATCATCAACATCAACCGTCGCA
CCAGCAAATAAGGCAAGAACTGGAGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACT
GAAGCTTCTGGTAGCCAGGGTTCTGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGT
GCTGCTTCCCAACCCACTACTCCAGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAA
GCTACTCCAAAAGAAGAATGCGGCACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCA
GCTGCGACATTGAAGTGTGGTGCCTACACTATCGTCTATGCACCTATAAAAGACCAAACA
GATCCCGCACCAAGATATATCTCTGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGAT
AATACAGTTAAAATCAAGGTTAACGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGT
AGTCCAACTGAAAATGGCGGATCTGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTC
TCAGAGGAAACCAGTGAAGCTGCTGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGT
GGTAAAAGAATTGAAGTGGCTGTACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTAC
AGTTTGGTTGCAGACGATAAACCTTTCTATACCGGCGCAAACAGCGGCACTACCAATGGT
GTCTACAGGTTGAATGAGAACGGAGACTTGGTGA

>4076

GGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCATCATCATCATCATCATC
ATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACTGGAGAAGACGC
AGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGGAGCCAGGGTTCTGAAGAGGA
AGATAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCGCTACTCCAGCTCAAAG
TGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGCGGCACTTCATT
TGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGTGCCTACACTAT
CGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAAGATATATCTCTGGTGAAGT
TACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTTAACGGTCAGGA
TTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGATCTGCGGGTCA
GGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCTGCTGCAACCGT
CGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCTGTACCAAACGT
CGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAACCTTTCTATAC
CGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAACGGAGACTTGG

>4077

TCATGCGTCATCGTCATCGTCATCATCATCATCATCATCATCATCATCATCATCATC
AACATCAACCGTCGCACCAGCAAATAAGGCAAGAACTGGAGAAGACGCAGAAGGCAGTCA
AGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGTTCTGAAGAGGAAGGTAGTGAAGA

CGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACTCCAGCTCAAAGTGAAGGCGCAAC
TACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGCGGCACTTCATTTGTAATGTGGTT
CGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGTGCCTACACTATCGTCTATGCACC
TATAAAAGACCAAACAGATCCCGCACCAAGATATATCTCTGGTGAAGTTACATCTGTAAC
CTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTTAACGGTCAGGATTTCAGCACTCT
CTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGATCTGCGGGTCAGGCTTCATCAAG
ATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCTGCTGCAACCGTCGATTTGTTTGC
CTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCTGTACCAAACGTCGAAGATGCATC
TAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAACCTTTCTATACCGGCGCAAACAG
CGGCACTACCAATGGTGTCTACAGGTTGAATGAGAACGGAGACTTGGTGA

>4077 (sequence 2)

TAAAGGATGTTCCTGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCAT
CATCATCATCATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGG
CAAGAACTGGAGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTA
GCCAGGGTTCTGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAAC
CCACTACTCCAGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAG
AAGAATGCGGCACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGA
AGTGTGGTGCCTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAA
GATATATCTCTGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAA
TCAAGGTTAACGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAA
ATGGCGGATCTGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCA
GTGAAGCTGCTGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTG
AAGTGGCTGTACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAG
ACGATAAACCTTTCTATACCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGA
ATGAGAACGGAGACT

>4078

ATGTATTCCTCGTTGAGGGCTCATCATCGTCATGCGTCATCGTCATCATCATGCATCATC
ATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACTGG
AGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGTTC
TGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACTCC
AGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGCGG
CACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGTGC
CTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAAGATATATCTC
TGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTTAA

CGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGATC
TGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCTGC
TGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCTGT
ACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAACC
TTTCTATACCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAACGG
AGACTTGGTTGATA

>4079

TCATCATCGTCATGCGTCATCGTCATCATCATCATCATCATCATCATCATCATCATCATC
AACATCAACCGTCGCACCTGCAAATAAGGCAAGAACTGGAGAAGACGCAGAAGGCAGTCA
AGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGTTCTGAAGAGGAAGGTAGTGAAGA
CGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACTCCAGCTCAAAGTGAAGGCGCAAC
TACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGCGGCACTTCATTTGTAATGTGGTT
CGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGTGCCTACACTATCGTCTATGCACC
TATAAAAGACCAAACAGATCCCGCACCAAGATATATCTCTGGTGAAGTTACATCTGTAAC
CTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTTAACGGTCAGGATTTCAGCACTCT
CTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGATCTGCGGGTCAGGCTTCATCAAG
ATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCTGCTGCAACCGTCGATTTGTTTGC
CTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCTGTACCAAACGTCGAAGATGCATC
TAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAACCTTTCTATACCGGCGCAAACAG
CGGCACTACCAATGGTGTCTACAGGTTGGATGAGAACGGAGACTTGGTGA

>4079 (sequence 2)

TAAAGGATGTTCCTGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCAT
CATCATCATCATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGG
CAAGAACTGGAGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTA
GCCAGGGTTCTGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAAC
CCACTACTCCAGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAG
AAGAATGCGGCACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGA
AGTGTGGTGCCTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAA
GATATATCTCTGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAA
TCAAGGTTAACGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAA
ATGGCGGATCTGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCA
GTGAAGCTGCTGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTG

AAGTGGCTGTACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAG
ACGATAAACCTTTCTATACCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGA
ATGAGAACGGAGACTTGGTTGAT

>4080

GAGGGCTCATCATCGTCATCGTCATCGTCATCATCATCATCATCATCATCATCATCATCA
TCATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACTGGAGAAGACGCAGAA
GGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGTTCTGAAGAGGAAGGT
AGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACTCCAGCTCAAAGTGAA
GGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGCGGCACTTCATTTGTA
ATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGTGCCTACACTATCGTC
TATGCACCTATAAAAGACCAAACAGATCCCGCACCAAGATATATCTCTGGTGAAGTTACA
TCTGTAACCTTTGAAAAGAGAGATAATACWGTTAAAATCAAGGTTAACGGTCAGGATTTC
AGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAATTGGCGGATCTGCGGGTCAGGCT
TCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCTGCTGCAACCGTCGAT
TTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCTGTACCAAACGTCGAA
GATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAACCTTTCTATACCGGC
GCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAACGGA

>4081

GGGCTCATCATCGTCATCGTCATCGTCATCATCATCATCATCATCATCATCATCATCATC
ATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACTGGAGAAGACGCAGAAGG
CAGTCAAGATTCTAGTGGTACCGAAGCTTCTGGTAGCCAGGGTTCTGAAGAGGAAGGTAG
TGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACTCCAGCTCAAAGTGAAGG
CGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGCGGCACTTCATTTGTAAT
GTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGTGCCTACACTATCGTCTA
TGCACCTATAAAGGACCAAACAGATCCCGCACCAAGATATATCTCTGGTGAAGTTACATC
TGTAACCTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTTAACGGTCAGGATTTCAG
CACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGATCCGCGGGTCAGGCTTC
ATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCTGCTGCAACCGTCGATTT
GTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCTGTACCAAACGTCGAAGA
TGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAACCTTTCTATACCGGCGC
AAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAACGGAGACTTGG

>4083

TCCTGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCATCATCATCATC
ATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACTGG
AGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGTTC
TGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACTCC
AGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGCGG
CACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGTGC
CTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAAGATATATCTC
TGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTTAA
CGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGATC
TGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCTGC
TGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCTGT
ACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAACC
TTTCTATACCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAACGG
AGACTTGGTTGAT

>4084

GTTCCTGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCATCATCATCA
TCATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACT
GGAGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGT
TCTGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACT
CCAGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGC
GGCACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGT
GCCTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAAGATATATC
TCTGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTT
AACGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAAWGGCGGA
TCTGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCT
GCTGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCT
GTACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAA
CCTTTCTATACCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAAC
GGAGACT

>4084 (sequence 2)

TAAAGGATGTTCCTGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCAT
CATCATCATCATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGG

CAAGAACTGGAGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTA
GCCAGGGTTCTGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAAC
CCACTACTCCAGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAG
AAGAATGCGGCACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGA
AGTGTGGTGCCTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAA
GATATATCTCTGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAA
TCAAGGTTAACGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAA
ATGGCGGATCTGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCA
GTGAAGCTGCTGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTG
AAGTGGCTGTACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAG
ACGATAAACCTTTCTATACCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGA
ATGAGAACGGAGACTTGGTTGAT

>4085

GTTCCTGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCATCATCATCA
TCATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACT
GGAGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGT
TCTGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACT
CCAGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGC
GGCACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGT
GCCTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAAGATATATC
TCTGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTT
AACGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGA
TCTGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCT
GCTGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCT
GTACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAA
CCTTTCTATTCCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAAC
GGAGACTTG

>4088

TCCTGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCATCATCATCATC
ATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACTGG
AGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGTTC
TGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACTCC
AGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGCGG
CACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGTGC

CTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAAGATATATCTC
TGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTTAA
CGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGATC
TGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCTGC
TGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCTGT
ACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAACC
TTTCTATTCCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAACGG
AGACT

>4088 (sequence 2)

TAAAGGATGTTCCTGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCAT
CATCATCATCATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGG
CAGAACTGGAGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTAG
CCAGGGTTCTGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACC
CACTACTCCAGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGA
AGAATGCGGCACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAA
GTGTGGTGCCTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAAG
ATATATCTCTGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAAT
CAAGGTTAACGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAA
TGGCGGATCTGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAG
TGAAGCTGCTGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGA
AGTGGCTGTACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGA
CGATAAACCTTTCTATACCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAA
TGAGAACGGAGA

>4089

GTTCCTGTTGAGGGCTCATCATCGTCATCGTCATCGTCATCGTCATCATCATCATCATCA
TCATCATCATCATCATCATCATCAACATCAACCGTCGCACCAGCAAATAAGGCAAGAACT
GGAGAAGACGCAGAAGGCAGTCAAGATTCTAGTGGTACTGAAGCTTCTGGTAGCCAGGGT
TCTGAAGAGGAAGGTAGTGAAGACGATGGCCAAACTAGTGCTGCTTCCCAACCCACTACT
CCAGCTCAAAGTGAAGGCGCAACTACCGAAACCATAGAAGCTACTCCAAAAGAAGAATGC
GGCACTTCATTTGTAATGTGGTTCGGAGAAGGTACCCCAGCTGCGACATTGAAGTGTGGT
GCCTACACTATCGTCTATGCACCTATAAAAGACCAAACAGATCCCGCACCAAGATATATC
TCTGGTGAAGTTACATCTGTAACCTTTGAAAAGAGTGATAATACAGTTAAAATCAAGGTT
AACGGTCAGGATTTCAGCACTCTCTCTGCTAATTCAAGTAGTCCAACTGAAAATGGCGGA
TCTGCGGGTCAGGCTTCATCAAGATCAAGAAGATCACTCTCAGAGGAAACCAGTGAAGCT

GCTGCAACCGTCGATTTGTTTGCCTTTACCCTTGATGGTGGTAAAAGAATTGAAGTGGCT
GTACCAAACGTCGAAGATGCATCTAAAAGAGACAAGTACAGTTTGGTTGCAGACGATAAA
CCTTTCTATACCGGCGCAAACAGCGGCACTACCAATGGTGTCTACAGGTTGAATGAGAAC
GGAGACTTGGTTGATAA

3. Sequences of *Cryptosporidium* spp. isolates amplified at HSP70 gene (new primer). The number denotes isolate from each specimen.

>4080

TGATACTCACTTGGGTGGTGAAGACTTTGACAACAGATTAGTTGAATTTTGCGTTCAAGA
CTTCAAGAGAAAGAATCGTGGAATGGACTTGACTACAAATGCCCGTGCATTAAGAAGACT
TAGAACTCAGTGTGAGCGTGCAAAGAGAACTTTATCTTCATCTACACAAGCCACTGTTGA
ATTGGATTCTCTTTATGAAGGTATTGACTACTCTGTAGCTATTAGTCGTGCAAGATTTGA
AGAACTTTGTTCCGATTACTTCCGTGCCACTTTATCACCAGTAGAGAAGGTACTCAAGGA
TGCTGGTATGGACAA

>4079

TGATACTCACTTGGGTGGTGAAGATTTTGATAACAGACTCGTAGAATTCTGTGTACAAGA
TTTCAAGAGAAAGAATAGAGGTATGGATTTAACCACAAATGCTAGAGCTTTAAGAAGACT
CAGAACTCAATGCGAGCGTGCAAAGAGAACTTTGTCATCTTCTACTCAAGCTACAATTGA
GTTAGATTCACTCTATGAAGGTATTGATTATTCAGTTGCCATCAGTAGAGCTAGATTCGA
AGAACTCTGCGCTGATTACTTCCGTGCAACTTTAGCTCCAGTTGAGAAAGTACTCAAGGA
TGCTGGTATGGACAA

>4087

TGATACTCACTTGGGTGGTGAAGATTTTGATAACAGACTCGTAGAATTCTGTGTACAAGA
TTTCAAGAGAAAGAATAGAGGTATGGATTTAACCACAAATGCTAGAGCTTTAAGAAGACT
CAGAACTCAATGCGAGCGTGCAAAGAGAACTTTGTCATCTTCTACTCAAGCTACAATTGA
GTTAGATTCACTCTATGAAGGTATTGATTATTCAGTTGCCATCAGTAGAGCTAGATTCGA
AGAACTCTGCGCTGATTACTTCCGTGCAACTTTAGCTCCAGTTGAGAAAGTACTCAAGGA
TGCTGGTATGGACAA

>4087 (sequence 2)

TGATACTCACTTGGGTGGTGAAGATTTTGATAACAGACTCGTAGAATTCTGTGTACAAGA
TTTCAAGAGAAAGAATAGAGGTATGGATTTAACCACAAATGCTAGAGCTTTAAGAAGACT
CAGAACTCAATGCGAGCGTGCAAAGAGAACTTTGTCATCTTCTACTCAAGCTACAATTGA
GTTAGATTCACTCTATGAAGGTATTGATTATTCAGTTGCCATCAGTAGAGCTAGATTCGA
AGAACTCTGCGCTGATTACTTCCGTGCAACTTTAGCTCCAGTTGAGAAAGTACTCAAGGA
TGCTGGTATGGACAA

4. Sequences of *Cryptosporidium* spp. isolates amplified at Actin gene. The number denotes isolate from each specimen.

>4079

CTGGTGACGATGCTCCTCGATGTGTATTTCCATCAATAGTAGGTCGTCCAAAAATGCCAG
GAGTTATGGTTGGTATGGACCAAAAGGATTGTTATGTAGGTGACGAGGCTCAATCAAAGA
GAGGTATATTGACATTAAAATATCCAATTGAACATGGTATTGTTACAAATTGGGAGGATA
TGGAAAAGATATGGCATCATACATTTTACAATGAATTACGTGTCGCTCCGGAAGAGCATC
CAGTATTGTTAACAGAGGCCCCAATGAATCCAAAGGTAAATCGTGAAAGAATGACACAAA
TTATGTTTGAGACATTTAATGTACCAGCAATGTATGTTAATATTCAAGCTGTTTTGTCTT
TGTATGCCTCAGGTCGTACAACAGGTATTGTATTAGATAGTGGAGATGGTGTTTCGCACA
CCGTTCCAATTTATGAAGGTTATGCCCTTCCTCATGCAATTATGAGGTTGGATTTGGCTG
GTCGTGACTTGACAGACTTCCTCATGAAAATTTTACATGATCGTGGCTACAGCTTTACGA
CAACAGCTGAAAGAGAAATAGTGAGGGATATCAAGGAAAAACTTTGCTATATCGCTTTGG
ACTATGAAGAGGAAATGAAGAAGTCTCAGGAATCTTCAGAAATTGAAAAGACATATGAAT
TACCTGATGGACATGTAATTACTGTGGGAAGTGAGAGATTTAGATGCCCAGAAGCTTTGT
TCCAGCCTGGTTTATTGGGTAAAGAGGCTGTTGGGTATTGGTGAGACCACTTTTCCAATC
TATTATGAAGTGCGATCTTTGATATTCGTAAGGATCTCTACGCAAATATTGTTTTGTCTG
GTGGTACGACAATGTATCCAGGCATTGGTGAGAGAATGACAAAAGAACTTACTTCATTAG
CACCTTCTACAATGAAGATTAAAGTTGTCGCCCCACCAGAGCGTAAGTATTCCGTTTGGA
TTGGTGGTTCAATTTTATCTTCGCTCTCAACGTTCCAACAAATGTGG

>4080

CAAAACAATATTTGCGTAGAGATCCCTACGAATATCAAAGATCGCACTTCATAATAGATT
GGAAAGTGGTCTCACCAATACCAACAGCCTCTTTACCCAATAAACCAGGTTGGAACAAAG
CTTCTGGGCATCTAAATCTCTCACTTCCCACAGTAATTACATGTCCATCAGGGTAATTCA
TATGTCTTTTCAATTTCTGAAGATTCCTGAGACTTCTTCATTTCCTCTTCATAGTCCAAA
AGCGATATAGCAAAGTTTTTCCTTGATATCCCTCACTATTTCTCTTTCAGCTGTTGTCGT
AAAGCTGTAGCCACGATCATGTAAAATTTTCATGAGGAAGTCTGTCAAGTCACGACCAGC
CAAATCCAACCTCATAATTGCATGAGGAAGGGCATAACCTTCATAAATTGGAACGGTGTG
CGAAACACCATCTCCACTATCTAATACAATACCTGTTGTACGACCTGAGGCATACAAAGA
CAAAACAGCTTGAATATTAACATACATTGCTGGTACATTAAATGTCTCAAACATAATTTG
TGTCATTCTTTCACGATTTACCTTTGGATTCATTGGGGCCTCTGTTAACAATACTGGATG
CTCTTCCGGAGCGACACGTAATTCATTGTAAAATGTATGATGCCATATCTTTTCCATATC
CTCCCAATTTGTAACAATACCATGTTCAATTGGATATTTTAATGTCAATATACCTCTCTT
TGATTGAGCCTCGTCACCTACATAACAATCCTTTTGGTCCATACCAACCATAACTCCTGG
CATTTTTGGACGACCTACTATTGATGGAAATACACATCGAGGAGCAT

>4087

TATGTAGGTGACGAGGCTCAATCAAAGAGAGGTATATTGACATTAAAATATCCAATTGAG
CATGGTATTGTTACAAATTGGGAGGATATGGAAAAGATATGGCATCATACATTTTACAAT
GAATTACGTGTCGCTCCGGAAGAGCATCCTGTATTGTTAACAGAGGCCCCATTGAATCCA
AAGGTAAATCGAGAAAGAATGACACAAATTATGTTTGAGACATTTAATGTACCAGCAATG
TATGTTAATATTCAAGCTGTTTTGTCTTTGTATGCCTCAGGTCGTACAACAGGTATTGTA
TTAGATAGTGGAGATGGTGTTTCGCACACCGTTCCAATTTATGAAGGTTATGCCCTTCCT
CATGCA

>4079 (sequence 2)

TTACGAATATCAAGATCGCACTTCATAATAGATTGGAAAGTGGTCTCACCAATACCAACA
GCCTCTTTACCCAATAAACCAGGTKGAAACAAAGCTTCTGGGCATCTAAATCTCTCACTT
CCCACAGTAATTACATGTCCATCAGGTAATTCATATGTCTTTTCAATTTCTGAAGATTCC
TGAGACTTCTTCATTTCCTCTTCATAGTCCAAAGCGATATAGCAAAGTTTTTCCTTGATA
TCCCTCACTATTTCTCTTTCAGCTGTTGTCGTAAAGCTGTAGCCACGATCATGTAAAATT
TTCATGAGGAAGTCTGTCAAGTCACGACCAGCCAAATCCAACCTCATAATTGCATGAGGA
AGGGCATAACCTTCATAAATTGGAACGGTGTGCGAAACACCATCTCCACTATCTAATACA
ATACCTGTTGTACGACCTGAGGCATACAAAGACAAAACAGCTTGAATATTAACATACATT
GCTGGTACATTAAATGTCTCAAACATAATTTGTGTCATTCTTTCACGATTTACCTTTGGA

TTCATTGGGGCCTCTGTTAACAATACTGGATGCTCTTCCGGAGCGACACGTAATTCATTG
TAAAATGTATGATGCCATATCTTTTCCATATCCTCCCAATTTGTAACAATACCATGTTCA
ATTGGATATTTTAATGTCAATATACCTCTCTTTGATTGAGCCTCGTCACCTACATAACAA
TCCTTTTGGTCCATACCAACCATAACTCCTGACATTTTTGGACGACCTACTATTGATGGA
AATACACATCGAGGAGCATCGTCACCAGC

>4080 (sequence 2)

CAAAACAATATTTGCGTAGAGATCCCTACGAATATCAAAGATCGCACTTCATAATAGATT
GGAAAGTGGTCTCACCAATACCAACAGCCTCTTTACCCAATAAACCAGGTTGGAACAAAG
CTTCTGGGCATCTAAATCTCTCACTTCCCACAGTAATTACATGTCCATCAGGGTAATTCA
TATGTCTTTTCAATTTCTGAAGATTCCTGAGACTTCTTCATTTCCTCTTCATAGTCCAAA
AGCGATATAGCAAAGTTTTTCCTTGATATCCCTCACTATTTCTCTTTCAGCTGTTGTCGT
AAAGCTGTAGCCACGATCATGTAAAATTTTCATGAGGAAGTCTGTCAAGTCACGACCAGC
CAAATCCAACCTCATAATTGCATGAGGAAGGGCATAACCTTCATAAATTGGAACGGTGTG
CGAAACACCATCTCCACTATCTAATACAATACCTGTTGTACGACCTGAGGCATACAAAGA
CAAAACAGCTTGAATATTAACATACATTGCTGGTACATTAAATGTCTCAAACATAATTTG
TGTCATTCTTTCACGATTTACCTTTGGATTCATTGGGGCCTCTGTTAACAATACTGGATG
CTCTTCCGGAGCGACACGTAATTCATTGTAAAATGTATGATGCCATATCTTTTCCATATC
CTCCCAATTTGTAACAATACCATGCTCAATTGGATATTTTAATGTCAATATACCTCTCTT
TGATTGAGCCTCGTCACCTACATAACAATCCTTTTGGTCCATACCAACCATAACTCCTGG
CATTTTTGGACGACCTACTATTGATGGAAATACACATCGAGGAGCAT

>4087 (sequence 2)

TATGTAGGTGACGAGGCTCAATCAAAGAGAGGTATATTGACATTAAAATATCCAATTGAG
CATGGTATTGTTACAAAYTGGGAGGATATGGAAAAGATATGGCATCATACATTTTACAAT
GAATTACGTGTCGCTCCGGAAGAGCATCCTGTATTGTTAACAGAGGCCCCATTGAATCCA
AAGGTAAATCGAGAAAGAATGACACAAATTATGTTTGAGACATTTAATGTACCAGCAATG
TATGTTAATATTCAAGCTGTTTTGTCTTTGTATGCCTCAGGTCGTACAACAGGTATTGTA
TTAGATAGTGGAGATGGTGTTTCGCACACCGTTCCAATTTATGAAGGTTATGCCCTTCCT
CATGCA

Appendix E

Supplementary materials for Chapter 5

Table E. 1. Sampling location (Shannon-"S" and Dannevirke-"D") and month of *C. jejuni* and *C. coli* DNA samples used in the massMLST study. The DNA samples were obtained from faecal samples collected from cattle and sheep in farms of that location, and from water samples collected during the chapter 4 study. Two known *C. jejuni* sequence types (**) were used as a control.

Sample ID	Sampling location	PCR Plate ID*	Sampling Month	Faecal samples			*Campylobacter* spp
				Farm ID	Species	Age group (month)	
S709	S	A4	August	SF2	Sheep	3 to 12	*C. jejuni*
S711	S	A5	August	SF2	Sheep	3 to 12	*C. jejuni*
S713	S	A6	August	SF2	Sheep	3 to 12	*C. jejuni*
S714	S	B5	August	SF2	Sheep	3 to 12	*C. jejuni*
S715	S	B7	August	SF2	Sheep	3 to 12	*C. jejuni*
S716	S	C12	August	SF2	Beef	3 to 12	*C. jejuni*
S718	S	A10	August	SF2	Beef	3 to 12	*C. jejuni*
S719	S	A7	August	SF2	Beef	3 to 12	*C. jejuni*
S723	S	B11	August	SF2	Beef	4 to 12	*C. coli*
S745	D	A8	August	DF1	Beef	>12	*C. jejuni*
S752	D	B3	August	DF1	Beef	>12	*C. jejuni*
S771	D	B12	August	DF3	Sheep	3 to 12	*C. coli*
S773	D	C10	August	DF3	Sheep	3 to 12	*C. jejuni*
S774	D	C1	August	DF3	Sheep	3 to 12	*C. coli*
S775	D	A9	August	DF3	Sheep	3 to 12	*C. jejuni*
S776	D	C2	August	DF3	Sheep	3 to 12	*C. coli*
S777	D	C3	August	DF3	Sheep	3 to 12	*C. coli*
S788	D	C11	August	DF4	Sheep	3 to 12	*C. jejuni*
S789	D	F7	August	DF4	Sheep	3 to 12	*C. jejuni*
S794	D	D1	August	DF4	Sheep	3 to 12	*C. jejuni*
S799	D	B1	August	DF4	Sheep	3 to 12	*C. jejuni*

S800	D	D2	August	DF4	Sheep	3 to 12	*C. jejuni*
S802	S	D3	August	SF4	Beef	3 to 12	*C. jejuni*
S805	S	F8	August	SF4	Beef	3 to 12	*C. jejuni*
S829	D	D4	August	DF5	Dairy	<3	*C. jejuni*
S833	S	D5	August	SF2	Beef	3 to 12	*C. jejuni*
S837	S	D9	September	SF2	Sheep	3 to 12	*C. jejuni*
S840	S	B2	September	SF2	Sheep	3 to 12	*C. jejuni*
S843	S	D6	September	SF2	Beef	>12	*C. jejuni*
S845	S	G8	September	SF2	Beef	>12	*C. jejuni*
S862	D	D7	September	DF6	Dairy	3 to 12	*C. jejuni*
S869	D	D8	September	DF6	Dairy	3 to 12	*C. jejuni*
S873	D	D9	September	DF6	Dairy	3 to 12	*C. jejuni*
S874	D	G5	September	DF1	Beef	>12	*C. jejuni*
S876	D	F10	September	DF1	Beef	>12	*C. jejuni*
S878	D	D10	September	DF1	Beef	>12	*C. jejuni*
S883	D	E11	September	DF1	Sheep	>12	*C. coli*
S886	D	C7	September	DF1	Sheep	>12	*C. coli*
S895	S	A3	September	SF4	Beef	3 to 12	*C. coli*
S906	D	E12	September	DF3	Sheep	>12	*C. coli*
S911	D	C8	September	DF3	Sheep	>12	*C. coli*
S917	D	F1	September	DF3	Sheep	>12	*C. coli*
S918	D	C9	September	DF3	Sheep	>12	*C. coli*
S922	D	D12	September	DF4	Sheep	3 to 12	*C. jejuni*
S923	D	F11	September	DF4	Sheep	3 to 12	*C. jejuni*
S925	D	F12	September	DF4	Sheep	3 to 12	*C. jejuni*
S933	D	G1	September	DF4	Sheep	3 to 12	*C. jejuni*
S939	D	F2	September	DF5	Dairy	<3	*C. coli*
S946	D	G2	September	DF5	Dairy	<3	*C. jejuni*
S949	D	F3	September	DF5	Dairy	<3	*C. coli*
S959	S	E1	September	SF2	Beef	3 to 12	*C. jejuni*
S961	S	G3	September	SF2	Beef	3 to 12	*C. jejuni*
S981	D	E2	September	DF6	Dairy	>12	*C. jejuni*
S994	D	A2	October	DF6	Dairy	>12	*C. jejuni*
S996	D	B4	October	DF1	Beef	>12	*C. jejuni*
S1017	S	G12	October	SF4	Sheep	<3	*C. jejuni*
S1025	S	G4	October	SF4	Beef	3 to 12	*C. jejuni*
S1032	D	E3	October	DF3	Sheep	>12	*C. jejuni*
S1036	D	F5	October	DF3	Sheep	>12	*C. coli*
S1037	D	F6	October	DF3	Sheep	>12	*C. coli*

S1041	D	B6	October	DF4	Sheep	>12	*C. jejuni*
S1043	D	F4	October	DF4	Sheep	>12	*C. coli*
S1053	D	G6	October	DF5	Dairy	<3	*C. jejuni*
S1061	D	G9	October	DF5	Dairy	<3	*C. jejuni*
S1065	D	G10	October	DF5	Dairy	<3	*C. jejuni*
S1094	S	G11	October	SF2	Beef	>12	*C. jejuni*
S1117	D	H1	October	DF6	Dairy	3 to 12	*C. jejuni*
S1123	D	H12	November	DF6	Dairy	3 to 12	*C. jejuni*
S1126	D	H2	November	DF6	Dairy	3 to 12	*C. jejuni*
S1144	D	E4	November	DF3	Sheep	>12	*C. coli*
S1146	D	G7	November	DF3	Sheep	>12	*C. coli*
S1147	D	E8	November	DF3	Sheep	>12	*C. coli*
S1150	D	C4	November	DF3	Sheep	>12	*C. coli*
S1152	D	A1	November	DF3	Sheep	<3	*C. coli*
S1155	D	E9	November	DF3	Sheep	>12	*C. coli*
S1159	D	H3	November	DF4	Beef	3 to 12	*C. jejuni*
S1165	D	H4	November	DF4	Beef	3 to 12	*C. jejuni*
S1175	D	C5	November	DF5	Dairy	3 to 12	*C. coli*
S1176	D	C6	November	DF5	Dairy	3 to 12	*C. coli*
S1179	D	E10	November	DF5	Dairy	3 to 12	*C. coli*
S1182	D	E5	November	DF5	Dairy	3 to 12	*C. jejuni*
S1185	D	D11	November	DF5	Dairy	3 to 12	*C. jejuni*
S1187	D	H5	November	DF5	Dairy	3 to 12	*C. jejuni*
S1189	D	H6	November	DF5	Dairy	3 to 12	*C. jejuni*
S1197	D	B8	November	DF6	Dairy	3 to 12	*C. jejuni*
S1201	D	H7	November	DF6	Dairy	3 to 12	*C. jejuni*
S1205	D	H8	November	DF6	Dairy	3 to 12	*C. jejuni*
S1216	D	H9	November	DF1	Beef	3 to 12	*C. jejuni*
S1219	D	H10	November	DF1	Beef	3 to 12	*C. jejuni*
S1224	D	H11	November	DF1	Beef	3 to 12	*C. jejuni*
W717a	River	B9	-	-	water	-	*C. jejuni*
W717b	River	E6	-	-	water	-	*C. jejuni*
W725a	River	B10	-	-	water	-	*C. jejuni*
W725b	River	E7	-	-	water	-	*C. jejuni*
H1579**	Control	A11	-		Human	-	*C. jejuni*
P1262a**	Control	A12	-		Poultry	-	*C. jejuni*

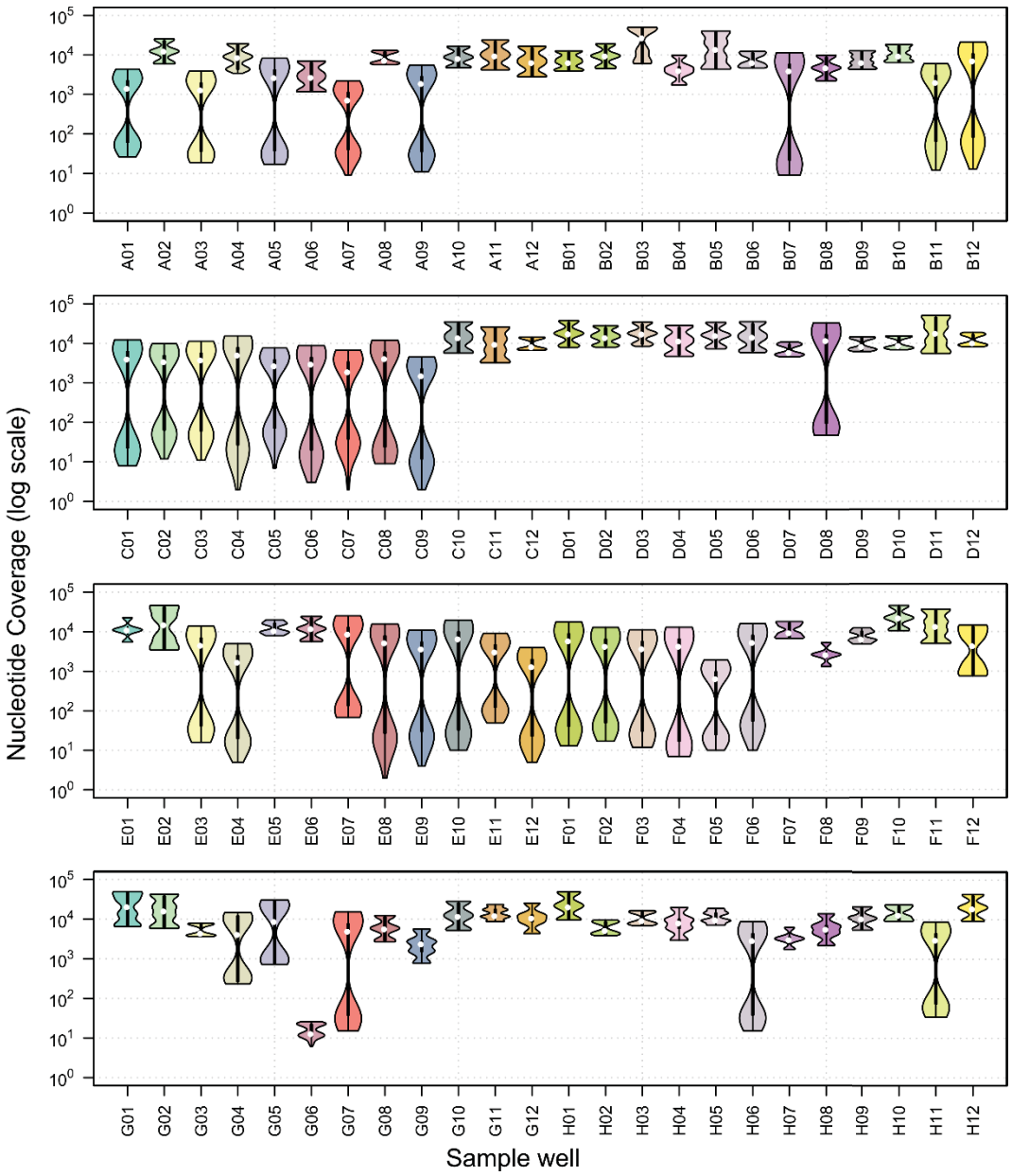


Figure E.1: Violin plots of nucleotide sequences coverage in log scale for each of 96 samples by plate location for a *glnA Campylobacter* MLST allele. The A01 to H12 represents wells in a 96-well plate, each of which contains a separate sample. The white dot is the median nucleotide sequences coverage and the dark black lines are the interquartile range along with whiskers.

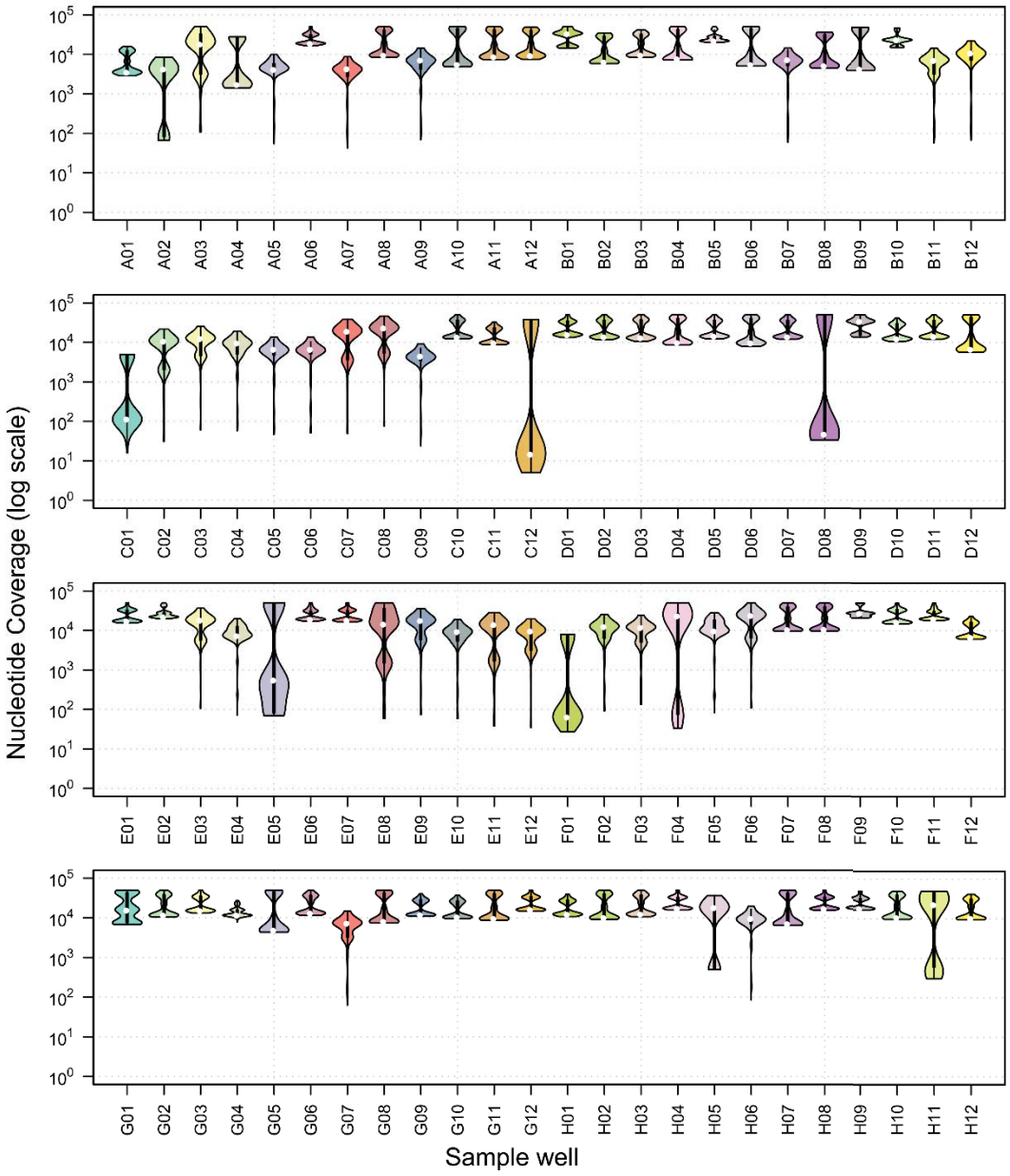


Figure E.2: Violin plots of nucleotide sequences coverage in log scale for each of 96 samples by plate location for a *aspA Campylobacter* MLST allele. The A01 to H12 represents wells in a 96-well plate, each of which contains a separate sample. The white dot is the median nucleotide sequences coverage and the dark black lines are the interquartile range along with whiskers.

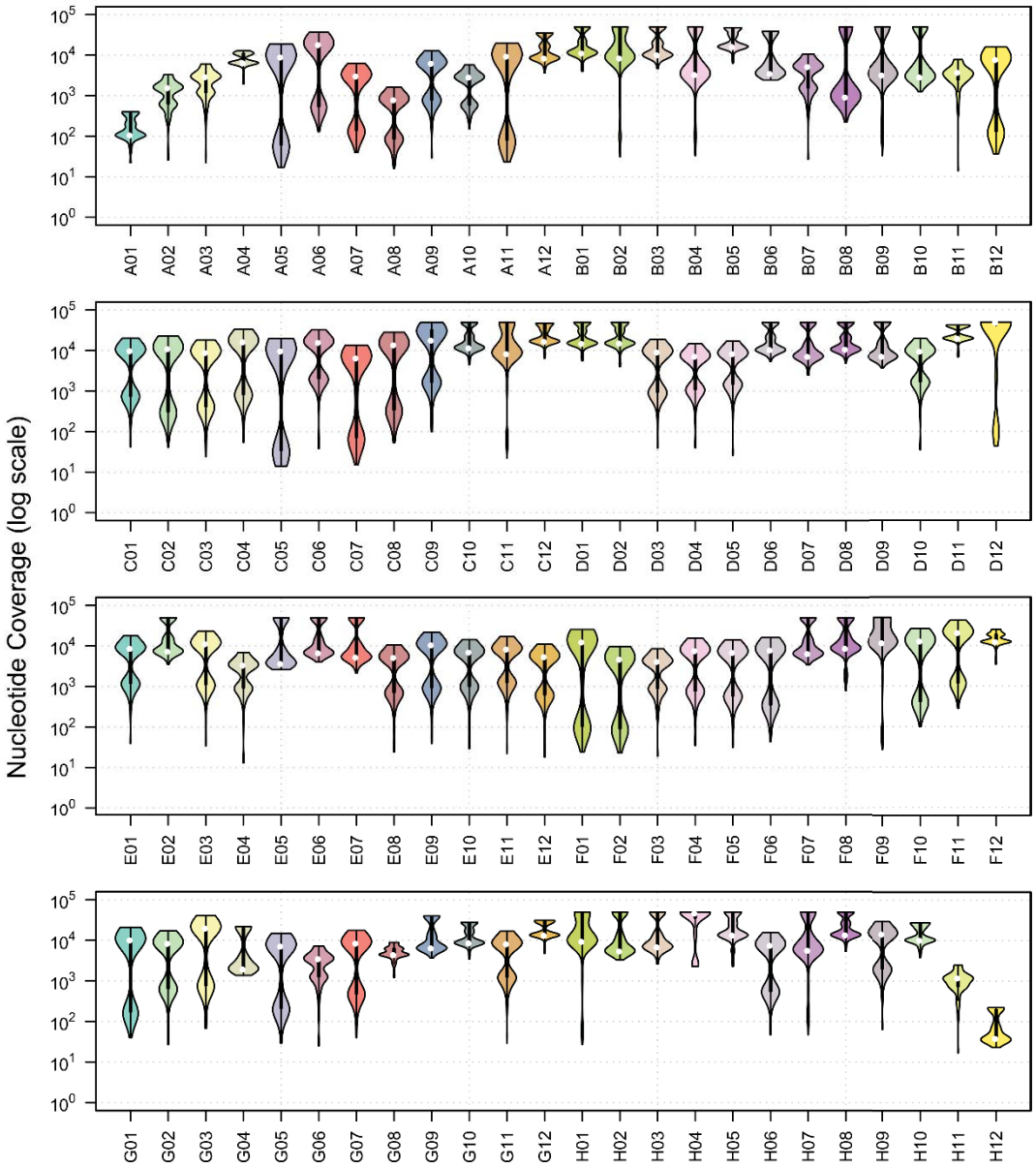


Figure E.3: Violin plots of nucleotide sequences coverage in log scale for each of 96 samples by plate location for a *uncA Campylobacter* MLST allele. The A01 to H12 represents wells in a 96-well plate, each of which contains a separate sample. The white dot is the median nucleotide sequences coverage and the dark black lines are the interquartile range along with whiskers.

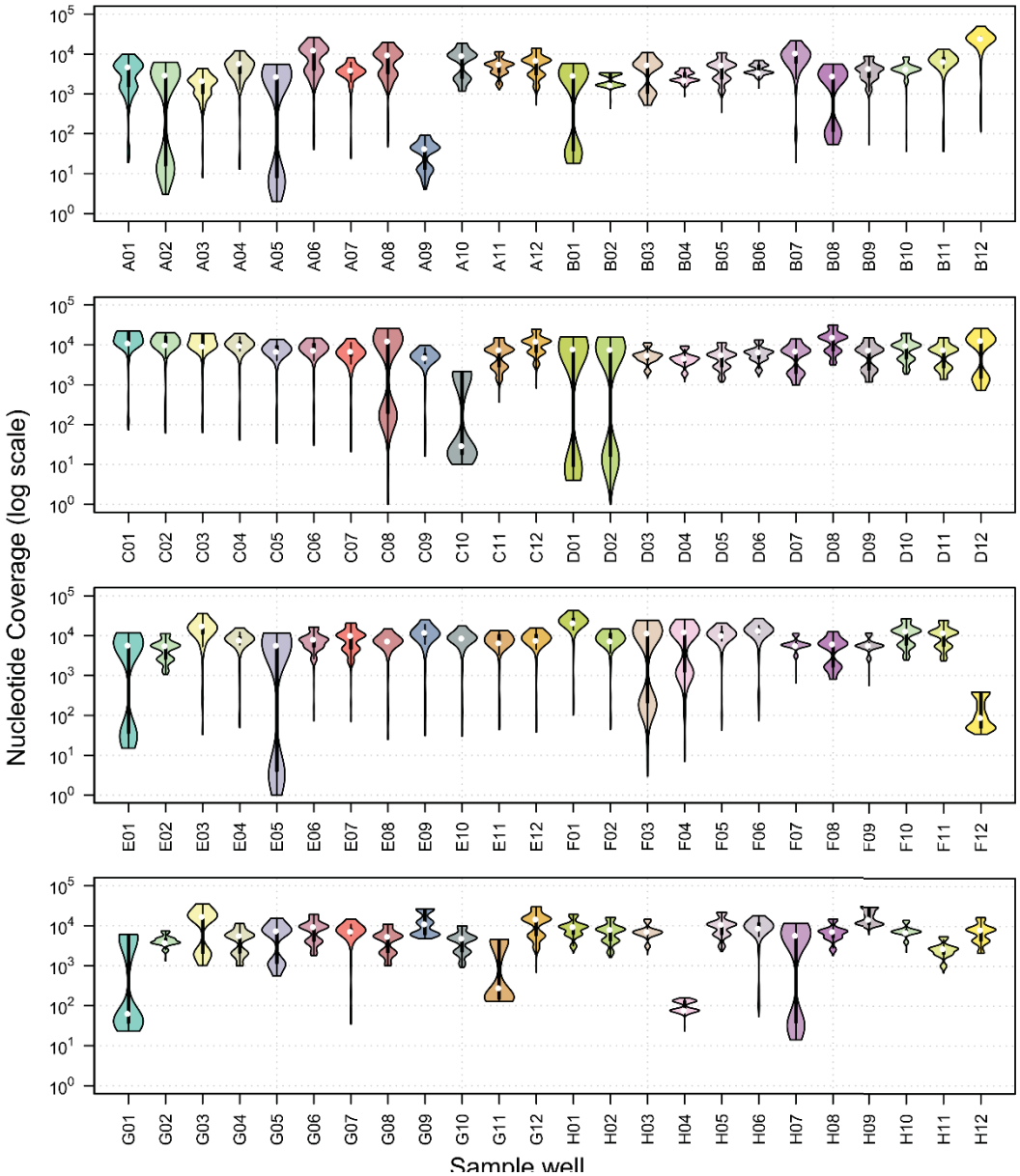


Figure E.4: Violin plots of nucleotide sequences coverage in log scale for each of 96 samples by plate location for a *pgm Campylobacter* MLST allele. The A01 to H12 represents wells in a 96-well plate, each of which contains a separate sample. The white dot is the median nucleotide sequences coverage and the dark black lines are the interquartile range along with whiskers.

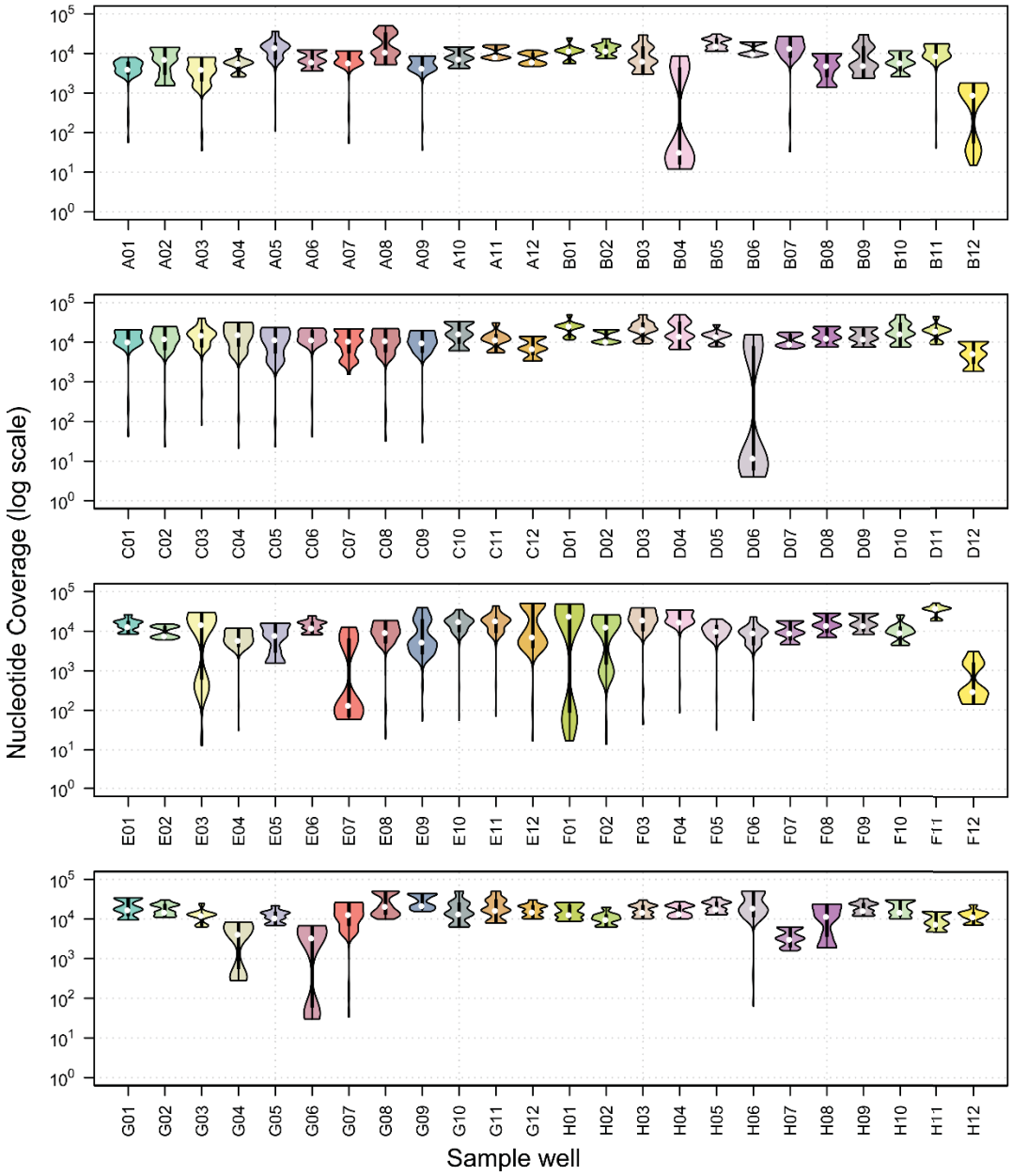


Figure E.5: Violin plots of nucleotide sequences coverage in log scale for each of 96 samples by plate location for a *gltA Campylobacter* MLST allele. The A01 to H12 represents wells in a 96-well plate, each of which contains a separate sample. The white dot is the median nucleotide sequences coverage and the dark black lines are the interquartile range along with whiskers.

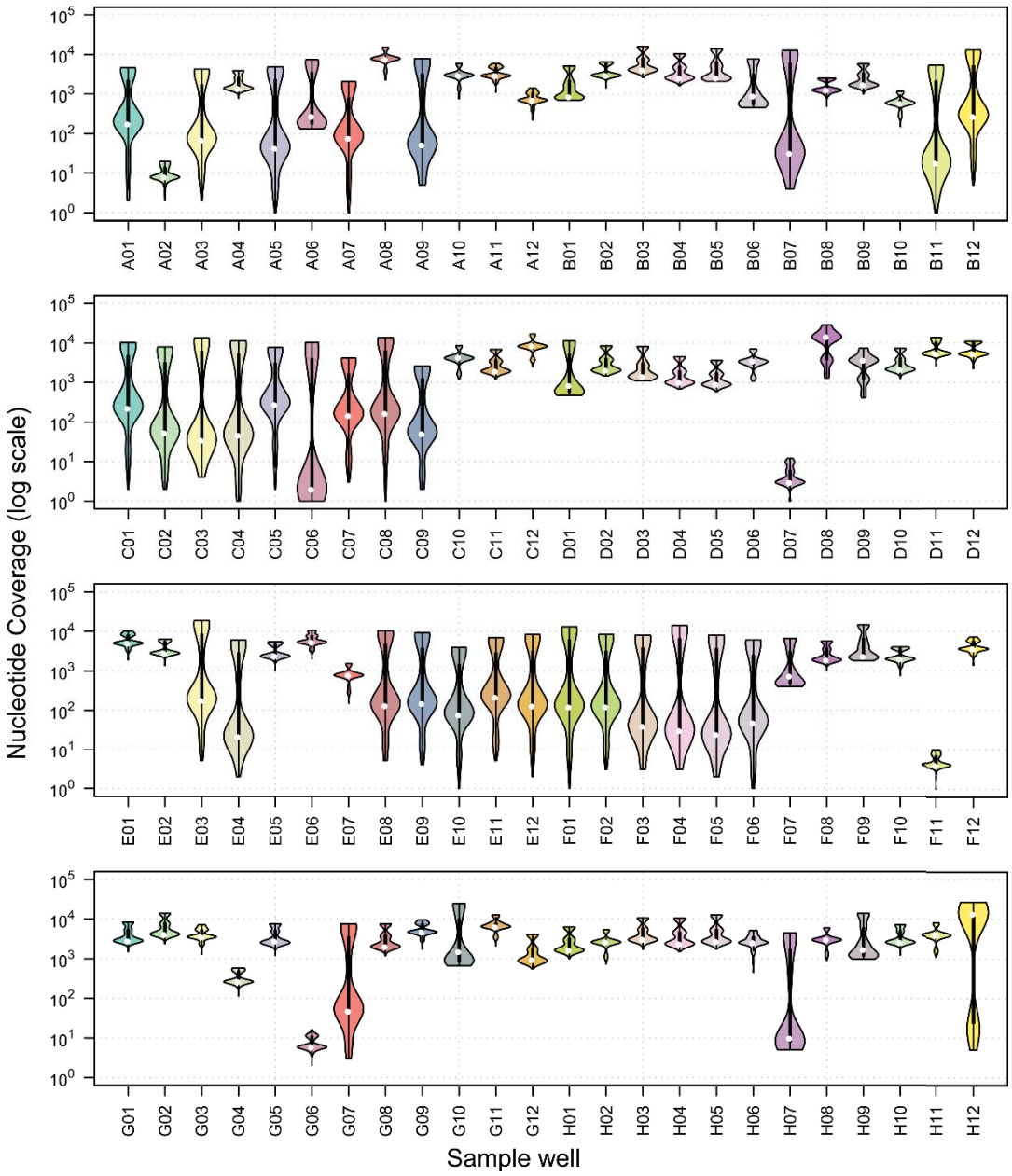


Figure E.6: Violin plots of nucleotide sequences coverage in log scale for each of 96 samples by plate location for a *tkt Campylobacter* MLST allele. The A01 to H12 represents wells in a 96-well plate, each of which contains a separate sample. The white dot is the median nucleotide sequences coverage and the dark black lines are the interquartile range along with whiskers.

Table E.2: The assigned allele numbers for each housekeeping gene of processed and processed trimmed sequences generated for each massMLST sample. The complete profile with defined sequence type (ST-best case based on non-ambiguous alleles) and clonal complex (CC) is highlighted in green. Red indicates the discrepant alleles between processed and processed trimmed sequences and yellow shows the extra allele determined based on the highest bit-scores for a given sample. (#) indicates the number of best bit-score determined for that sample, (*) denotes the possible ST and CC determined for some samples, "UD**" represents the ST/CC that could not be determined, and UA explains those unassigned ST or CC.

Sample ID	Processed sequences									ST	CC	Trimmed sequences								
	max Alleles	#	aspA	glnA	gltA	glyA	pgm	tkt	uncA			max Alleles	#	aspA	glnA	gltA	glyA	pgm	tkt	uncA
H1579	2	1 2	1	4	2	15	6	3	28 65	61*	61*	2	1 2	1	4	2	15	6	3	28 65
P1262a	2	1 2	4	7	10	247 4	1	7	1	45	45*	3	1 2	4	7	10	247 4	1	7	1
S709a	4	1 2 3 4	1	22 266 385 82	3	247 4	93	9	3	UD**	42*	4	1 2 3 4	1	22 266 385 82	3	247 4 60	93	9	3
S711a	3	1 2 3	32	107 24 438	30 83	242	225 321	77	28 65	UD**	UD**	6	1 2 3 4 5 6	159 215 273 32	188 260	222	242	316 416 5 64 7 87	164 77	65
713a	2	1 2	1	2	3	247 4	93	9	3	**UD****	**42***	3	1 2 3	1	2	3	247 4 60	93	9	3

S714a	2	1 2	1	2	59	247 4	5	9	3	**UD****	**42***	3	1 2 3	1	2	59	247 4 60	93	9
S715a	2	1 2	159 32	1	222	242	118	493 77	17 28	**UD****	**828***	3	1 2 3	159 215 32	115 300	222	242	118 209 413	164 77	42
S716a	2	1 2	282	2	3	247 4	5	9	3	**UD***	**42***	3	1 2 3	79	2	3	247 4 60	5	9	3
S718a	2	1 2	207	399	2	15	6	1	5	**UD****	**48***	1	1	2	399	2	15	6	1	5
S719a	2	1 2	196 33	251	44	242	253	44	28 65	**UD****	**828***	3	1 2 3	196 273 33	251	222 44	242	28	173	65
S723a	3	1 2 3	196 33	283	222 44	272 310 82	118	77	17 28	**UD***	**UD***	2	1 2	196 33	283	222 44	242	118 413	273 411	42
S745a	2	1 2	1	23	2	15	93	3	28 65	**UD***	**UD***	1	1	1	23	2	15	93	3	65
S752a	2	1 2	207	1	177 2	3	2	1	5	**1823***	**21**	2	1 2	2	1	177 2	3	2	1	5
S771a	2	1 2	33	283	127	242	118 189	44	28 65	**3222***	**UA***	3	1 2 3	196 33	283	127 222 44	242	118 413	44 493	28 65
S773a	2	1 2	2	1	5	3	293 349	3	5	**UD***	**21**	2	1 2	2	1	5	3	11 56	3	5
S774a	3	1 2 3	273 32 337	82	222	310 82	118	324	17 28	**UD***	**UD***	3	1 2 3	337	188 260	222	242	118 209 413	164 77	42

S775a	3	1 2 3	196 33	82	222 44	272 310 82	6	44	17 28	**UD***	**UD***	4	1 2 3 4	196 33	283	222 44	242	28 365 6	164 44 493 77	42
S776a	2	1 2	33	283	222 44	310 82	118	44	17	**3222***	**UA**	4	1 2 3 4	196 33	283	222 44	242	118 209 413	164 44 493 77	42
S777a	2	1 2	33	188 260	222	242	113 118	43	17 28	UD**	UD**	4	1 2 3 4	196 33	188 260	222	242	113 118 209 413	164 77	42
S788a	2	1 2	1	2	3	247 4	5	9	3	**UD****	**42***	3	1 2 3	1	2	3	247 4 60	5	9	3
S789a	2	1 2	1	2	3	247 4	5	9	3	**UD****	**42***	3	1 2 3	1	2	3	247 4 60	5	9	3
S794a	1	1 2	1	2	3	247 4	5	9	3	**UD****	**42***	1	1	1	2	3		5	9	3
S799a	1	1	2	1	12	3	2	1	5	**50**	**21**	1	1	2	1	12	3	2	1	5
S800b	2	1 2	1	2	3	247 4	5	9	3	**UD****	**42***	2	1 2	1	2	3	247 4	560	9	3
802a	2	1 2	1	4	2	15	6	3	17 28	**UD****	**61***	1	1	1	4	2	15	6	3	42
S805a	2	1 2	207	1	5	3 353	2	3	5	**190***	**21***	2	1 2	2	1	5 353	3	2	3	5
S829	2	1 2	1	14	2	15	6	1	17 28	**UD****	**61***	1	1	1	14	2	15	6	1	42

S833a	2	12	1	4	2	15	6	3	1728	UD**	61*	1	1	1	4	2	15	6	3	42
S837a	2	12	1	2	3	2474	5	9	3	UD**	42*	1	1	1	2	3	61	5	9	3
S840	2	12	1	2	3	2474	5	9	3	UD**	42*	1	1	1	2	3	61	5	9	3
S843	2	12	207	1	195	3353	2	1	5	UD**	21*	2	12	207	1		3353	2	1	5
S845	1	1	1	4	5	3	6	3	5	3425*	61	2	12	1	4	5	3	1066	3	5
S862a	1	1	2	1	21	3	2	255	5	UD**	21*	1	1	207	1	21	3	2		5
S869a	2	12	207	1	21	3353	2	1	5	UD**	21*	2	12	207	1	21	3353	2	1	5
S873a	1	1	2	1	21	3	2	1	5	53	21	1	1	2	1	21	3	2	1	5
S874a	1	1	1	4	5	3	6	3	17	UD**	61*	1	1	1	4	5	3	6	3	42
S876a	2	12	1	4	2	3629	6	3	65	UD**	61*	2	12	1	4	2	3353	6	3	5
S878a	2	12	1	4	2	134	6	3	1728	UD**	61*	2	12	1	4	2	134	6	3	42
S883a	4	1234	19633	22266 38582	22244	242	253	32444	1728	UD**	UD**	4	1234	19633	153188 260	22244	207208 242454	118413	16477	42
S886a	2	12	33	283	22244	242	118	44	2865	3222*	UD**	4	1234	19633	283	127	242	118413	16444 49377	28

S895a	2	1 2	196 33	82	222	310 82	253	44	17 28	**UD****	**UD****	4	1 2 3 4	196 33	283	222 44	242	253	164 44 493 77	42
S906a	2	1 2	32	188 260	222	310 82	118	44	17 28	**UD****	**UD****	4	1 2 3 4	159 215 32	188 260	222	242	118 209 413	164 44 493 77	42
S911a	5	1 2 3 4 5	33	22 266 340 385 82	222	310 82	113 209	43	17 28	**UD****	**UD****	4	1 2 3 4	196 33	188 260	222	242	113 118 209 413	164 493 77	42
S917a	4	1 2 3 4	126 248 261 283	22 266 385 82	222	82	104 189	35	28 65	**UD****	**UD****	5	1 2 3 4 5	157	188 260	222	242 272 310 354 82	118 413	164 333 35 493 77	42 65
S918a	4	1 2 3 4	33	218 262	222	310 82	104 189	333 35 493 77	36	**UD****	**UD****	5	1 2 3 4 5	196 33	188 260 38 437	222	242	118 209 413	164 333 35 493 77	36 42
S922a	1	1	1	4	2	15	6	3	5	**new**	**61**	1	1	1	4	2	15	6	3	5
S923a	5	1 2 3 4 5	1	4	2	10 119 121 380 89	6	1	5	**2776***	**61***	2	1 2	1	4	2	3 407	6		5

S925a	2	1 2	1	2	3	247 4	5	9	3	**UD****	**42***	3	1 2 3	1	2	3	247 4 60	5	9	3
S933a	2	1 2	1	2	3	247 4	5	9	3	**UD****	**42***	3	1 2 3	1	2	3	247 4 60	5	9	3
S939a	5	1 2 3 4 5	33	22 266 385 389 82	222	82	118	44	65	**UD****	**UD****	4	1 2 3 4	196 33	188 260	222	242	118 209 413	164 44 493 77	65
S946a	2	1 2	1	14	2	15	6	1	17 28	**UD****	**61***	3	1 2 3	1	14	2	15	6	1	17 28 42
S949a	4	1 2 3 4	196 33	22 266 385 82	222	310 82	113 209	43	17 28	**UD****	**UD****	4	1 2 3 4	196 33	188 260	222	242	113 118 209 413	164 493 77	42
S959a	2	1 2	1	4	2	15	6	3	17 28	**UD****	**61***	1	1	1	4	2	15	6	3	42
S961a	2	1 2	1	4	2	15	6	3	17 28	**UD****	**61***	1	1	1	4	2	15	6	3	42
S981b	3	1 2 3	9	2	4	15 337 62	4	5	6	**257***	**257***	3	1 2 3	9	2	4	15 337 62	4	5	6
S994a	3	1 2 3	69	399	2	104 130	118 209	317 40 443	17 28	**UD****	**61***	3	1 2 3	1	399	2	134 147 529	93		42
S996a	2	1 2	207	1	16 371	3 353	2	1	5	**21***	**21***	2	1 2	2	1	1	3 353	2	1	5

S1017a	2	1 2	4	7	14 43	247 4	1	7	1	**45***	**45***	3	1 2 3	4	7	14 43	247 4 60	1	7	1
S1025a	1	1	2	1	12	3	2	1	5	**50**	**21**	1	1	2	1	12	3	2	1	5
S1032a	4	1 2 3 4	33	82	222	242 272 310 82	113	43	17 28	**UD****	**828***	4	1 2 3 4	196 33	188 260	222	242	113 118 209 413	164 493 77	42
S1036a	3	1 2 3	33	399	222	119 406	113 118	164 493 77	17 28	**UD****	**828***	4	1 2 3 4	196 33	188 260	222	89	113 118 209 413	164 77	42
1037b	2	1 2	33	283	222 44	406	118 189	44	17 28	**3222***	**UA**	4	1 2 3 4	196 33	188	222 44	45 89	118 209 413	164 44 493 77	42
S1041a	1	1	207	1	1	380	2	1	5	**4337***	**21***	1	1	2	1	1	380	2	1	5
S1043a	4	1 2 3 4	159 215 32	22 266 385 82	30 83	406	118	43	17 28	**UD****	**UD****	4	1 2 3 4	159 215 32	22 266 385 82	222		118 209 413	164 77	42
S1053a	2	1 2	1	384	2	297 332	6	3	17 28	**61***	**61***	3	1 2 3	1	1	2	362 9	6	1 298 358	42
S1061a	1	1	2	1	1	380	2	1	5	**4337**	**21**	1	1	2	1	1	380	2	1	5
S1065a	1	1	2	1	1	380	2	1	5	**4337**	**21**	1	1	2	1	1	380	2	1	5
S1094a	2	1 2	1	4	2	15 297	6	3	17 28	**61***	**61***	2	1 2	1	4	2	15 297	6	3	42
S1117b	2	1 2	79	1	21	3	2	1	5	**UD****	**21***	2	1 2	79	1	21	3 353	2	1	5

S1123a	2	1 2	1	4	2	15	6	3	195 63	**61***	**61***	2	1 2	1	4	2	15	6		195 63
S1126a	2	1 2	207	1	21	3 353	2	1	5	**53***	**21***	2	1 2	2	1	21	3 353	2	1	5
S1144a	2	1 2	33	82	222	242	113	493 77	17 28	**UD****	**828***	4	1 2 3 4	273	188 260	222	242	113 118 209 413	273 411	42
S1146a	4	1 2 3 4	196 33	22 266 385 82	222	1 249	113	43	17 28	**UD****	**828***	4	1 2 3 4	196 33	188 260	222	89	113 118 209 413	164 493 77	42
S1147a	2	1 2	196 33	188 260	222	310 82	113 118	43	17 28	**UD****	**828***	4	1 2 3 4	196 33	188 260	222	242	113 118 209 413	164 493 77	42
S1150a	4	1 2 3 4	33	188 260	222	3	113 118	493 77	17 28	**UD****	**828***	4	1 2 3 4	196 33	188 260 38 437	222	242	113 118 209 413	164 493 77	42
S1152	2	1 2	1	82	356	242	118	333 35	36 42	**UD****	**UD****	5	1 2 3 4 5	1	188 260 38 437	356	32	118 209 413	164 333 35 493 77	370

S1155a	6	1 2 3 4 5 6	33	165 264 340 50 57 61	222 44	310 82	118	44	17 28	UD**	UD**	6	1 2 3 4 5 6	196 33	165 264 340 50 57 61	222 **44**	242	118 209 413	164 44 493 77	42
S1159a	2	1 2	1	4	2	15 297	6	3	5	**61***	**61***	2	1 2	1	4	2	15 297	6	3	5
S1165a	2	1 2	2	1	1	3 353	2	1	5	**21***	**21***	2	1 2	2	1	1	3 353	2	1	5
S1175a	2	1 2	33	283	222 44	310 82	118	44	28 65	**3222***	**UA**	3	1 2 3	196 33	283	222 44	242	118 209 413	44 493	28
S1176a	2	1 2	33	188 260	30 83	310 82	118	273 411	68	**UD****	**828***	6	1 2 3 4 5	273	188 260	222	242	118 209 413	171 206 325 373 44	42 68
S1179a	2	1 2	33	82	222	310 82	118	44	68	**UD****	**828***	4	1 2 3	196 33	188 260	222	242	118 209 413	164 44 493	42 68
S1182a	4	1 2 3 4	2	462	1	380	20 346 80 90	1	5	**UD****	**21***	2	1 2	264	462	1	380	20 346	1	5
1185a	1	1	2	1	1	380	2	1	5	**4337**	**21**	1	1	2	1	1	380	2	1	5
S1187a	1	1	264	1	1	380	2	1	5	**New**	**21**	1	1	264	1	1	380	2	1	5

S1189a	2	1 2	196 33	82	222 44	378	118	273 411	17	UD**	UD**	10	1 2 3 4 5 6 7 8 9 10	196 33	283	222 44	359 362 373 389 393 406 407 449 462 506	118 209 413	411	42
S1197a	2	1 2	207	1	5	3 353	2	1	5	19*	21*	2	1 2	2	1	5	3 353	2	1	5
S1201a	4	1 2 3 4	207	22 266 385 82	127	71	2	273	5	UD**	61*	4	1 2 3 4	2	22 266 385 82	127		2	273 411	5
S1205a	1	1	2	4	2	15	6	1	5	New	61	1	1	2	4	2	15	6	1	5
S1216a	2	1	1	4	2	15	6	3	17	61*	61*	2	1	1	4	2	15	6	3	42
		2				297			28				2							
S1219a	2	1 2	172	16 462	371	71	496	3	5	UD**	UD**	2	1 2	2	462	5	3 353	496	3	5
S1224a	2	1 2	1	4	2	134	6	3	17 28	UD**	61*	2	1 2	1	4	2	362 9	6	3	42

W717a	2	1 2	1	6	5	247 4	1 261	293	3	**UD****	**3640***	3	1 2 3	1	6	5	247 4 60	1 261	293	3
W717b	1	1	175	251	216	282	389	293	102	**2381**	**UA**	1	1	175	251	216	282	389	293	102
W725a	2	1 2	1	41	21	247 4	1 261	7	3	**UD****	**3640***	2	1 2	1	41	21	61	1 261	7	3
W725b	2	1 2	1	6	2	247 4	261	7	3	**UD****	**3640***	4	1 2 3 4	1	6	2	247 4 60 61	261	7	3

Bibliography

Abeywardena, H., Jex, A. R., Firestone, S. M., McPhee, S., Driessen, N., Koehler, A. V, … Gasser, R. B. (2013). Assessing calves as carriers of *Cryptosporidium* and *Giardia* with zoonotic potential on dairy and beef farms within a water catchment area by mutation scanning. Electrophoresis, 34(15), 2259–2267.

Abrahams, C. A., Agbodaze, D., Nakano, T., Afari, E. A., & Longmatey, H. E. K. (1990). Prevalence and Antibiogram of *Campylobacter-jejuni* in Domestic-Animals in Rural Ghana. Archives of Environmental Health, 45(1), 59–62.

Abulencia, C. B., Wyborski, D. L., Garcia, J. A., Podar, M., Chen, W., Chang, S. H., ... & Keller, M. (2006). Environmental whole-genome amplification to access microbial populations in contaminated sediments. Applied and environmental microbiology, 72(5), 3291-3301.

Abulreesh H.H., Paget T.A., Goulder R. (2006). *Campylobacter* in waterfowl and aquatic environments: incidence and methods of detection. Environ Sci Technol, 40, 7122-7131.

Adegbola, R. A., Demba, E., Deveer, G., & Todd, J. (1994). *Cryptosporidium* Infection in Gambian Children Less-Than 5 Years of Age. Journal of Tropical Medicine and Hygiene, 97(2), 103–107.

Adesiyun, A. A., Kaminjolo, J. S., Loregnard, R., & Kitsonpiggott, W. (1992). *Campylobacter* Infections in Calves, Piglets, Lambs and Kids in Trinidad. British Veterinary Journal, 148(6), 547–556.

Adessi, C., Matton, G., Ayala, G., Turcatti, G., Mermod, J. J., Mayer, P., & Kawashima, E. (2000). Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. Nucleic acids research, 28(20), e87-e87.

Adhikari, B., Connolly, J. H., Madie, P., & Davies, P. R. (2004). Prevalence and clonal diversity of *Campylobacter jejuni* from dairy farms and urban sources. New Zealand Veterinary Journal, 52(6), 378–383.

Adler, D. (2005). Vioplot: Violin plot. R package version 0.2.

Aiello, A. E., Xiao, L. H., Limor, J. R., Liu, C., Abrahamsen, M. S., & Lal, A. A. (1999). Microsatellite analysis of the human and bovine genotypes of *Cryptosporidium parvum*. Journal of Eukaryotic Microbiology, 46(5), 46s–47s.

Al Mawly, J., Prattley, D., French, N. P., Lopez-Villalobos, N., Hedgespeth, B., & Grinberg, A. (2013). Utility of halofuginone lactate for the prevention of natural cryptosporidiosis

of calves, in the presence of co-infection with rotavirus and Salmonella Typhimurium. Veterinary Parasitology, 197(1-2), 59–67.

Al Mawly, J., Grinberg, A., Prattley, D., Moffat, J., & French, N. (2014). Prevalence of endemic enteropathogens of calves in New Zealand dairy farms. New Zealand Veterinary Journal, 1–6.

Al Mawly, J., Grinberg, A., Prattley, D., Moffat, J., Marshall, J., & French, N. (2015a). Risk factors for neonatal calf diarrhoea and enteropathogen shedding in New Zealand dairy farms. The Veterinary Journal, 203(2), 155–160.

Al Mawly, J., Grinberg, A., Velathanthiri, N., & French, N. (2015b). Cross sectional study of prevalence, genetic diversity and zoonotic potential of *Cryptosporidium parvum* cycling in New Zealand dairy farms. Parasites & Vectors, 8(1), 1–7.

Aldeyarbi, H. M., & Karanis, P. (2014). Electron microscopic study of in vitro axenic development of *Cryptosporidium parvum*. In Fifth International *Giardia* and *Cryptosporidium* Conference. Uppsala, Sweden. May (pp. 27-30).

Alfredson, D. A., & Korolik, V. (2007). Antibiotic resistance and resistance mechanisms in *Campylobacter jejuni* and *Campylobacter coli*. FEMS Microbiology Letters, 277(2), 123-132.

Aloisio, F., Filippini, G., Antenucci, P., Lepri, E., Pezzotti, G., Cacciò, S. M., & Pozio, E. (2006). Severe weight loss in lambs infected with *Giardia duodenalis* assemblage B. Veterinary parasitology, 142(1), 154-158.

Altekruse, S. F., Hunt, J. M., Tollefson, L. K., & Madden, J. M. (1994). Food and Animal Sources of Human *Campylobacter jejuni* Infection. Journal of the American Veterinary Medical Association, 204(1), 57–61.

Alves, M., Matos, O., & Antunes, F. (2003). Microsatellite analysis of *Cryptosporidium hominis* and *C. parvum* in Portugal: A preliminary study. Journal of Eukaryotic Microbiology, 50, 529–530.

Alves, M., Xiao, L. H., Sulaiman, I., Lal, A. A., Matos, O., & Antunes, F. (2003). Subgenotype analysis of *Cryptosporidium* isolates from humans, cattle, and zoo ruminants in Portugal. Journal of Clinical Microbiology, 41(6), 2744–2747.

Alves, M., Xiao, L. H., Antunes, F., & Matos, O. (2006). Distribution of *Cryptosporidium* subtypes in humans and domestic and wild ruminants in Portugal. Parasitology Research, 99(3), 287–292.

Amadi, B., Kelly, P., Mwiya, M., Mulwazi, E., Sianongo, S., Changwe, F., … Chintu, C. (2001). Intestinal and systemic infection, HIV, and mortality in Zambian children with persistent diarrhea and malnutrition. Journal of Pediatric Gastroenterology and Nutrition, 32(5), 550–554.

Amar, C., Pedraza-Diaz, S., & McLauchlin, J. (2001). Extraction and genotyping of *Cryptosporidium parvum* DNA from fecal smears on glass slides stained conventionally for direct microscope examination. Journal of Clinical Microbiology, 39(1), 401–403.

Amar, C. F. L., Chalmers, R. M., Elwin, K., Tynan, P., & McLauchlin, J. (2002). Blinded evaluation of DNA extraction and genotyping of stained *Cryptosporidium* on glass slides. Letters in Applied Microbiology, 35(6), 486–488.

Amar, C. F. L., East, C. L., Gray, J., Iturriza-Gomara, M., Maclure, E. A., & McLauchlin, J. (2007). Detection by PCR of eight groups of enteric pathogens in 4,627 faecal samples: re-examination of the English case-control Infectious Intestinal Disease Study (1993-1996). European Journal of Clinical Microbiology & Infectious Diseases, 26(5), 311–323.

Amer, S., Zidan, S., Adamu, H., Ye, J., Roellig, D., Xiao, L., Feng, Y. (2013). Prevalence and characterization of *Cryptosporidium* spp. in dairy cattle in Nile River delta provinces, Egypt. Exp Parasitol 135, 518-523.

Anceno, A. J., Katayama, H., Houpt, E. R., Chavalitshewinkoon-Petmitr, P., Chuluun, B., & Shipin, O. V. (2007). IMS-free DNA extraction for the PCR-based quantification of *Cryptosporidium parvum* and *Giardia lamblia* in surface and waste water. International journal of environmental health research, 17(4), 297-310.

Anderson, B. C. (1998). Cryptosporidiosis in bovine and human health. Journal of Dairy Science, 81(11), 3036–3041.

Angus, K. W., Appleyard, W. T., Menzies, J. D., Campbell, I., & Sherwood, D. (1982). An Outbreak of Diarrhea Associated with Cryptosporidiosis in Naturally Reared Lambs. Veterinary Record, 110(6), 129–130.

Anonymous. (2001). Regional Plan for Beds of Rivers and Lakes and Associated Activities: Horizon Regional Council. Report no. 21/EXT/441.

Anonymous. (2007). Manawatu raw water pilot survey- A report. Palmerston North: Massey University.

Anonymous. (2011). A Report Prepared for Gastro-Intestinal Protozoa Research and Services. Palmerston North: Massey University.

Anonymous. (2011a). Horizon Regional Council. http://www.horizons.govt.nz/, accessed on 2 January 2012.

Anonymous. (2011b). Horizon Regional Council. http://www.tararuadc.govt.nz/Home, accessed on 2 January 2012.

Anonymous. (2013). A Report Prepared for Gastro-Intestinal Protozoa Research and Services. Palmerston North: Massey University.

Arrowood, M. J., & Sterling, C. R. (1989). Comparison of Conventional Staining Methods and Monoclonal Antibody-Based Methods for *Cryptosporidium* Oocyst Detection. Journal of Clinical Microbiology, 27(7), 1490–1495.

Assure Quality Ltd (2011). National Agribase Database: Assure Quality New Zealand Limited.

Atherholt, T. B., LeChevallier, M. W., Norton, W. D., & Rosen, J. S. (1998). Effect of rainfall of *Giardia* and Crypto. American Water Works Association. Journal, 90(9), 66.

Atwill, E. R., Harp, J. A., Jones, T., Jardon, P. W., Checel, S., & Zylstra, M. (1998). Evaluation of periparturient dairy cows and contact surfaces as a reservoir of *Cryptosporidium parvum* for calfhood infection. American Journal of Veterinary Research, 59(9), 1116–1121.

Bae, W., Kaya, K. N., Hancock, D. D., Call, D. R., Park, Y. H., & Besser, T. E. (2005). Prevalence and antimicrobial resistance of thermophilic *Campylobacter* spp. from cattle farms in Washington State. Applied and Environmental Microbiology, 71(1), 169–174.

Baffone, W., Casaroli, A., Citterio, B., Pierfelici, L., Campana, R., Vittoria, E., … Donelli, G. (2006). *Campylobacter jejuni* loss of culturability in aqueous microcosms and ability to resuscitate in a mouse model. International Journal of Food Microbiology, 107(1), 83–91.

Bagshaw, C. S. (2002). Factors influencing direct deposition of cattle faecal material in riparian zones. Ministry of Agriculture and Forestry.

Baig, F., Knopp, M., & Rajabally, Y. A. (2012). Diagnosis, epidemiology and treatment of inflammatory neuropathies. British Journal of Hospital Medicine, 73(7), 380–385.

Baker, I.K., & Carbonell, P.L. (1974). *Cryptosporidium agni* sp. n. from lambs, and *Cryptosporidium bovis* sp.n. from a calf, with observations on the oocyst. Z. Parasitenkd. 44, 289–298.

Baker, M., Wilson, N., Ikram, R., Chambers, S., Shoemack, P., & Cook, G. (2006). Regulation of chicken contamination is urgently needed to control New Zealand's serious campylobacteriosis epidemic. The New Zealand Medical Journal (Online), 119(1243).

Baker, M., Sneyd, E., Wilson, N. (2007a). Is the major increase in notified campylobacteriosis in New Zealand real? Epidemiol Infect 135:163-170.

Baker, M., Wilson, N., Edwards, R. (2007b). *Campylobacter* infection and chicken: an update on New Zealand's largest 'common source outbreak'. Journal of the New Zealand Medical Association 120.

Baker, M.G., Kvalsvig, A., Zhang, J., Lake, R., Sears, A., Wilson, N. (2012). Declining Guillain-Barré Syndrome after Campylobacteriosis Control, New Zealand, 1988–2010. Emerg Infect Dis 18:226.

Balatbat, A. B., Jordan, G. W., Tang, Y. J., & Silva, J. (1996). Detection of *Cryptosporidium parvum* DNA in human feces by nested PCR. Journal of Clinical Microbiology, 34(7), 1769–1772.

Barer, M. R., & Harwood, C. R. (1999). Bacterial viability and culturability. Advances in microbial physiology, 41, 93-137.

Baron, J. L., Harris, J. K., Holinger, E. P., Duda, S., Stevens, M. J., Robertson, C. E., … Stout, J. E. (2015). Effect of monochloramine treatment on the microbial ecology of Legionella and associated bacterial populations in a hospital hot water system. Systematic and Applied Microbiology, 38(3), 198–205.

Baxby, D., Hart, C. A., & Taylor, C. (1983). Human Cryptosporidiosis - a Possible Case of Hospital Cross Infection. British Medical Journal, 287(6407), 1760–1761.

Becher, K. A., Robertson, I. D., Fraser, D. M., Palmer, D. G., & Thompson, R. C. A. (2004). Molecular epidemiology of *Giardia* and *Cryptosporidium* infections in dairy calves originating from three sources in Western Australia. Veterinary Parasitology, 123(1-2), 1–9.

Berk, R. N., Wall, S. D., Mcardle, C. B., Mccutchan, J. A., Clemett, A. R., Rosenblum, J. S., … Megibow, A. J. (1984). Cryptosporidiosis of the Stomach and Small-Intestine in Patients with Aids. American Journal of Roentgenology, 143(3), 549–554.

Bezirtzoglou, C., Dekas, K., & Charvalos, E. (2011). Climate changes, environment and infection: facts, scenarios and growing awareness from the public health community within Europe. Anaerobe, 17(6), 337-340.

Bhat, S. A., Juyal, P. D., & Singla, L. D. (2012). Prevalence of Cryptosporidiosis in Neonatal Buffalo Calves in Ludhiana District of Punjab, India. Asian Journal of Animal and Veterinary Advances, 7(6), 512–520.

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., ... & Karch, H. (2011). Characterisation of the *Escherichia coli* strain associated with an outbreak of haemolytic uraemic syndrome in Germany, 2011: a microbiological study. The Lancet infectious diseases, 11(9), 671-676.

Bifulco, J. M., & Schaefer, F. W. (1993). Antibody-magnetite method for selective concentration of *Giardia lamblia* cysts from water samples. Applied and environmental microbiology, 59(3), 772-776.

Bjorkman, C., Svensson, C., Christensson, B., & de Verdier, K. (2003). *Cryptosporidium parvum* and *Giardia intestinalis* in calf diarrhoea in Sweden. Acta Veterinaria Scandinavica, 44(3-4), 145–152.

Black, R. E., Levine, M. M., Clements, M. L., Hughes, T. P., & Blaser, M. J. (1988). Experimental *Campylobacter jejuni* infection in humans. Journal of infectious diseases, 157(3), 472-479.

Blewett, D. A., Wright, S. E., Casemore, D. P., Booth, N. E., & Jones, C. E. (1993). Infective dose size studies on *Cryptosporidium parvum* using gnotobiotic lambs. Water Science and Technology, 27(3-4), 61-64.

Blodgett, R. (2010). FDA, Bacterial Analytical Manual, Appendix 2 Most Probable Number from Serial Dilutions.

Boers, S. A., Van der Reijden, W. A., & Jansen, R. (2012). High-throughput multilocus sequence typing: bringing molecular typing to the next level. PLoS One, 7(7), e39630.

Bolton, F. J., Coates, D., Hinchliffe, P. M., & Robertson, L. (1983). Comparison of Selective Media for Isolation of *Campylobacter-jejuni*-Coli. Journal of Clinical Pathology, 36(1), 78–83.

Bomfim, T. C. B., Huber, F., Gomes, R. S., & Alves, L. L. (2005). Natural infection by *Giardia* sp. and *Cryptosporidium* sp. in dairy goats, associated with possible risk factors of the studied properties. Veterinary parasitology, 134(1), 9-13.

Bonadonna, L., Briancesco, R., Cataldo, C., Divizia, M., Donia, D., & Pana, A. (2002). Fate of bacterial indicators, viruses and protozoan parasites in a wastewater multi-component treatment system. Microbiologica, 25(4), 413–420.

Borchardt, M. A., Bradbury, K. R., Gotkowitz, M. B., Cherry, J. A., & Parker, B. L. (2007). Human enteric viruses in groundwater from a confined bedrock aquifer. Environmental science & technology, 41(18), 6606-6612.

Bouzid, M., Hunter, P. R., Chalmers, R. M., & Tyler, K. M. (2013). *Cryptosporidium* pathogenicity and virulence. Clinical microbiology reviews, 26(1), 115-134.

Bowman, D. D. (2009). Essential veterinary education in water-borne transmission of disease. Revue Scientifique Et Technique-Office International Des Epizooties, 28(2), 589–596.

Brady M. T., Oleske J. M., Williams P. L., Elgie C., Mofenson L. M., Dankner W. M., Van Dyke R. B. (2010). Declines in Mortality Rates and Changes in Causes of Death in HIV-1-Infected Children During the HAART Era. Jaids-J Acq Imm Def, 53, 86-94.

Britton, E., Hales, S., Venugopal, K., & Baker, M. G. (2010). The impact of climate variability and change on cryptosporidiosis and giardiasis rates in New Zealand. Journal of water and health, 8(3), 561-571.

Bronowski, Christina, Chloe E. James, and Craig Winstanley. "Role of environmental survival in transmission of *Campylobacter jejuni*." FEMS microbiology letters 356, no. 1 (2014): 8-19.

Brook, E. J., Hart, C. A., French, N. P., & Christley, R. M. (2009). Molecular epidemiology of *Cryptosporidium* subtypes in cattle in England. The Veterinary Journal, 179(3), 378-382.

Brown, P. E., Christensen, O. F., Clough, H. E., Diggle, P. J., Hart, C. A., Hazel, S., … French, N. P. (2004). Frequency and spatial distribution of environmental *Campylobacter* spp. Applied and Environmental Microbiology, 70(11), 6501–6511.

Brzuszkiewicz, E., Thurmer, A., Schuldes, J., Leimbach, A., Liesegang, H., Meyer, F. D., Boelter, J., Petersen, H., Gottschalk, G. & Daniel, R. (2011). Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the

emergence of a new pathotype: Entero-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC). Archives of Microbiology, 193, 883-891.

Budu-Amoako, E., Greenwood, S. J., Dixon, B. R., Barkema, H. W., & McClure, J. T. (2012a). *Giardia* and *Cryptosporidium* on Dairy Farms and the Role these Farms May Play in Contaminating Water Sources in Prince Edward Island, Canada. Journal of Veterinary Internal Medicine, 26(3), 668–673.

Budu-Amoako, E., Greenwood, S. J., Dixon, B. R., Barkema, H. W., & McClure, J. T. (2012b). Occurrence of *Cryptosporidium* and *Giardia* on beef farms and water sources within the vicinity of the farms on Prince Edward Island, Canada. Veterinary Parasitology, 184(1), 1–9.

Budu-Amoako, E., Greenwood, S. J., Dixon, B. R., Sweet, L., Ang, L., Barkema, H. W., & McClure, J. T. (2012c). Molecular Epidemiology of *Cryptosporidium* and *Giardia* in Humans on Prince Edward Island, Canada: Evidence of Zoonotic Transmission From Cattle. Zoonoses and Public Health, 59(6), 424–433.

Bukhari, Z., & Smith, H. V. (1995). Effect of 3 Concentration Techniques on Viability of *Cryptosporidium-parvum* Oocysts Recovered from Bovine Feces. Journal of Clinical Microbiology, 33(10), 2592–2595.

Buswell, C. M., Herlihy, Y. M., Lawrence, L. M., McGuiggan, J. T. M., Marsh, P. D., Keevil, C. W., & Leach, S. A. (1998). Extended survival and persistence of *Campylobacter* spp. in water and aquatic biofilms and their detection by immunofluorescent-antibody and -rRNA staining. Applied and Environmental Microbiology, 64(2), 733–741.

Bybee, S. M., Bracken-Grissom, H., Haynes, B. D., Hermansen, R. A., Byers, R. L., Clement, M. J., ... & Crandall, K. A. (2011). Targeted amplicon sequencing (TAS): a scalable next-gen approach to multilocus, multitaxa phylogenetics. Genome biology and evolution, 3, 1312-1323.

Caccio, S., Homan, W., Camilli, R., Traldi, G., Kortbeek, T., & Pozio, E. (2000). A microsatellite marker reveals population heterogeneity within human and animal genotypes of *Cryptosporidium parvum*. Parasitology, 120, 237–244.

Caccio, S., Spano, F., & Pozio, E. (2001). Large sequence variation at two microsatellite loci among zoonotic (genotype C) isolates of *Cryptosporidium parvum*. International Journal for Parasitology, 31(10), 1082–1086.

Caccio, S. M., Thompson, R. C. A., McLauchlin, J., & Smith, H. V. (2005). Unravelling *Cryptosporidium* and *Giardia* epidemiology. Trends in Parasitology, 21(9), 430–437.

Caccio, S. M., Rinaldi, L., Cringoli, G., Condoleo, R., & Pozio, E. (2007). Molecular identification of *Cryptosporidium parvum* and *Giardia duodenalis* in the Italian water buffalo (Bubalus bubalis). Veterinary parasitology, 150(1), 146-149.

Caccio, S. M., & Ryan, U. (2008). Molecular epidemiology of giardiasis. Molecular and Biochemical Parasitology, 160(2), 75–80.

Cacciò, S. M., & Widmer, G. (Eds.). (2013). *Cryptosporidium*: parasite and disease. Springer Science & Business Media.

Cacciò, S. M., & Putignani, L. (2014). Epidemiology of human cryptosporidiosis. In *Cryptosporidium*: parasite and disease (pp. 43-79). Springer Vienna.

Cama, V., Gilman, R. H., Vivar, A., Ticona, E., Ortega, Y., Bern, C., & Xiao, L. H. (2006). Mixed *Cryptosporidium* infections and HIV. Emerging Infectious Diseases, 12(6), 1025–1028.

Cama, V. A., Bern, C., Roberts, J., Cabrera, L., Sterling, C. R., Ortega, Y., … Xiao, L. H. (2008). *Cryptosporidium* species and subtypes and clinical manifestations in children, Peru. Emerging Infectious Diseases, 14(10), 1567–1574.

Campana, M. G., García, N. R., Rühli, F. J., & Tuross, N. (2014). False positives complicate ancient pathogen identifications using high-throughput shotgun sequencing. BMC Research Notes, 7(1), 111.

Cardarelli-Leite, P., Blom, K., Patton, C. M., Nicholson, M. A., Steigerwalt, A. G., Hunter, S. B., … Swaminathan, B. (1996). Rapid identification of *Campylobacter* species by restriction fragment length polymorphism analysis of a PCR-amplified fragment of the gene coding for 16S rRNA. Journal of Clinical Microbiology, 34(1), 62–67.

Cardenas, E., & Tiedje, J. M. (2008). New tools for discovering and characterizing microbial diversity. Current Opinion in Biotechnology, 19(6), 544-549.

Carter P. E., McTavish S. M., Brooks H. J. L., Campbell D., Collins-Emerson J. M., Midwinter A. C., French N. P. (2009). Novel clonal complexes with an unknown animal reservoir dominate *Campylobacter jejuni* isolates from river water in New Zealand. Appl. Environ. Microbiol. 75:6038-6046.

Casemore, D. P., Armstrong, M., & Sands, R. L. (1985). Laboratory Diagnosis of Cryptosporidiosis. Journal of Clinical Pathology, 38(12), 1337–1341.

Castro-Hermida, J. A., Gonzalez-Losada, Y. A., & Ares-Mazas, E. (2002). Prevalence of and risk factors involved in the spread of neonatal bovine cryptosporidiosis in Galicia (NW Spain). Veterinary Parasitology, 106(1), 1–10.

Castro-Hermida, J. A., Delafosse, A., Pors, I., Ares-Mazás, E., & Chartier, C. (2005). *Giardia duodenalis* and *Cryptosporidium parvum* infections in adult goats and their implications for neonatal kids. Veterinary Record: Journal of the British Veterinary Association, 157(20).

Castro-Hermida, J. A., García-Presedo, I., Almeida, A., González-Warleta, M., Da Costa, J. M. C., & Mezo, M. (2011). *Cryptosporidium* spp. and *Giardia duodenalis* in two areas of Galicia (NW Spain). Science of the total environment, 409(13), 2451-2459.

Centers for Disease Control and Prevention (CDC). (2013a). Incidence and trends of infection with pathogens transmitted commonly through food-foodborne diseases active surveillance network, 10 US sites, 1996-2012. MMWR. Morbidity and mortality weekly report, 62(15), 283.

Centers for Disease Control and Prevention (CDC). (2013b). Summary of Notifiable Diseases —United States, 2011. MMWR. Morbidity and mortality weekly report, 60(53), 26-40.

Centers for Disease Control and Prevention (CDC). (2014). Incidence and trends of infection with pathogens transmitted commonly through foodborne diseases active surveillance network, 10 US sites, 2006-2013. MMWR. Morbidity and mortality weekly report, 63(15), 328.

Chaisowwong W, Kusumoto A, Hashimoto M, Harada T, Maklon K, Kawamoto K. 2011. Physiological characterization of *Campylobacter jejuni* under cold stresses conditions: its potential for public threat. The Journal of Veterinary Medical Science.

Chalmers, R. M., Ferguson, C., Caccio, S., Gasser, R. B., El-Osta, Y. G. A., Heijnen, L., … Stevens, M. (2005a). Direct comparison of selected methods for genetic categorisation of *Cryptosporidium parvum* and *Cryptosporidium hominis* species.International Journal for Parasitology, 35(14), 1615.

Chalmers, R. M., Thomas, A. L., Butler, B. A., & Morel, M. C. G. D. (2005b). Identification of *Cryptosporidium parvum* genotype 2 in domestic horses. Veterinary Record, 156(2), 49–50.

Chalmers, R. M., & Davies, A. P. (2010). Minireview: Clinical cryptosporidiosis. Experimental Parasitology, 124(1), 138–146.

Chalmers, R. M., Campbell, B. M., Crouch, N., Charlett, A., & Davies, A. P. (2011). Comparison of diagnostic sensitivity and specificity of seven *Cryptosporidium* assays used in the UK. Journal of Medical Microbiology, 60(11), 1598–1604.

Chalmers, R.M., Katzer, F. (2013). Looking for *Cryptosporidium*: the application of advances in detection and diagnosis. Trends Parasitol.

Chang'a, J. S., Robertson, L. J., Mtambo, M. M. A., Mdegela, R. H., Loken, T., & Reksen, O. (2011). Unexpected results from large-scale cryptosporidiosis screening study in calves in Tanzania. Annals of Tropical Medicine and Parasitology, 105(7), 513–519.

Chanyalew, Y., Asrat, D., Amavisit, P., & Loongyai, W. (2013). Prevalence and Antimicrobial Susceptibility of Thermophilic *Campylobacter* Isolated from Sheep at Debre Birhan, North-Shoa, Ethiopia. Kasetsart Journal (Natural Science), 47, 551-560.

Châtre, P., Haenni, M., Meunier, D., Botrel, M. A., Calavas, D., & Madec, J. Y. (2010). Prevalence and antimicrobial resistance of *Campylobacter jejuni* and *Campylobacter coli* isolated from cattle between 2002 and 2006 in France. Journal of Food Protection®, 73(5), 825-831.

Chute, C. G., Smith, R. P., & Baron, J. A. (1987). Risk factors for endemic giardiasis. American journal of public health, 77(5), 585-587.

Claesson, M. J., Wang, Q., O'Sullivan, O., Greene-Diniz, R., Cole, J. R., Ross, R. P., & O'Toole, P. W. (2010). Comparison of two next-generation sequencing technologies for resolving

highly complex microbiota composition using tandem variable 16S rRNA gene regions. Nucleic acids research, 38(22), e200-e200.

Clark, C. G., & Diamond, L. S. (2002). Methods for cultivation of luminal parasitic protists of clinical importance. Clinical microbiology reviews, 15(3), 329-341.

Clark, C. G., Price, L., Ahmed, R., Woodward, D. L., Melito, P. L., Rodgers, F. G., … Ellis, A. (2003). Characterization, of waterborne outbreak-associated *Campylobacter jejuni*, Walkerton, Ontario. Emerging Infectious Diseases, 9(10), 1232–1241.

Clavel, A., Arnal, A. C., Sanchez, E. C., Cuesta, J., Letona, S., Amiguet, J. A., … GomezLus, R. (1996). Respiratory cryptosporidiosis: Case series and review of the literature. Infection, 24(5), 341–346.

Cleaveland, S., Laurenson, M. K., & Taylor, L. H. (2001). Diseases of humans and their domestic mammals: pathogen characteristics, host range and the risk of emergence. Philosophical Transactions of the Royal Society B: Biological Sciences, 356(1411), 991–999.

Close, M., Dann, R., Ball, A., Pirie, R., Savill, M., & Smith, Z. (2008). Microbial groundwater quality and its health implications for a border-strip irrigated dairy farm catchment, South Island, New Zealand. Journal of Water and Health, 6(1), 83–98.

Coklin, T., Farber, J., Parrington, L., & Dixon, B. (2007). Prevalence and molecular characterization of *Giardia duodenalis* and *Cryptosporidium* spp. in dairy cattle in Ontario, Canada. Veterinary parasitology, 150(4), 297-305.

Coklin, T., Uehlinger, F. D., Farber, J. M., Barkema, H. W., O'Handley, R. M., & Dixon, B. R. (2009). Prevalence and molecular characterization of *Cryptosporidium* spp. in dairy calves from 11 farms in Prince Edward Island, Canada. Veterinary parasitology, 160(3), 323-326.

Collins, D. M., & De Lisle, G. W. (1985). Letters to the editor: Typing of '*Campylobacter* fetus fetus isolated from sheep abortions in New Zealand.

Collins, R., Ross, C., & Donnison, A. (2002). Riparian attenuation of faecal microbes. National Institute of Water and Atmospheric Research Ltd (NIWA) Client Report MAF02202 to MAF (New Zealand).

Collins, D. (2003). Pretesting survey instruments: an overview of cognitive methods. Quality of life research, 12(3), 229-238.

Collins, D. B. G., Bras, R. L., & Tucker, G. E. (2004). Modeling the effects of vegetation-erosion coupling on landscape evolution. Journal of Geophysical Research: Earth Surface, 109(F3).

Collins, R., Mcleod, M., Hedley, M., Donnison, A., Close, M., Hanly, J., … Matthews, L. (2007). Best management practices to mitigate faecal contamination by livestock of New Zealand waters. New Zealand Journal of Agricultural Research, 50(2), 267–278.

Constable P. (2014): Cryptosporidiosis. In: The Merck Veterinary Manual Online. Retrieved June 8, 2014, http://www.merckvetmanual.com/mvm/digestive_system/cryptosporidiosis/overview_of_cryptosporidiosis.html?qt=cryptosporidium&alt=sh.

Cools, I., Uyttendaele, M., Caro, C., D'Haese, E., Nelis, H. J., & Debevere, J. (2003). Survival of *Campylobacter jejuni* strains of different origin in drinking water. Journal of Applied Microbiology, 94(5), 886–892.

Cooper, C. M. (1993). Biological Effects of Agriculturally Derived Surface Water Pollutants on Aquatic Systems—A Review. J. Environ. Qual., 22(3), 402-408.

Corry, J.E.L., Atabay, H.I., 2001. Poultry as a source of *Campylobacter* and related organisms. J Appl Microbiol, 90, 96s-114s.

Cortés, C., de la Fuente, R., Contreras, A., Sánchez, A., Corrales, J. C., Martínez, S., & Orden, J. A. (2006). peer reviewed. Irish veterinary journal, 59(7), 391-393.

Coupe, S., Sarfati, C., Hamane, S., & Derouin, F. (2005). Detection of *Cryptosporidium* and identification to the species level by nested PCR and restriction fragment length polymorphism. Journal of Clinical Microbiology, 43(3), 1017–1023.

Cox, M. P., Peterson, D. A., & Biggs, P. J. (2010). SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data. BMC bioinformatics, 11(1), 1.

Craik, S. A., Weldon, D., Finch, G. R., Bolton, J. R., & Belosevic, M. (2001). Inactivation of *Cryptosporidium parvum* oocysts using medium- and low-pressure ultraviolet radiation. Water Research, 35(6), 1387–1398.

Craun, G. F., Brunkard, J. M., Yoder, J. S., Roberts, V. A., Carpenter, J., Wade, T., Calderon, R. L., Roberts, J. M., Beach, M. J.& Roy, S. L. (2010). Causes of outbreaks associated with drinking water in the United States from 1971 to 2006. Clinical Microbiology Reviews, 23(3), 507-528.

Current, W. L., Reese, N. C., Ernst, J. V, Bailey, W. S., Heyman, M. B., & Weinstein, W. M. (1983). Human Cryptosporidiosis in Immunocompetent and Immunodeficient Persons - Studies of an Outbreak and Experimental Transmission. New England Journal of Medicine, 308(21), 1252–1257.

Curriero, F. C., Patz, J. A., Rose, J. B., & Lele, S. (2001). The association between extreme precipitation and waterborne disease outbreaks in the United States, 1948-1994. American journal of public health, 91(8), 1194-1199.

Danciger, M., & Lopez, M. (1975). Numbers of *Giardia* in the feces of infected children. The American journal of tropical medicine and hygiene, 24(2), 237-242.

Dasti, J. I., Tareen, A. M., Lugert, R., Zautner, A. E., & Groß, U. (2010). *Campylobacter jejuni*: a brief overview on pathogenicity-associated factors and disease-mediating mechanisms. International Journal of Medical Microbiology, 300(4), 205-211.

Davies, C. M., Kaucner, C., Deere, D., & Ashbolt, N. J. (2003). Recovery and enumeration of *Cryptosporidium parvum* from animal fecal matrices. Applied and environmental microbiology, 69(5), 2842-2847.

Davies-Colley, R. J., Nagels, J. W., Smith, R. A., Young, R. G., & Phillips, C. J. (2004). Water quality impact of a dairy cow herd crossing a stream. New Zealand Journal of Marine and Freshwater Research, 38(4), 569-576.

Davies-Colley, R., Elizabeth L., and John N. (2008). "Stormflow-dominated loads of faecal pollution from an intensively dairy-farmed catchment." Water Science & Technology 57 (10).

Debruyne, L., Broman, T., Bergström, S., Olsen, B., On, S. L., & Vandamme, P. (2010). *Campylobacter volucris* sp. nov., isolated from black-headed gulls (Larus ridibundus). International journal of systematic and evolutionary microbiology, 60(8), 1870-1875.

Delafosse, A., Castro-Hermida, J. A., Baudry, C., Ares-Mazas, E., & Chartier, C. (2006). Herd-level risk factors for *Cryptosporidium* infection in dairy-goat kids in western France. Preventive Veterinary Medicine, 77(1-2), 109–121.

Dempster, R. P., Wilkins, M., Green, R. S., & de Lisle, G. W. (2011). Serological survey of Toxoplasma gondii and *Campylobacter fetus fetus* in sheep from New Zealand. New Zealand Veterinary Journal, 59(4), 155–159.

Deusch, O., O'Flynn, C., Colyer, A., Morris, P., Allaway, D., Jones, P. G., & Swanson, K. S. (2014). Deep illumina-based shotgun sequencing reveals dietary effects on the structure and function of the fecal microbiome of growing kittens.

Devane, M. L., Nicol, C., Ball, A., Klena, J. D., Scholes, P., Hudson, J. A., … Savill, M. G. (2005). The occurrence of *Campylobacter* subtypes in environmental reservoirs and potential transmission routes. Journal of Applied Microbiology, 98(4), 980–990.

Di Giovanni, G. D., Hashemi, F. H., Shaw, N. J., Abrams, F. A., LeChevallier, M. W., & Abbaszadegan, M. (1999). Detection of infectious *Cryptosporidium parvum* oocysts in surface and filter backwash water samples by immunomagnetic separation and integrated cell culture-PCR. Applied and Environmental Microbiology, 65(8), 3427–3432.

Di Giovanni, G. D., Hoffman, R. M., & Sturbaum, G. D. (2010). *Cryptosporidium* genotyping method for regulatory microscope slides. Water Research Foundation, Denver.

Diaz, P., Quilez, J., Chalmers, R. M., Panadero, R., Lopez, C., Sanchez-Acedo, C., ... & Diez-Banos, P. (2010). Genotype and subtype analysis of *Cryptosporidium* isolates from calves and lambs in Galicia (NW Spain). Parasitology, 137(08), 1187-1193.

Diaz de Ramírez, A. D. (2012). Dynamic of infections by *Cryptosporidium* spp. and *Giardia* spp. in Buffaloes (Bubalus bubalis) During the First Three Months of Life. REVISTA CIENTIFICA, 22(6).

Díaz-Sánchez, S., Sánchez, S., Herrera-León, S., Porrero, C., Blanco, J., Dahbi, G., ... & Vidal, D. (2013). Prevalence of Shiga toxin-producing *Escherichia coli*, Salmonella spp. and *Campylobacter* spp. in large game animals intended for consumption: relationship with management practices and livestock influence. Veterinary microbiology, 163(3), 274-281.

DiGiorgio, C. L., Gonzalez, D. A., & Huitt, C. C. (2002). *Cryptosporidium* and *Giardia* recoveries in natural waters by using environmental protection agency method 1623. Applied and Environmental Microbiology, 68(12), 5952–5955.

Dingle, K. E., Colles, F. M., Wareing, D. R. A., Ure, R., Fox, A. J., Bolton, F. E., Bootsma, H. J., Willems, R. J. L., Urwin, R. & Maiden, M. C. J. (2001). Multilocus sequence typing system for *Campylobacter jejuni*. Journal of Clinical Microbiology, 39(1), 14–23.

Dingle, K. E., Colles, F. M., Ure, R., Wagenaar, J. A., Duim, B., Bolton, F. J., ... & Maiden, M. C. (2002). Molecular characterization of *Campylobacter jejuni* clones: a basis for epidemiologic investigation. Emerging infectious diseases, 8(9), 949-55.

Dingle, K. E., Colles, F. M., Falush, D., Maiden, M. C. J. (2005). Sequence typing and comparison of population biology of *Campylobacter coli* and *Campylobacter jejuni*. J Clin Microbiol 43:340-347.

Dixon, B., Parrington, L., Cook, A., Pintar, K., Pollari, F., Kelton, D., & Farber, J. (2011). The potential for zoonotic transmission of *Giardia duodenalis* and *Cryptosporidium* spp. from beef and dairy cattle in Ontario, Canada. Veterinary parasitology, 175(1), 20-26.

Donnison, A., & Ross, C. (2009). Survival and retention of *Escherichia coli* O157:H7 and *Campylobacter* in contrasting soils from the Toenepi catchment. New Zealand Journal of Agricultural Research, 52(2), 133–144.

Donnison, A. (2003). Isolation of thermotolerant *Campylobacter*–Review & methods for New Zealand laboratories. Ministry of Health of New Zeeland.

Dorner, S. M., Huck, P. M. & Slawson, R. M. (2004). Estimating potential environmental loadings of *Cryptosporidium* spp. and *Campylobacter* spp. from livestock in the Grand River Watershed, Ontario, Canada. Environmental Science & Technology, 38(12), 3370-3380.

Dubinsky, E. A., Esmaili, L., Hulls, J. R., Cao, Y., Griffith, J. F., & Andersen, G. L. (2012). Application of phylogenetic microarray analysis to discriminate sources of fecal pollution. Environmental science & technology, 46(8), 4340-4347.

Duim, B., Ang, C. W., van Belkum, A., Rigter, A., van Leeuwen, N. W. J., Endtz, H. P., & Wagenaar, J. A. (2000). Amplified fragment length polymorphism analysis of *Campylobacter jejuni* strains isolated from chickens and from patients with gastroenteritis or Guillain-Barre or Miller Fisher syndrome. Applied and Environmental Microbiology, 66(9), 3917–3923.

Duncanson, M., Russell, N., Weinstein, P., Baker, M., Skelly, C., Hearnden, M., & Woodward, A. (2000). Rates of notified cryptosporidiosis and quality of drinking water supplies in Aotearoa, New Zealand. Water Research, 34(15), 3804–3812.

DuPont, H. L., Chappell, C. L., Sterling, C. R., Okhuysen, P. C., Rose, J. B., & Jakubowski, W. (1995). The infectivity of *Cryptosporidium parvum* in healthy volunteers. New England Journal of Medicine, 332(13), 855-859.

Dwivedi, K. K., Prasad, G., Saini, S., Mahajan, S., Lal, S., & Baveja, U. K. (2007). Enteric opportunistic parasites among HIV infected individuals: Associated risk factors and immune status. Japanese Journal of Infectious Diseases, 60(2-3), 76–81.

Eamens, G. J., Hornitzky, M. A., Walker, K. H., Hum, S. I., Vanselow, B. A., Bailey, G. D., ... & Gill, P. A. (2003). A study of the foodborne pathogens: *Campylobacter*, *Listeria* and *Yersinia* in faeces from slaughter-age cattle and sheep in Australia.

Eberhart-Phillips, J., Walker, N., Garrett, N., Bell, D., Sinclair, D., Rainger, W., & Bates, M. (1997). Campylobacteriosis in New Zealand: results of a case-control study. Journal of Epidemiology and Community Health, 51(6), 686-691.

Eckmann, L. (2002). Small bowel infections. Current Opinion in Gastroenterology, 18(2), 197–202.

Edwards, R. A., Rodriguez-Brito, B., Wegley, L., Haynes, M., Breitbart, M., Peterson, D. M., ... & Rohwer, F. (2006). Using pyrosequencing to shed light on deep mine microbial ecology. BMC genomics, 7(1), 1.

EFSA, 2014. The European union summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in 2012. EFSA J. 12, 3547, 312 pp.

Ellis-Iversen, J., Pritchard, G. C., Wooldridge, M., & Nielen, M. (2009b). Risk factors for *Campylobacter jejuni* and *Campylobacter coli* in young cattle on English and Welsh farms. Preventive Veterinary Medicine, 88(1), 42–48.

Elwin, K., Robinson, G., Hadfield, S. J., Fairclough, H. V, Iturriza-Gomara, M., & Chalmers, R. M. (2012). A comparison of two approaches to extracting *Cryptosporidium* DNA from human stools as measured by a real-time PCR assay. Journal of Microbiological Methods, 89(1), 38–40.

Enright, M. C., & Spratt, B. G. (1999). Multilocus sequence typing. Trends in microbiology, 7(12), 482-487.

Espelage, W., an der Heiden, M., Stark, K., & Alpers, K. (2010). Characteristics and risk factors for symptomatic *Giardia lamblia* infections in Germany. BMC public health, 10(1), 1.

ESR. (2007). The Institute of Environmental Science and Research Ltd. Notifiable and Other Diseases in New Zealand: Annual Report 2006. Porirua, New Zealand.

ESR. (2011). The Institute of Environmental Science and Research Ltd. Notifiable and Other Diseases in New Zealand: Annual Report 2010.

ESR. (2011a). The Institute of Environmental Science and Research Ltd. Annual Summary of Outbreaks in New Zealand 2010 (Vol. 8). Porirua, New Zealand

ESR. (2011b). The Institute of Environmental Science and Research Ltd. Notifiable and Other Diseases in New Zealand: Annual Report 2010 (Vol. 4). Porirua, New Zealand

ESR. (2012a). The Institute of Environmental Science and Research Ltd. Notifiable and Other Diseases in New Zealand: Annual Report 2012. Porirua, New Zealand.

ESR. (2012b). The Institute of Environmental Science and Research Ltd. Annual Summary of Outbreaks in New Zealand: Annual Report 2012. Porirua, New Zealand.

ESR. (2013). The Institute of Environmental Science and Research Ltd. Notifiable and Other Diseases in New Zealand: Annual Report 2013. Porirua, New Zealand.

ESR. (2014). The Institute of Environmental Science and Research Ltd. Notifiable and Other Diseases in New Zealand: Annual Report 2013. Porirua, New Zealand.

ESR. (2014). The Institute of Environmental Science and Research Ltd. Notifiable and Other Diseases in New Zealand: Annual Report 2013. Porirua, New Zealand.

ESRI. (2002). ArcView 3.2a: Environmental Systems Research Institute.

Esteban, J. G., Aguirre, C., Flores, A., Strauss, W., Angles, R., & Mas-Coma, S. (1998). High *Cryptosporidium* prevalences in healthy Aymara children from the northern Bolivian Altiplano. American Journal of Tropical Medicine and Hygiene, 58(1), 50–55.

European Centre for Disease Control and Prevention (ECDC). 2012. Food-borne diseases. Available from: http://www.ecdc.europa.eu/en/healthtopics/climate_change /health_ effects/Pages/food_borne_diseases.aspx. Accessed 2013 September 15.

Evans, M., Roberts, R., Ribeiro, C., Gardner, D., Kembrey, D. (1996). A milk-borne *Campylobacter* outbreak following an educational farm visit. Epidemiol Infect, 117, 457-462.

Ewing, B., & Green, P. (1998). Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome research, 8(3), 186-194.

Eyles, R., Niyogi, D., Townsend, C., Benwell, G., Weinstein, P. (2003). Spatial and temporal patterns of contamination underlying public health risk in the Taieri River, New Zealand. J Environ Qual 32:1820-1828.

FAO. 2006c. World Agriculture towards 2030/2050. Interim report. Global Perspective Studies Unit. Rome, June.

FAO (2014). 'Food and Agriculture Organization of the United Nations, FAOSTAT database.' Available at http://faostat3.fao.org/home/E.

Farthing, M. J. G. (2000). Diarrhoea: a significant worldwide problem. International Journal of Antimicrobial Agents, 14(1), 65-69.

Faustini, A., Marinacci, C., Fabrizi, E., Marangi, M., Recchia, O., Pica, R., ... & Perucci, C. A. (2006). The impact of the Catholic Jubilee in 2000 on infectious diseases. A case-control study of giardiasis, Rome, Italy 2000–2001. Epidemiology and infection, 134(03), 649-658.

Fayer, R. (Ed.). (1997). *Cryptosporidium* and cryptosporidiosis. Boca Raton: CRC Press

Fayer, R., Gasbarre, L., Pasquali, P., Canals, A., Almeria, S., & Zarlenga, D. (1998a). *Cryptosporidium parvum* infection in bovine neonates: dynamic clinical, parasitic and immunologic patterns. International Journal for Parasitology, 28(1), 49–56.

Fayer, R., Graczyk, T. K., Lewis, E. J., Trout, J. M.& Farley, C. A. (1998b). Survival of infectious *Cryptosporidium parvum* oocysts in seawater and Eastern oysters (Crassostrea virginica) in the Chesapeake Bay. Applied and Environmental Microbiology, 64(3), 1070-1074.

Fayer, R. (2000). Cryptosporidiosis in cattle: A review (Reprinted). Large Animal Practice, 21(2), 22–23.

Fayer, R. (2004). *Cryptosporidium*: a water-borne zoonotic parasite. Veterinary Parasitology, 126(1-2), 37-56.

Fayer, R., & Trout, J. M. (2005). Zoonotic protists in the marine environment. In Oceans and Health: Pathogens in the Marine Environment (pp. 143-163). Springer US.

Fayer, R., Santin, M., & Xiao, L. H. (2005). *Cryptosporidium bovis* n. sp (Apicomplexa : Cryptosporidiidae) in cattle (Bos taurus). Journal of Parasitology, 91(3), 624–629.

Fayer, R., Santín, M., Trout, J. M., Greiner, E. (2006). Prevalence of species and genotypes of *Cryptosporidium* found in 1-2-year-old dairy cattle in the eastern United States. Vet Parasitol 135, 105-112.

Fayer, R., Santin, M., Trout, J. (2007). A two-year longitudinal study of *Cryptosporidium* species and genotypes in dairy cattle. J Phycol 43, 55-55.

Fayer, R. & Xiao, L. (Eds.). (2008a). *Cryptosporidium* and cryptosporidiosis (2nd ed.). Boca Raton: Taylor & Francis.

Fayer, R., Santin, M., & Trout, J. M. (2008b). *Cryptosporidium ryanae* n. sp (Apicomplexa: Cryptosporidiidae) in cattle (Bos taurus). Veterinary Parasitology, 156(3-4), 191–198.

Fayer, R., & Santin, M. (2009). *Cryptosporidium xiaoi* n. sp (Apicomplexa: Cryptosporidiidae) in sheep (Ovis aries). Veterinary Parasitology, 164(2-4), 192–200.

Fayer, R. (2010). Taxonomy and species delimitation in *Cryptosporidium*. Exp Parasitol 124, 90-97.

Federighi, M., Tholozan, J. L., Cappelier, J. M., Tissier, J. P., Jouve, J. L. (1998). Evidence of non-coccoid viable but non-culturable *Campylobacter jejuni* cells in microcosm water by direct viable count, CTC-DAPI double staining, and scanning electron microscopy. Food Microbiol, 15, 539-550.

Fedurco, M., Romieu, A., Williams, S., Lawrence, I., & Turcatti, G. (2006). BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies. Nucleic acids research, 34(3), e22-e22.

Feng, Y. Y., Ong, S. L., Hu, J. Y., Song, L. F., Tan, X. L., & Ng, W. J. (2003). Effect of particles on the recovery of *Cryptosporidium* oocysts from source water samples of various turbidities. Applied and Environmental Microbiology, 69(4), 1898–1903.

Feng, Y., Ortega, Y., He, G., Das, P., Xu, M., Zhang, X., ... & Xiao, L. (2007). Wide geographic distribution of *Cryptosporidium bovis* and the deer-like genotype in bovines. Veterinary parasitology, 144(1), 1-9.

Feng, Y., & Xiao, L. (2011). Zoonotic potential and molecular epidemiology of *Giardia* species and giardiasis. Clinical Microbiology Reviews, 24(1), 110–40.

Feng, Y., Zhao, X., Chen, J., Jin, W., Zhou, X., Li, N., Wang, L.& Xiao, L. (2011). Occurrence, Source, and Human Infection Potential of *Cryptosporidium* and *Giardia* spp. in Source and Tap Water in Shanghai, China. Applied and Environmental Microbiology, 77(11), 3609-3616.

Fitzenberger, J., Uphoff, H., Gawrich, S., & Hauri, A. M. (2010). Urban-rural differences of age- and species-specific campylobacteriosis incidence, Hesse, Germany, July 2005-June 2006. Eurosurveillance, 15(42), 12–18.

Fitzgerald, C., Helsel, L. O., Nicholson, M. A., Olsen, S. J., Swerdlow, D. L., Flahart, R., ... & Fields, P. I. (2001). Evaluation of methods for subtyping *Campylobacter jejuni* during an outbreak involving a food handler. Journal of Clinical Microbiology, 39(7), 2386-2390.

Fleiss, J. L. (1981). Balanced incomplete block designs for inter-rater reliability studies. Applied Psychological Measurement, 5(1), 105-112.

Fletcher, S. M., Stark, D., Harkness, J., & Ellis, J. (2012). Enteric protozoa in the developed world: a public health perspective. Clinical microbiology reviews, 25(3), 420-449.

Fong, T. T., & Lipp, E. K. (2005). Enteric viruses of humans and animals in aquatic environments: health risks, detection, and potential water quality assessment tools. Microbiology and molecular biology reviews, 69(2), 357-371.

Fong, T. T., Mansfield, L. S., Wilson, D. L., Schwab, D. J., Molloy, S. L., & Rose, J. B. (2007). Massive microbiological groundwater contamination associated with a waterborne outbreak in Lake Erie, South Bass Island, Ohio. Environmental health perspectives, 856-864.

Fox, J., Kahn, H., Regli, S., Grubs, B., Lustik, M., & Owens, M. (1996). Evaluation of a method for monitoring protozoa in raw water. American Statistical Association - 1996 Proceedings of the Section on Statistics and the Environment, 37–42.

Franzen, O., Jerlström-Hultqvist, J., Castro, E., Sherwood, E., Ankarklev, J., Reiner, D. S., ... & Svärd, S. G. (2009). Draft genome sequencing of giardia *intestinalis* assemblage B isolate GS: is human giardiasis caused by two different species?. PLoS Pathog, 5(8).

Fraser, D., Bilenko, N., Deckelbaum, R. J., Dagan, R., El-On, J., & Naggan, L. (2000). *Giardia lamblia* carriage in Israeli Bedouin infants: Risk factors and consequences. Clinical Infectious Diseases, 30(3), 419–424.

French, N. P., Till, D., Learmonth, J.& Pita, A. (2007). Manawatu raw water pilot survey. Palmerston North: Massey University.

French, N. P., Midwinter, A., Holland, B., Collins-Emerson, J., Pattison, R., Colles, F., Carter, P. (2009). Molecular epidemiology of *Campylobacter jejuni* isolates from wild-bird fecal material in children's playgrounds. Appl Environ Microb 75:779-783.

French, N. P., Marshall, J., Mohan, V. (2011). New and emerging data on typing of *Campylobacter* spp. strains in animals, environmental matrices and humans. Technical report. Ministry of Agriculture and Forestry.

French, N. P., Till, D., Learmonth, J.& Pita, A. (2007). Manawatu raw water pilot survey. Palmerston North: Massey University.French, N., & Marshall, J. (2014). Source Attribution Studies for Campylobacteriosis in New Zealand.

Friedrich, A., Marshall, J. C., Biggs, P. J., Midwinter, A. C., & French, N. P. (2015). Seasonality of *Campylobacter jejuni* isolates associated with human campylobacteriosis in the Manawatu region, New Zealand. Epidemiology and infection, 1-9.

Frost, J., Gillespie, I., O'brien, S. (2002). Public health implications of *Campylobacter* outbreaks in England and Wales, 1995–9: epidemiological and microbiological investigations. Epidemiol Infect, 128, 111-118.

Fugett, E. B., Schoonmaker-Bopp, D., Dumas, N. B., Corby, J., & Wiedmann, M. (2007). Pulsed-field gel electrophoresis (PFGE) analysis of temporally matched Listeria monocytogenes isolates from human clinical cases, foods, ruminant farms, and urban and natural environments reveals source-associated as well as widely distributed PFGE types. Journal of clinical microbiology, 45(3), 865-873.

Garcia, L. S., Bruckner, D. A., Brewer, T. C., & Shimizu, R. Y. (1983). Techniques for the Recovery and Identification of *Cryptosporidium* Oocysts from Stool Specimens. Journal of Clinical Microbiology, 18(1), 185–190.

Garcia, M. M., Lior, H., Stewart, R. B., Ruckerbauer, G. M., Trudel, J. R. R., & Skljarevski, A. (1985). Isolation, Characterization, and Serotyping of *Campylobacter-jejuni* and *Campylobacter*-Coli from Slaughter Cattle. Applied and Environmental Microbiology, 49(3), 667–672.

Garcia, L. S., & Shimizu, R. Y. (1997). Evaluation of nine immunoassay kits (enzyme immunoassay and direct fluorescence) for detection of *Giardia lamblia* and *Cryptosporidium parvum* in human fecal specimens. Journal of Clinical Microbiology, 35(6), 1526–1529.

García-Presedo, I., Pedraza-Díaz, S., González-Warleta, M., Mezo, M., Gómez-Bautista, M., Ortega-Mora, L. M., & Castro-Hermida, J. A. (2013). The first report of *Cryptosporidium bovis*, *C. ryanae* and *Giardia duodenalis* sub-assemblage A-II in roe deer (Capreolus capreolus) in Spain. Veterinary parasitology, 197(3), 658-664.

Gardner, T. B., & Hill, D. R. (2001). Treatment of giardiasis. Clinical Microbiology Reviews, 14(1), 114.

Gasser, R. B., El-Osta, Y. G. A., Prepens, S., & Chalmers, R. M. (2004). An improved "cold SSCP" method for the genotypic and subgenotypic characterization of *Cryptosporidium*. Molecular and Cellular Probes, 18(5), 329–332.

Gatei, W., Greensill, J., Ashford, R. W., Cuevas, L. E., Parry, C. M., Cunliffe, N. A., Beeching, N. J., Hart, C. A. (2003). Molecular analysis of the 18S rRNA gene of *Cryptosporidium* parasites from patients with or without human immunodeficiency virus infections living in Kenya, Malawi, Brazil, the United Kingdom, and Vietnam. J Clin Microbiol 41, 1458-1462.

Gatei, W., Das, P., Dutta, P., Sen, A., Cama, V., Lal, A. A., & Xiao, L. H. (2007). Multilocus sequence typing and genetic structure of *Cryptosporidium hominis* from children in Kolkata, India. Infection Genetics and Evolution, 7(2), 197–205.

Geurden, T., Vercruysse, J., & Claerebout, E. (2006a). Field testing of a fenbendazole treatment combined with hygienic and management measures against a natural *Giardia* infection in calves. Veterinary Parasitology, 142(3-4), 367–371.

Geurden, T., Claerebout, E., Dursin, L., Deflandre, A., Bernay, F., Kaltsatos, V., & Vercruysse, J. (2006b). The efficacy of an oral treatment with paromomycin against an experimental infection with *Giardia* in calves. Veterinary parasitology, 135(3), 241-247.

Geurden, T., Thomas, P., Casaert, S., Vercruysse, J., & Claerebout, E. (2008). Prevalence and molecular characterisation of *Cryptosporidium* and *Giardia* in lambs and goat kids in Belgium. Veterinary parasitology, 155(1), 142-145.

Geurden, T., & Olson, M. (2011). *Giardia* in pets and farm animals, and their zoonotic potential (pp. 71-92). Springer Vienna.

Geurden, T., Vanderstichel, R., Pohle, H., Ehsan, A., von Samson-Himmelstjerna, G., Morgan, E. R., ... & Claerebout, E. (2012). A multicentre prevalence study in Europe on *Giardia duodenalis* in calves, with molecular identification and risk factor analysis. Veterinary parasitology, 190(3), 383-390.

Ghai, R., Rodriguez-Valera, F., McMahon, K. D., Toyama, D., Rinke, R., de Oliveira, T. C. S., … Henrique-Silva, F. (2011). Metagenomics of the water column in the pristine upper course of the Amazon river. PloS One, 6(8), e23785.

Giacoboni, G. I., Itoh, K., Hirayama, K., Takahashi, E., & Mitsuoka, T. (1993). Comparison of fecal *Campylobacter* in calves and cattle of different ages and areas in Japan. The

Journal of Veterinary Medical Science/the Japanese Society of Veterinary Science, 55(4), 555–559.

Gilpin, B., Scholes, P., Robson, B. & Savill, M. (2008). The transmission of thermotolerant *Campylobacter* spp. to people living or working on dairy farms in New Zealand. Zoonoses and Public Health, 55(7), 352-360.

Gilpin, B., Robson, B., Lin, S., Scholes, P., & On, S. (2012). Pulsed-Field Gel Electrophoresis Analysis of More than One Clinical Isolate of *Campylobacter* spp. from Each of 49 Patients in New Zealand. Journal of Clinical Microbiology, 50(2), 457-459.

Gilpin, B. J., Walsh, G., On, S. L., Smith, D., Marshall, J. C., & French, N. P. (2013). Application of molecular epidemiology to understanding campylobacteriosis in the Canterbury region of New Zealand. Epidemiology and Infection, 141(6), 1253–1266.

Glaberman, S., Moore, J. E., Lowery, C. J., Chalmers, R. M., Sulaiman, I., Elwin, K., … Xiao, L. H. (2002). Three drinking-water-associated cryptosporidiosis outbreaks, Northern Ireland. Emerging Infectious Diseases, 8(6), 631–633.

Gleick, P. H. (2002). Dirty-water: Estimated Deaths from Water-related Diseases 2000-2020. Pacific Institute for Studies in Development, Environment, and Security.

Gómez-Muñoz, M. T., Navarro, C., Garijo-Toledo, M. M., Dea-Ayuela, M. A., Fernández-Barredo, S., Pérez-Gracia, M. T., ... & Borrás, R. (2009). Occurrence and genotypes of *Giardia* isolated from lambs in Spain. Parasitology international, 58(3), 297-299.

Gomez-Munoz, M. T., Camara-Badenes, C., Martinez-Herrero, M. D., Dea-Ayuela, M. A., Perez-Gracia, M. T., Fernandez-Barredo, S., … Fayer, R. (2012). Multilocus genotyping of *Giardia* in lambs from Spain reveals a high heterogeneity. Research in Veterinary Science, 93(2), 836–842.

Gonzalez, M., & Hanninen, M. L. (2012). Effect of temperature and antimicrobial resistance on survival of *Campylobacter jejuni* in well water: application of the Weibull model. Journal of Applied Microbiology, 113(2), 284–293.

Gow, S., & Waldner, C. (2006). An examination of the prevalence of and risk factors for shedding of *Cryptosporidium* spp. and *Giardia* spp. in cows and calves from western Canadian cow-calf herds. Veterinary Parasitology, 137(1-2), 50–61.

Grace, D., Gilbert, J., Randolph, T., & Kang'ethe, E. (2012). The multiple burdens of zoonotic disease and an ecohealth approach to their assessment. Tropical Animal Health and Production, 44(1), 67–73.

Green, C. G., Krause, D. O., & Wylie, J. L. (2006). Spatial analysis of *Campylobacter* infection in the Canadian province of Manitoba. International Journal of Health Geographics, 5, 2.

Grinberg, A., Pomroy, W. E., Weston, J. F., Ayanegui-Alcerreca, A.& Knight, D. (2005). The occurrence of *Cryptosporidium parvum*, *Campylobacter* and Salmonella in newborn

dairy calves in the Manawatu region of New Zealand. New Zealand Veterinary Journal, 53(5), 315-320

Grinberg, A., Learmonth, J., Kwan, E., Pomroy, W., Lopez Villalobos, N., Gibson, I., Widmer, G. (2008). Genetic diversity and zoonotic potential of *Cryptosporidium parvum* causing foal diarrhoea. J. Clin. Microbiol., JCM.00936-00908.

Grinberg, A., Pomroy, W.E., Squires, R.A., Scuffham, A., Pita, A., Kwan, E. (2011). Retrospective cohort study of an outbreak of cryptosporidiosis caused by a rare *Cryptosporidium parvum* subgenotype. Epidemiology & Infection 139, 1542-1550.

Grinberg, A., Biggs, P., Dukkipati, V., George, T. (2013). Extensive intra-host genetic diversity uncovered in *Cryptosporidium parvum* using Next Generation Sequencing. Infection, Genetics and Evolution.

Gronewold, A. D., Borsuk, M. E., Wolpert, R. L., & Reckhow, K. H. (2008). An assessment of fecal indicator bacteria-based water quality standards. Environmental Science & Technology, 42(13), 4676–4682.

Grove-White, D. H., Leatherbarrow, a J. H., Cripps, P. J., Diggle, P. J., & French, N. P. (2010a). Temporal and farm-management-associated variation in the faecal-pat prevalence of *Campylobacter jejuni* in ruminants. Epidemiology and Infection, 138(4), 549–58.

Hadfield, S. J., Robinson, G., Elwin, K., & Chalmers, R. M. (2011). Detection and Differentiation of *Cryptosporidium* spp. in Human Clinical Samples by Use of Real-Time PCR. Journal of Clinical Microbiology, 49(3), 918–924

Hadfield, S. J., Pachebat, J. A., Swain, M. T., Robinson, G., Cameron, S. J. S., Alexander, J., … Chalmers, R. M. (2015). Generation of whole genome sequences of new *Cryptosporidium hominis* and *Cryptosporidium parvum* isolates directly from stool samples. BMC Genomics, 16(1), 650.

Hamnes, I. S., Gjerde, B., & Robertson, L. (2006). Prevalence of *Giardia* and *Cryptosporidium* in dairy calves in three areas of Norway. Veterinary Parasitology, 140(3-4), 204–216.

Handelsman, J., Rondon, M. R., Brady, S. F., Clardy, J. & Goodman, R. M. (1998). Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. Chemistry & Biology, 5, R245-R249.

Hänninen, M. L., Niskanen, M., & Korhonen, L. (1998). Water as a Reservoir for *Campylobacter jejuni* Infection in Cows Studied by Serotyping and Pulsed-field Gel Electrophoresis (PFGE). Journal of Veterinary Medicine, Series B, 45(1-10), 37-42.

Hänninen, M. L., Sarelli, L., Sukura, A., On, S. L., Harrington, C. S., Matero, P., & Hirvelä-Koski, V. (2002). *Campylobacter hyointestinalis* subsp. *hyointestinalis*, a common *Campylobacter* species in reindeer. Journal of applied microbiology, 92(4), 717-723.

Hannon, S. J., Allan, B., Waldner, C., Russell, M. L., Potter, A., Babiuk, L. A., & Townsend, H. G. G. (2009). Prevalence and risk factor investigation of *Campylobacter* species in beef

cattle feces from seven large commercial feedlots in Alberta, Canada. Canadian Journal of Veterinary Research-Revue Canadienne De Recherche Veterinaire, 73(4), 275–282.

Haque, R., Roy, S., Siddique, A., Mondal, U., Rahman, S. M. M., Mondal, D., … Petri, W. A. (2007). Multiplex real-time PCR assay for detection of Entamoeba histolytica, *Giardia* intestinalis, and *Cryptosporidium* spp. American Journal of Tropical Medicine and Hygiene, 76(4), 713–717.

Harismendy, O., Ng, P. C., Strausberg, R. L., Wang, X., Stockwell, T. B., Beeson, K. Y., ... & Frazer, K. A. (2009). Evaluation of next generation sequencing platforms for population targeted sequencing studies. Genome biol, 10(3), R32.

Hazen, T. C., Rocha, A. M., & Techtmann, S. M. (2013). Advances in monitoring environmental microbes. Current opinion in biotechnology, 24(3), 526-533.

Hearnden, M., Skelly, C., Eyles, R., & Weinstein, P. (2003). The regionality of campylobacteriosis seasonality in New Zealand. International Journal of Environmental Health Research, 13(4), 337-348.

Heitman, T. L., Frederick, L. M., Viste, J. R., Guselle, N. J., Morgan, U. M., Thompson, R. C. A., & Olson, M. E. (2002). Prevalence of *Giardia* and *Cryptosporidium* and characterization of *Cryptosporidium* spp. isolated from wildlife, human, and agricultural sources in the North Saskatchewan River Basin in Alberta, Canada. Canadian Journal of Microbiology, 48(6), 530–541.

Helmi, K., Skraber, S., Burnet, J. B., Leblanc, L., Hoffmann, L. & Cauchie, H. M. (2011). Two-year monitoring of *Cryptosporidium parvum* and *Giardia lamblia* occurrence in a recreational and drinking water reservoir using standard microscopic and molecular biology techniques. Environmental Monitoring and Assessment, 179(1), 163-175.

Helmy, Y. A., Krucken, J., Nockler, K., von Samson-Himmelstjerna, G., & Zessin, K. H. (2013). Molecular epidemiology of *Cryptosporidium* in livestock animals and humans in the Ismailia province of Egypt. Veterinary Parasitology, 193(1-3), 15–24.

Helmy, Y. A., Krucken, J., Nockler, K., von Samson-Himmelstjerna, G., & Zessin, K. H. (2014b). Comparison between two commercially available serological tests and polymerase chain reaction in the diagnosis of *Cryptosporidium* in animals and diarrhoeic children. Parasitology Research, 113(1), 211–216.

Helmy, Y. A., Klotz, C., Wilking, H., Krücken, J., Nöckler, K., Von Samson-Himmelstjerna, G., ... & Aebischer, T. (2014a). Epidemiology of *Giardia duodenalis* infection in ruminant livestock and children in the Ismailia province of Egypt: insights by genetic characterization. Parasit Vectors, 7(1), 321.

Hibler, C. P. (1987). Inactivation of *Giardia* cysts with chlorine at 0.50 C to 5.00 C. American Water Works Association.

Hintze, J. L., & Nelson, R. D. (1998). Violin plots: a box plot-density trace synergism. The American Statistician, 52(2), 181-184.

Hokajarvi, A. M., Pitkanen, T., Siljanen, H. M. P., Nakari, U. M., Torvinen, E., Siitonen, A., & Miettinen, I. T. (2013). Occurrence of thermotolerant *Campylobacter* spp. and adenoviruses in Finnish bathing waters and purified sewage effluents. Journal of Water and Health, 11(1), 120–134.

Holinger, E. P., Ross, K. A., Robertson, C. E., Stevens, M. J., Harris, J. K. & Pace, N. R. (2014). Molecular analysis of point-of-use municipal drinking water microbiology. Water Research, 49, 225-235.

Homan, W. L., & Mank, T. G. (2001). Human giardiasis: genotype linked differences in clinical symptomatology. International journal for parasitology, 31(8), 822-826.

Homan, W. L., Gilsing, M., Bentala, H., Limper, L., & van Knapen, F. (1998). Characterization of *Giardia duodenalis* by polymerase-chain-reaction fingerprinting. Parasitology research, 84(9), 707-714.

Hoque, M. E., Hope, V. T., Kjellström, T., Scragg, R., & Lay-Yee, R. (2002). Risk of giardiasis in Aucklanders: a case—control study. International journal of infectious diseases, 6(3), 191-197.

Horman, A., Rimhanen-Finne, R., Maunula, L., von Bonsdorff, C. H., Torvela, N., Heikinheimo, A., & Hanninen, M. L. (2004). *Campylobacter* spp., *Giardia* spp., *Cryptosporidium* noroviruses and indicator organisms in surface water in southwestern Finland, 2000-2001. Applied and Environmental Microbiology, 70(1), 87–95.

Hosmer D., & Stanley L. W. (2000). Applied logistic regression. John Wiley & amp.

Howe, A. C., Jansson, J. K., Malfatti, S. A., Tringe, S. G., Tiedje, J. M., & Brown, C. T. (2014). Tackling soil diversity with the assembly of large, complex metagenomes. Proceedings of the National Academy of Sciences, 111(13), 4904–4909.

HPA. (2011). *Campylobacter* infections per year in England and Wales, 1989-2010. Health Protection Agency. Retrieved on 24 June 2011, from http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/*Campylobacter*/EpidemiologicalData/campyDataEw.

HPA. (2012a). *Cryptosporidium* infections per year in England and Wales, 1989-2010. Health Protection Agency. Retrieved on 5 May 2013, from http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/*Cryptosporidium*/EpidemiologicalData.

HPA. (2012b). *Giardia* infections per year in England and Wales, 1989-2010. Health Protection Agency. Retrieved on 5 May 2013, from http://www.hpa.org.uk/Topics/InfectiousDiseases/InfectionsAZ/*Giardia*/EpidemiologicalData.

Hrudey, S. E., Huck, P. M., Payment, P., Gillham, R. W., & Hrudey, E. J. (2002). Walkerton: Lessons learned in comparison with waterborne outbreaks in the developed world. Journal of Environmental Engineering and Science, 1(6), 397–407.

Hrudey, S. E., Payment, P., Huck, P. M., Gillham, R. W., & Hrudey, E. J. (2003). A fatal waterborne disease epidemic in Walkerton, Ontario: comparison with other waterborne outbreaks in the developed world. Water Science and Technology, 47(3), 7–14.

Hrudey, S. E., & Hrudey, E. J. (2007a). A nose for trouble - the role of off-flavours in assuring safe drinking water. Water Science and Technology, 55(5), 239-247.388. doi: Doi 10.2166/Wst.2007.185

Hrudey, S. E., & Hrudey, E. J. (2007b). Published case studies of waterborne disease outbreaks - Evidence of a recurrent threat. Water Environment Research, 79(3), 233-245. doi: Doi 10.2175/106143006x95483

Hsu, B. M., Huang, C. P., Hsu, Y. F., Jiang, G. Y., & Hsu, C. L. L. (2001a). Evaluation of two concentration methods for detecting *Giardia* and *Cryptosporidium* in water. Water Research, 35(2), 419–424.

Hsu, B. M., & Huang, C. P. (2001b). Performances of the immunomagnetic separation method for *Cryptosporidium* in water under various operation conditions. Biotechnology Progress, 17(6), 1114–1118.

Hudson, J. A., Nicol, C., Wright, J., Whyte, R., & Hasell, S. K. (1999). Seasonal variation of *Campylobacter* types from human cases, veterinary cases, raw chicken, milk and water. Journal of Applied Microbiology, 87(1), 115-124.

Hugenholtz, P., & Tyson, G. W. (2008). Microbiology: metagenomics. Nature, 455(7212), 481-483.

Hum, S. (1987). Bovine Abortion Due to *Campylobacter-fetus*. Australian Veterinary Journal, 64(10), 319–320.

Human Genome Sequencing, C. 2004. Finishing the euchromatic sequence of the human genome. Nature, 431, 931-945.

Humphrey, T. J., & Beckett, P. (1987). *Campylobacter jejuni* in dairy cows and raw milk. Epidemiology and infection, 98(03), 263-269.

Hunt, C. L., Ionas, G., & Brown, T. J. (2000). Prevalence and strain differentiation of *Giardia intestinalis* in calves in the Manawatu and Waikato regions of North Island, New Zealand. Veterinary Parasitology, 91(1), 7-13.

Hunter, P. R., & Nichols, G. (2002). Epidemiology and clinical features of *Cryptosporidium* infection in immunocompromised patients. Clinical Microbiology Reviews, 15(1), 145.

Hunter, P. R., Andersson, Y., Von Bonsdorff, C. H., Chalmers, R. M., Cifuentes, E., Deere, D., ... & Prescott, A. (2003). Surveillance and investigation of contamination incidents and waterborne outbreaks. Assessing microbial safety of drinking water, 205.

Huson, D. H., Auch, A. F., Qi, J., & Schuster, S. C. (2007). MEGAN analysis of metagenomic data. Genome Research, 17, 377-386.

Huson, D. H., & Xie, C. (2013). A poor man's BLASTX—high-throughput metagenomic protein database search using PAUDA. Bioinformatics.

Imre, K., Luca, C., Costache, M., Sala, C., Morar, A., Morariu, S., … Darabus, G. (2013). Zoonotic *Cryptosporidium parvum* in Romanian newborn lambs (Ovis aries). Veterinary Parasitology, 191(1-2), 119–122.

Inglis, G. D., Kalischuk, L. D., & Busz, H. W. (2004). Chronic shedding of *Campylobacter* species in beef cattle. Journal of applied microbiology, 97(2), 410-420.

Inglis, G. D., McAllister, T. A., Larney, F. J., Topp, E. (2010). Prolonged survival of *Campylobacter* species in bovine manure compost. Appl Environ Microb, 76,1110-1119.

Insulander, M., Lebbad, M., Stenstrom, T. A., & Svenungsson, B. (2005). An outbreak of cryptosporidiosis associated with exposure to swimming pool water. Scandinavian Journal of Infectious Diseases, 37(5), 354–360.

Irons, P. C., Schutte, A. P., Van Der Walt, M. L., & Bishop, G. C. (2004). Genital campylobacteriosis in cattle. Infectious diseases of livestock, 3, 1459-1468.

Ishii, S., & Sadowsky, M. J. (2008). *Escherichia coli* in the environment: Implications for water quality and human health. Microbes and Environments, 23, 101-108.

Izzo, M. M., Kirkland, P. D., Mohler, V. L., Perkins, N. R., Gunn, A. A., & House, J. K. (2011). Prevalence of major enteric pathogens in Australian dairy calves with diarrhoea. Australian veterinary journal, 89(5), 167-173.

Jafari, H., Jalali, M. H. R., Shapouri, M. S. A., & Hajikolaii, M. R. H. (2014). Determination of *Giardia duodenalis* genotypes in sheep and goat from Iran. Journal of parasitic diseases, 38(1), 81-84.

Jamieson, R., Gordon, R., Joy, D.& Lee, H. (2004). Assessing microbial pollution of rural surface waters: A review of current watershed scale modeling approaches. Agricultural water management, 70(1), 1-17.

Jenkins, M. B., Bowman, D. D., & Ghiorse, W. C. (1998). Inactivation of *Cryptosporidium* parvum oocysts by ammonia. Applied and environmental microbiology, 64(2), 784-788.

Jenkins, M. B., Bowman, D. D., Fogarty, E. A., & Ghiorse, W. C. (2002). *Cryptosporidium parvum* oocyst inactivation in three soil types at various temperatures and water potentials. Soil biology and Biochemistry, 34(8), 1101-1109.

Jeong, J.-Y., Park, H.-D., Lee, K.-H., Weon, H.-Y., & Ka, J.-O. (2011). Microbial community analysis and identification of alternative host-specific fecal indicators in fecal and river water samples using pyrosequencing. The Journal of Microbiology, 49(4), 585–594.

Jerlstrom-Hultqvist, J., Franzen, O., Ankarklev, J., Xu, F. F., Nohynkova, E., Andersson, J. O., … Andersson, B. (2010). Genome analysis and comparative genomics of a *Giardia intestinalis* assemblage E isolate. Bmc Genomics, 11.

Jex, A. R., Stanley, K. K., Lo, W., Littman, R., Verweij, J. J., Campbell, B. E., … Gasser, R. B. (2012). Detection of diarrhoeal pathogens in human faeces using an automated, robotic platform. Molecular and Cellular Probes, 26(1), 11–15.

Jiang, J., Alderisio, K. A., & Xiao, L. (2005). Distribution of *Cryptosporidium* genotypes in storm event water samples from three watersheds in New York. Applied and Environmental Microbiology, 71(8), 4446-4454.

Johnson, E. H., Muirhead, D. E., Windsor, J. J., King, G. J., Al-Busaidy, R., & Cornelius, R. (1999). Atypical outbreak of caprine cryptosporidiosis in the Sultanate of Oman. The Veterinary Record, 145(18), 521-524.

Johnston, S. P., Ballard, M. M., Beach, M. J., Causer, L., & Wilkins, P. P. (2003). Evaluation of three commercial assays for detection of *Giardia* and *Cryptosporidium* organisms in fecal specimens. Journal of Clinical Microbiology, 41(2), 623–626.

Jokipii, L., Pohjola, S., & Jokipii, A. M. M. (1983). *Cryptosporidium* - a Frequent Finding in Patients with Gastrointestinal Symptoms. Lancet, 2(8346), 358–360.

Jokipii, A. M. M., Hemila, M., & Jokipii, L. (1985). Prospective-Study of Acquisition of *Cryptosporidium*, *Giardia lamblia*, and Gastrointestinal Illness. Lancet, 2(8453), 487–489.

Jokinen, C., Edge, T. A., Ho, S., Koning, W., Laing, C., Mauro, W., … Gannon, V. P. J. (2011). 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 Research, 45(3), 1247–1257.

Jokinen, C. C., Edge, T. A., Koning, W., Laing, C. R., Lapen, D. R., Miller, J., ... & Topp, E. (2012). Spatial and temporal drivers of zoonotic pathogen contamination of an agricultural watershed. Journal of environmental quality, 41(1), 242-252.Jones K. (2001). Campylobacters in water, sewage and the environment. J Appl Microbiol 90:68s-79s.

Jones, K., Betaieb, M., & Telford, D. R. (1990). Thermophilic campylobacters in surface waters around Lancaster, UK: negative correlation with *Campylobacter* infections in the community. Journal of Applied Bacteriology, 69(5), 758-764.

Jones, K., & Hobbs, A. (1996). Campylobacters and faecal indicators in streams and rivers subject to farm run-off. In Campylobacters, Helicobacters, and Related Organisms (pp. 123-128). Springer US.

Jones, K., Howard, S., & Wallace, J. S. (1999). Intermittent shedding of thermophilic campylobacters by sheep at pasture. Journal of Applied Microbiology, 86(3), 531-536.

Joshua, G. W. P., Guthrie-Irons, C., Karlyshev, A. V, & Wren, B. W. (2006). Biofilm formation in *Campylobacter jejuni*. Microbiology-Sgm, 152, 387–396.

Júlio, C., Vilares, A., Oleastro, M., Ferreira, I., Gomes, S., Monteiro, L., ... & Ângelo, H. (2012). Prevalence and risk factors for *Giardia duodenalis* infection among children: a case study in Portugal. Parasit Vectors, 5(1), 22-29.

Karanis, P., Sotiriadou, I., Kartashev, V., Kourenti, C., Tsvetkova, N., & Stojanova, K. (2006). Occurrence of *Giardia* and *Cryptosporidium* in water supplies of Russia and Bulgaria. Environmental research, 102(3), 260-271.

Karanis, P., Kourenti, C., & Smith, H. (2007). Waterborne transmission of protozoan parasites: a worldwide review of outbreaks and lessons learnt. Journal of water and health, 5(1), 1-38.

Karanis, P., Eiji, T., Palomino, L., Boonrod, K., Plutzer, J., Ongerth, J., & Igarashi, I. (2010). First description of *Cryptosporidium bovis* in Japan and diagnosis and genotyping of *Cryptosporidium* spp. in diarrheic pre-weaned calves in Hokkaido. Veterinary parasitology, 169(3), 387-390.

Kassa, T., Gebre-Selassie, S., & Asrat, D. (2007). Antimicrobial susceptibility patterns of thermotolerant *Campylobacter* strains isolated from food animals in Ethiopia. Veterinary Microbiology, 119(1), 82–87.

Keeley, A., & Faulkner, B. R. (2008). Influence of land use and watershed characteristics on protozoa contamination in a potential drinking water resources reservoir. Water Research, 42(10), 2803-2813.

Keithlin, J., Sargeant, J., Thomas, M. K., & Fazil, A. (2014). Systematic review and meta-analysis of the proportion of *Campylobacter* cases that develop chronic sequelae. BMC public health, 14(1), 1.

Kemp, R., Leatherbarrow, A., Williams, N., Hart, C., Clough, H., Turner, J., Wright, E., French N. P. (2005). Prevalence and genetic diversity of *Campylobacter* spp. in environmental water samples from a 100-square-kilometer predominantly dairy farming area. Appl Environ Microb, 71, 1876-1882.

Khalifa, N. O., Afify, J. S., & Rabie, N. S. (2013). Zoonotic and Molecular Characterizations of *Campylobacter jejuni* and *Campylobacter coli* Isolated from Beef Cattle and Children. Global Veterinaria, 11(5), 585-591.

Khan, S. M., Debnath, C., Pramanik, A. K., Xiao, L., Nozaki, T., & Ganguly, S. (2010). Molecular characterization and assessment of zoonotic transmission of *Cryptosporidium* from dairy cattle in West Bengal, India. Veterinary parasitology, 171(1), 41-47.

Khan, S. M., Debnath, C., Pramanik, A. K., Xiao, L. H., Nozaki, T., & Ganguly, S. (2011). Molecular evidence for zoonotic transmission of *Giardia duodenalis* among dairy farm workers in West Bengal, India. Veterinary Parasitology, 178(3-4), 342–345.

King, B. J., & Monis, P. T. (2007). Critical processes affecting *Cryptosporidium* oocyst survival in the environment. Parasitology, 134(03), 309-323.

Kircher, M., & Kelso, J. (2010). High-throughput DNA sequencing–concepts and limitations. Bioessays, 32(6), 524-536.

Kirk, R., & Rowe, M. T. (1994). A Pcr Assay for the Detection of *Campylobacter-jejuni* and *Campylobacter*-coli in Water. Letters in Applied Microbiology, 19(5), 301–303.

Kirkpatrick, B. D., & Tribble, D. R. (2011). Update on human *Campylobacter jejuni* infections. Current opinion in gastroenterology, 27(1), 1-7.

Klein, D., Alispahic, M., Sofka, D., Iwersen, M., Drillich, M., & Hilbert, F. (2013). Prevalence and risk factors for shedding of thermophilic *Campylobacter* in calves with and without diarrhea in Austrian dairy herds. Journal of Dairy Science, 96(2), 1203–1210.

Klena, J. D., & Konkel, M. E. (2005). Methods for epidemiological analysis of *Campylobacter jejuni*. *Campylobacter*: Molecular and Cellular Biology, 165-179.

Koehler, A. V, Jex, A. R., Haydon, S. R., Stevens, M. A., & Gasser, R. B. (2014). *Giardia*/giardiasis—A perspective on diagnostic and analytical tools. Biotechnology Advances, 32(2), 280–289.

Koenraad, P. M. F. J., Jacobs-Reitsma, W. F., Van der Laan, T., Beumer, R. R., & Rombouts, F. M. (1995). Antibiotic susceptibility of *Campylobacter* isolates from sewage and poultry abattoir drain water. Epidemiology and infection, 115(03), 475-483.

Koenraad, P., Rombouts, F., Notermans, S. (1997). Epidemiological aspects of thermophilic *Campylobacter* in water-related environments: a review. Water Environ Res 69:52-63.

Koinari, M., Karl, S., Ng-Hublin, J., Lymbery, A., Ryan, U. (2013). Identification of novel and zoonotic *Cryptosporidium* species in fish from Papua New Guinea. Vet Parasitol 198, 1-9.

Kuczynski, J., Lauber, C. L., Walters, W. A., Parfrey, L. W., Clemente, J. C., Gevers, D., & Knight, R. (2012). Experimental and analytical tools for studying the human microbiome. Nature Reviews Genetics, 13(1), 47-58.

Kurniawan, A., Dwintasari, S. W., Connelly, L., Nichols, R. A., Yunihastuti, E., Karyadi, T., & Djauzi, S. (2013). *Cryptosporidium* species from human immunodeficiency–infected patients with chronic diarrhea in Jakarta, Indonesia. Annals of epidemiology, 23(11), 720-723.

Kvac, M., Kouba, M., & Vitovec, J. (2006). Age-related and housing-dependence of *Cryptosporidium* infection of calves from dairy and beef herds in South Bohemia, Czech Republic. Veterinary Parasitology, 137(3-4), 202–209.

Kváč, M., Hromadová, N., Květoňová, D., Rost, M., & Sak, B. (2011). Molecular characterization of *Cryptosporidium* spp. in pre-weaned dairy calves in the Czech Republic: absence of *C. ryanae* and management-associated distribution of C. *andersoni*, *C. bovis* and *C. parvum* subtypes. Veterinary parasitology, 177(3), 378-382.

Kvalsvig, A., Baker, M.G., Sears, A., French, N. (2014). Bacteria: *Campylobacter*. Encyc. Food Saf, 1, 369-380.

Kwan, P. S. L., Birtles, A., Bolton, F. J., French, N. P., Robinson, S. E., Newbold, L. S., … Fox, A. J. (2008a). Longitudinal study of the molecular epidemiology of *Campylobacter jejuni* in cattle on dairy farms. Applied and Environmental Microbiology, 74(12), 3626–3633.

Kwan, P. S., Barrigas, M., Bolton, F. J., French, N. P., Gowland, P., Kemp, R., ... & Fox, A. J. (2008b). Molecular epidemiology of *Campylobacter jejuni* populations in dairy cattle, wildlife, and the environment in a farmland area. Applied and environmental microbiology, 74(16), 5130-5138.

Lal, A., Baker, M. G., Hales, S., & French, N. P. (2013). Potential effects of global environmental changes on cryptosporidiosis and giardiasis transmission. Trends in parasitology, 29(2), 83-90.

Lalle, M., Jimenez-Cardosa, E., Cacciò, S. M., & Pozio, E. (2005). Genotyping of *Giardia duodenalis* from humans and dogs from Mexico using a β-giardin nested polymerase chain reaction assay. Journal of Parasitology, 91(1), 203-205.

Lane, S., & Lloyd, D. (2002). Current trends in research into the waterborne parasite *Giardia*. Critical Reviews in Microbiology, 28(2), 123–147.

Langhout, M. V, Reed, P., & Fox, M. (2010). Validation of Multiple Diagnostic Techniques to Detect *Cryptosporidium* Sp and *Giardia* Sp in Free-Ranging Western Lowland Gorillas (Gorilla Gorilla Gorilla) and Observations on the Prevalence of These Protozoan Infections in Two Populations in Gabon. Journal of Zoo and Wildlife Medicine, 41(2), 210–217.

Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome biol, 10(3), R25.

Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods, 9, 357-359.

Lasser, K. H., Lewin, K. J., & Ryning, F. W. (1979). Cryptosporidial enteritis in a patient with congenital hypogammaglobulinemia. Human pathology, 10(2), 234-240.

Lázaro, B., Cárcamo, J., Audícana, A., Perales, I., & Fernández-Astorga, A. (1999). Viability and DNA maintenance in nonculturable spiral *Campylobacter jejuni* cells after long-term exposure to low temperatures. Applied and environmental microbiology, 65(10), 4677-4681.

Lazou, T., Houf, K., Soultos, N., Dovas, C., & Iossifidou, E. (2014). *Campylobacter* in small ruminants at slaughter: Prevalence, pulsotypes and antibiotic resistance. International journal of food microbiology, 173, 54-61.

Learmonth, J. J., Ionas, G., Pita, A. B., & Cowie, R. S. (2003). Identification and genetic characterisation of *Giardia* and *Cryptosporidium* strains in humans and dairy cattle in the Waikato Region of New Zealand. Water Science and Technology, 47(3), 21–26.

Learmonth, J. J., Ionas, G., Ebbett, M. A., Kwan, E. S. (2004). Genetic characterization and transmission cycles of *Cryptosporidium* species isolated from humans in New Zealand. Appl Environ Microb 70, 3973-3978.

Lebbad, M., Mattsson, J. G., Christensson, B., Ljungström, B., Backhans, A., Andersson, J. O., & Svärd, S. G. (2010). From mouse to moose: multilocus genotyping of *Giardia* isolates from various animal species. Veterinary parasitology, 168(3), 231-239.

Lechevallier, M. W., Norton, W. D., & Lee, R. G. (1991). Occurrence of *Giardia* and *Cryptosporidium* Spp in Surface-Water Supplies. Applied and Environmental Microbiology, 57(9), 2610–2616.

Lee, S. C., Ngui, R., Tan, T. K., Roslan, M. A., Ithoi, I., & Lim, Y. A. (2014). Aquatic biomonitoring of *Giardia* cysts and *Cryptosporidium* oocysts in peninsular Malaysia. Environmental Science and Pollution Research, 21(1), 445-453.

Lemos, V., Graczyk, T. K., Alves, M., Lobo, M. L., Sousa, M. C., Antunes, F., & Matos, O. (2005). Identification and determination of the viability of *Giardia lamblia* cysts and *Cryptosporidium parvum* and *Cryptosporidium hominis* oocysts in human fecal and water supply samples by fluorescent in situ hybridization (FISH) and monoclonal antibodies. Parasitology Research, 98(1), 48–53.

Leoni, F., Gallimore, C. I., Green, J., & McLauchlin, J. (2006). Characterisation of small double stranded RNA molecule in *Cryptosporidium hominis*, *Cryptosporidium felis* and *Cryptosporidium meleagridis*. Parasitology International, 55(4), 299–306.

Li, H., & Durbin, R. (2010). Fast and accurate long-read alignment with Burrows–Wheeler transform. Bioinformatics, 26(5), 589-595.

Lim, Y. A. L., Ahmad, R. A., & Smith, H. V. (2008). Current status and future trends in *Cryptosporidium* and *Giardia* epidemiology in Malaysia. Journal of water and health, 6(2), 239-254.

Limor, J. R., Lal, A. A., & Xiao, L. H. (2002). Detection and differentiation of *Cryptosporidium* parasites that are pathogenic for humans by real-time PCR. Journal of Clinical Microbiology, 40(7), 2335–2338.

Lindergard, G., Nydam, D. V, Wade, S. E., Schaaf, S. L., & Mohammed, H. O. (2003). A novel multiplex polymerase chain reaction approach for detection of four human infective *Cryptosporidium* isolates: *Cryptosporidium parvum*, types H and C, *Cryptosporidium canis*, and *Cryptosporidium felis* in fecal and soil samples. Journal of Veterinary Diagnostic Investigation, 15(3), 262–267.

Lindsay, D. S., Upton, S. J., Owens, D. S., Morgan, U. M., Mead, J. R., & Blagburn, B. L. (2000). *Cryptosporidium andersoni* n. sp (Apicomplexa : Cryptosporiidae) from cattle, Bos taurus. Journal of Eukaryotic Microbiology, 47(1), 91–95.

Lindström, E. S., Kamst-Van Agterveld, M. P., & Zwart, G. (2005). Distribution of typical freshwater bacterial groups is associated with pH, temperature, and lake water retention time. Applied and environmental microbiology, 71(12), 8201-8206.

Liu, T. L., Tsai, C. H., Lee, W. B., & Chiang, J. H. (2013). Optimizing Information In Next-Generation-Sequencing (Ngs) Reads for Improving De Novo Genome Assembly. Plos One, 8.

Llarena, A. K., Sivonen, K., & Hänninen, M. L. (2014). *Campylobacter jejuni* prevalence and hygienic quality of retail bovine ground meat in Finland. Letters in applied microbiology, 58(5), 408-413.

Lopes, L. D., de Souza Lima, A. O., Taketani, R. G., Darias, P., da Silva, L. R. F., Romagnoli, E. M., … Mendes, R. (2015). Exploring the sheep rumen microbiome for carbohydrate-active enzymes. Antonie van Leeuwenhoek, 1–16.

Lorenz, E., Lastovica, A., & Owen, R. J. (1998). Subtyping of *Campylobacter jejuni* Penner serotypes 9, 38 and 63 from human infections, animals and water by pulsed field gel electrophoresis and flagellin gene analysis. Lett Appl Microbiol, 26, 179-182.

Lorenzo, M. J. L., Aresmazas, E., & Dematurana, I. V. M. (1993). Detection of Oocysts and Igg Antibodies to *Cryptosporidium*-Parvum in Asymptomatic Adult Cattle. Veterinary Parasitology, 47(1-2), 9–15.

Luján, H. D., & Svärd, S. (Eds.). (2011). *Giardia*: A model organism. Springer Science & Business Media.

Lunn, D. J., Thomas, A., Best, N., & Spiegelhalter, D. (2000). WinBUGS-a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and computing, 10(4), 325-337.

Lymer, D., Logue, J. B., Brussaard, C. P., BAUDOUX, A. C., Vrede, K., & LINDSTRÖM, E. S. (2008). Temporal variation in freshwater viral and bacterial community composition. Freshwater Biology, 53(6), 1163-1175.

Ma, P., & Soave, R. (1983). 3-Step Stool Examination for Cryptosporidiosis in 10 Homosexual Men with Protracted Watery Diarrhea. Journal of Infectious Diseases, 147(5), 824–828.

Mackenzie, W. R. (1994). A Massive Outbreak in Milwaukee of *Cryptosporidium* Infection Transmitted through the Public Water-Supply (Vol 331, Pg 161, 1994). New England Journal of Medicine, 331(15), 1035-1035.

Maddox-Hyttel, C., Langkjær, R. B., Enemark, H. L., & Vigre, H. (2006). *Cryptosporidium* and *Giardia* in different age groups of Danish cattle and pigs—occurrence and management associated risk factors. Veterinary parasitology, 141(1), 48-59.

Magoč, T., & Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. Bioinformatics, 27(21), 2957-2963.

Maiden, M. F. J., Tanner, A. C. R., Macuch, P. J., Murray, L., & Kent, R. L. (1998). Subgingival temperature and microbiota in initial periodontitis. Journal of Clinical Periodontology, 25(10), 786–793.

Maiden, M. C. J. (2006). Multilocus sequence typing of bacteria. Annual Review of Microbiology, 60, 561–588.

Maiden, M. C., & Dingle, K. E. (2008). Population biology of *Campylobacter jejuni* and related organisms. *Campylobacter*, 3, 27-40.

Maikai, B. V., Umoh, J. U., Kwaga, J. K., Lawal, I. A., Maikai, V. A., Cama, V., & Xiao, L. (2011). Molecular characterization of *Cryptosporidium* spp. in native breeds of cattle in Kaduna State, Nigeria. Veterinary parasitology, 178(3), 241-245.

Mallon, M. E., MacLeod, A., Wastling, J. M., Smith, H., & Tait, A. (2003a). Multilocus genotyping of *Cryptosporidium parvum* Type 2: population genetics and sub-structuring. Infection Genetics and Evolution, 3(3), 207–218.

Mallon, M., MacLeod, A., Wastling, J., Smith, H., Reilly, B., & Tait, A. (2003b). Population structures and the role of genetic exchange in the zoonotic pathogen *Cryptosporidium parvum* (vol 56, pg 407, 2003). Journal of Molecular Evolution, 56(6), 778.

Man, S. M. (2011). The clinical importance of emerging *Campylobacter* species. Nature Reviews Gastroenterology and Hepatology, 8(12), 669-685.

Mangrola, A., Dudhagara, P., Koringa, P., Joshi, C. G., Parmar, M., & Patel, R. (2015). Deciphering the microbiota of Tuwa hot spring, India using shotgun metagenomic sequencing approach. Genomics data, 4, 153-155.

Manning, G., Dowson, C. G., Bagnall, M. C., Ahmed, I. H., West, M., & Newell, D. G. (2003). Multilocus sequence typing for comparison of veterinary and human isolates of *Campylobacter jejuni*. Applied and Environmental Microbiology, 69(11), 6370-6379.

Mannering, S. A., West, D. M., Fenwick, S. G., Marchant, R. M., & O'Connell, K. (2006). Pulsed-field gel electrophoresis of *Campylobacter jejuni* sheep abortion isolates. Veterinary Microbiology, 115(1-3), 237–242.

Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. Trends in Genetics : TIG, 24(3), 133–41.

Martin-Cuadrado, A.-B., Garcia-Heredia, I., Moltó, A. G., López-Úbeda, R., Kimes, N., López-García, P., … Rodriguez-Valera, F. (2014). A new class of marine Euryarchaeota group II from the mediterranean deep chlorophyll maximum. The ISME Journal.

Mason, R. W., Hartley, W. J., & Tilt, L. (1981). Intestinal Cryptosporidiosis in a Kid Goat. Australian Veterinary Journal, 57(8), 386–388.

Mason, M. J., Humphrey, T. J., & Martin, K. W. (1999). Isolation of sublethally injured campylobacters from poultry and water sources. British journal of biomedical science, 56(1), 2.

Maurya, P. S., Rakesh, R. L., Pradeep, B., Kumar, S., Kundu, K., Garg, R., … Banerjee, P. S. (2013). Prevalence and risk factors associated with *Cryptosporidium* spp. infection in young domestic livestock in India. Tropical Animal Health and Production, 45(4), 941–946.

Mawdsley, J. L., Bardgett, R. D., Merry, R. J., Pain, B. F., & Theodorou, M. K. (1995). Pathogens in Livestock Waste, Their Potential for Movement through Soil and Environmental-Pollution. Applied Soil Ecology, 2(1), 1–15.

McAllister, T. A., & Topp, E. (2012). Role of livestock in microbiological contamination of water: commonly the blame, but not always the source. Animal Frontiers, 2(2), 17–27.

McCool, C. J., Townsend, M. P., Wolfe, S. G., Simpson, M. A., Olm, T. C., Jayawardhana, G. A., & Carney, J. V. (1988). Prevalence of bovine veneral disease in the Victoria River District of the Northern Territory: likely economic effects and practicable control measures. Australian veterinary journal, 65(5), 153-156.

McCuin, R. M., Bukhari, Z., Sobrinho, J., & Clancy, J. L. (2001). Recovery of *Cryptosporidium* oocysts and *Giardia* cysts from source water concentrates using immunomagnetic separation. Journal of Microbiological Methods, 45(2), 69–76.

McKergow, L. A., Tanner, C. C., Monaghan, R. M., & Anderson, G. (2007). Stocktake of diffuse pollution attenuation tools for New Zealand pastoral farming systems. NIWA Client Report HAM2007-161 to the Pastoral 21 Research Consortium, Hamilton, NZ.

McLauchlin, J., Pedraza-Diaz, S., Amar-Hoetzeneder, C., & Nichols, G. L. (1999). Genetic characterization of *Cryptosporidium* strains from 218 patients with diarrhea diagnosed as having sporadic cryptosporidiosis. Journal of Clinical Microbiology, 37(10), 3153–3158.

McLauchlin, J., Amar, C., Pedraza-Diaz, S., & Nichols, G. L. (2000). Molecular Epidemiological Analysis of Cryptosporidiumspp. in the United Kingdom: Results of GenotypingCryptosporidium spp. in 1,705 Fecal Samples from Humans and 105 Fecal Samples from Livestock Animals. Journal of Clinical Microbiology, 38(11), 3984-3990.

McNabb, S. J. N., Hensel, D. M., Welch, D. F., Heijbel, H., Mckee, G. L., & Istre, G. R. (1985). Comparison of Sedimentation and Flotation Techniques for Identification of *Cryptosporidium* Sp Oocysts in a Large Outbreak of Human Diarrhea. Journal of Clinical Microbiology, 22(4), 587–589.

McTavish, S. M., Pope, C. E., Nicol, C., Sexton, K., French, N., & Carter, P. E. (2008). Wide geographical distribution of internationally rare *Campylobacter* clones within New Zealand. Epidemiology and Infection, 136(09), 1244-1252.

Medema, G. J., Schets, F. M., Vandegiessen, A. W., & Havelaar, A. H. (1992). Lack of Colonization of 1-Day-Old Chicks by Viable, Non-Culturable *Campylobacter-jejuni*. Journal of Applied Bacteriology, 72(6), 512–516.

Meireles, M. V., de Oliveira, F. P., Teixeira, W. F. P., Coelho, W. M., & Mendes, L. C. N. (2011). Molecular characterization of *Cryptosporidium* spp. in dairy calves from the state of São Paulo, Brazil. Parasitology research, 109(3), 949-951.

Meisel, J. L., Perera, D. R., Meligro, C., & Rubin, C. E. (1976). Overwhelming watery diarrhea associated with a cryptosporidium in an immunosuppressed patient. Gastroenterology, 70(6), 1156-1160.

Meldrum, R. J., Griffiths, J. K., Smith, R. M. M., & Evans, M. R. (2005). The seasonality of human campylobacter infection and *Campylobacter* isolates from fresh, retail chicken in Wales. Epidemiology and Infection, 133(01), 49-52.

Mellmann, A., Harmsen, D., Cummings, C. A., Zentz, E. B., Leopold, S. R., Rico, A., Prior, K., Szczepanowski, R., Ji, Y., & Zhang, W. (2011). Prospective genomic characterization of the German enterohemorrhagic *Escherichia coli* O104: H4 outbreak by rapid next generation sequencing technology. Plos One, 6, e22751.

Mendonça, C., Almeida, A., Castro, A., de Lurdes Delgado, M., Soares, S., da Costa, J. M. C., & Canada, N. (2007). Molecular characterization of *Cryptosporidium* and *Giardia* isolates from cattle from Portugal. Veterinary parasitology, 147(1), 47-50.

Metzker, M. L. (2005). Emerging technologies in DNA sequencing. Genome Research, 15, 1767-1776.

Metzker, M. L. (2010). Sequencing technologies - the next generation. Nature Reviews. Genetics, 11(1), 31–46.

MfE/MoH. 2003. Microbiological water quality guidelines for marine and freshwater recreational areas. Published by the Ministry for the Environment and Ministry of Health, Wellington, New Zealand.

Mi, R., Wang, X., Huang, Y., Zhou, P., Liu, Y., Chen, Y., ... & Chen, Z. (2014). Prevalence and molecular characterization of *Cryptosporidium* in goats across four provincial level areas in China. PloS one, 9(10), e111164.

Miller, W. G., On, S. L., Wang, G., Fontanoz, S., Lastovica, A. J. & Mandrell, R. E. (2005). Extended multilocus sequence typing system for *Campylobacter coli*, *C. lari*, *C. upsaliensis*, and *C. helveticus*. Journal of Clinical Microbiology, 43, 2315-2329.

Minetti, C., Lamden, K., Durband, C., Cheesbrough, J., Fox, A., & Wastling, J. M. (2015). Determination of *Giardia duodenalis* assemblages and multi-locus genotypes in patients with sporadic giardiasis from England. Parasites & vectors, 8(1), 1-10.

Minetti, C., Taweenan, W., Hogg, R., Featherstone, C., Randle, N., Latham, S. M., & Wastling, J. M. (2014). Occurrence and diversity of *Giardia duodenalis* assemblages in livestock in the UK. Transboundary and emerging diseases, 61(6), e60-e67.

Minihan, D., Whyte, P., O'Mahony, M., Fanning, S., Doyle, M., & Collins, J. D. (2004a). An investigation of transport, lairage and hide cleanliness on *Campylobacter* prevalence in feedlot cattle and dressed carcasses. Journal of Food Safety, 24(1), 37–52.

Minihan, D., Whyte, P., O'Mahony, M., Fanning, S., McGill, K., & Collins, J. D. (2004b). *Campylobacter* spp. in Irish feedlot cattle: A longitudinal study involving pre-harvest and harvest phases of the food chain. Journal of Veterinary Medicine Series B-Infectious Diseases and Veterinary Public Health, 51(1), 28–33.

Mirhashemi, M. E., Zintl, A., Grant, T., Lucy, F., Mulcahy, G., & De Waal, T. (2016). Molecular epidemiology of *Cryptosporidium* species in livestock in Ireland. Veterinary Parasitology, 216, 18-22.

Mohammed, H. O., Wade, S. E., & Schaaf, S. (1999). Risk factors associated with *Cryptosporidium parvum* infection in dairy cattle in southeastern New York State. Veterinary Parasitology, 83(1), 1–13.

Mohran, Z. S., Arthur, R. R., Oyofo, B. A., Peruski, L. F., Wasfy, M. O., Ismail, T. F., & Murphy, J. R. (1998). Differentiation of *Campylobacter* isolates on the basis of sensitivity to boiling in water as measured by PCR-detectable DNA. Applied and Environmental Microbiology, 64(1), 363–365.

Molbak, K., Andersen, M., Aaby, P., Hojlyng, N., Jakobsen, M., Sodemann, M., & daSilva, A. P. J. (1997). *Cryptosporidium* infection in infancy as a cause of malnutrition: A community study from Guinea-Bissau, West Africa. American Journal of Clinical Nutrition, 65(1), 149–152.

Monis, P. T., & Thompson, R. C. A. (2003). *Cryptosporidium* and *Giardia*-zoonoses: fact or fiction?. Infection, genetics and evolution, 3(4), 233-244.

Monis, P. T., Andrews, R. H., Mayrhofer, G., & Ey, P. L. (2003). Genetic diversity within the morphological species *Giardia intestinalis* and its relationship to host origin. Infection, Genetics and Evolution, 3(1), 29-38.

Moodley, D., Jackson, T. F. H. G., Gathiram, V., & Vandenende, J. (1991). *Cryptosporidium* Infections in Children in Durban - Seasonal-Variation, Age Distribution and Disease Status. South African Medical Journal, 79(6), 295–297.

Moore, J., Caldwell, P., Millar, B. (2001). Molecular detection of *Campylobacter* spp. in drinking, recreational and environmental water supplies. Int J Hyg Envir Heal, 204, 185-189.

Moore, J. E., Wilson, T. S., Wareing, D. R. A., & Murphy, P. G. (2003). Occurence and characterisation of thermophilic *Campylobacter* spp. in a poultry processing plant in Northern Ireland. Irish Veterinary Journal, 56(2), 95–98.

Morgavi, D. P., Kelly, W. J., Janssen, P. H., & Attwood, G. T. (2013). Rumen microbial (meta) genomics and its application to ruminant production. Animal, 7(s1), 184–201.

Moriarty, E. M., Sinton, L. W., Mackenzie, M. L., Karki, N.& Wood, D. R. (2008). A survey of enteric bacteria and protozoans in fresh bovine faeces on New Zealand dairy farms. Journal of Applied Microbiology, 105(6), 2015-2025.

Moriarty, E. M., McEwan, N., Mackenzie, M., Karki, N., & Sinton, L. W. (2011a). Incidence and prevalence of microbial indicators and pathogens in ovine faeces in New Zealand. New Zealand Journal of Agricultural Research, 54(2), 71–81.

Moriarty, E. M., Mackenzie, M. L., Karki, N., & Sinton, L. W. (2011b). Survival of *Escherichia coli*, *Enterococci*, and *Campylobacter* spp. in Sheep Feces on Pastures. Applied and Environmental Microbiology, 77(5), 1797–1803.

Morse, T. D., Nichols, R. A. B., Grimason, A. M., Campbell, B. M., Tembo, K. C., & Smith, H. V. (2007). Incidence of cryptosporidiosis species in paediatric patients in Malawi. Epidemiology and Infection, 135(8), 1307–1315.

Motarjemi Y. 2014. Encylopedia of food safety, p. 2304. In A K, MG Baker, A Sears and N French (ed.), *Campylobacter*, First ed. Academic Press.

Mshelia, G. D., Amin, J. D., Woldehiwet, Z., Murray, R. D., & Egwu, G. O. (2010). Epidemiology of Bovine Venereal Campylobacteriosis: Geographic Distribution and Recent Advances in Molecular Diagnostic Techniques. Reproduction in Domestic Animals, 45(5), E221–E230.

Mtambo, M. M. A., Nash, A. S., Blewett, D. A., & Wright, S. (1992). Comparison of Staining and Concentration Techniques for Detection of *Cryptosporidium* Oocysts in Cat Fecal Specimens. Veterinary Parasitology, 45(1-2), 49–57.

Mueller-Doblies, D., Giles, M., Elwin, K., Smith, R. P., Clifton-Hadley, F. A., & Chalmers, R. M. (2008). Distribution of *Cryptosporidium* species in sheep in the UK. Veterinary Parasitology, 154(3), 214-219.

Muhid, A., Robertson, I., Ng, J., & Ryan, U. (2011). Prevalence of and management factors contributing to *Cryptosporidium* sp infection in pre-weaned and post-weaned calves in Johor, Malaysia. Experimental Parasitology, 127(2), 534–538.

Mullner, P., Jones, G., Noble A., Spencer, S. E. F., Hathaway, S., French, N. P. (2009). Source Attribution of Food Borne Zoonoses in New Zealand: A Modified Hald Model. Risk Anal 29:970-984.

Mullner, P., Shadbolt, T., Collins-Emerson, J., Midwinter, A., Spencer, S., Marshall, J., Carter, P., Campbell, D., Wilson, D.& Hathaway, S. (2010). Molecular and spatial epidemiology of human campylobacteriosis: source association and genotype-related risk factors. Epidemiology and Infection, 138(10), 1372-1383.

Mullner, P., Marshall, J. C., Spencer, S. E. F., Noble, A. D., Shadbolt, T., Collins-Emerson, J. M., Midwinter, A. C., Carter, P. E., Pirie, R., Wilson, D. J., Campbell, D. M., Stevenson, M. A., French, N. P. (2011). Utilizing a combination of molecular and spatial tools to assess the effect of a public health intervention. Prev Vet Med 102:242-253.

Mullner, P., Pleydell, E., Pirie, R., Baker, M. G., Campbell, D., Carter, P. E., & French, N. P. (2013). Molecular-based surveillance of campylobacteriosis in New Zealand--from source attribution to genomic epidemiology. Euro Surveillance : Bulletin Européen Sur Les Maladies Transmissibles = European Communicable Disease Bulletin, 18(3), 1–7.

Munoz, M., Alvarez, M., Lanza, I., & Carmenes, P. (1996). Role of enteric pathogens in the aetiology of neonatal diarrhoea in lambs and goat kids in Spain. Epidemiology and Infection, 117(1), 203–211.

Murakoshi, F., Xiao, L. H., Matsubara, R., Sato, R., Kato, Y., Sasaki, T., … Nakai, Y. (2012). Molecular characterization of *Cryptosporidium* spp. in grazing beef cattle in Japan. Veterinary Parasitology, 187(1-2), 123–128.

Murphy, L. R., Wallqvist, A., & Levy, R. M. (2000). Simplified amino acid alphabets for protein fold recognition and implications for folding. Protein Engineering, 13, 149-152.

Murphy, C., Carroll, C., & Jordan, K. N. (2006). Environmental survival mechanisms of the foodborne pathogen *Campylobacter jejuni*. Journal of Applied Microbiology, 100(4), 623–632.

Nachamkin, I., Szymanski, C. M.& Blaser, M. J. (Eds.). (2008). *Campylobacter* (3rd ed.). Washington, DC: ASM Press.

Nash, T. E., Herrington, D. A., Losonsky, G. A., & Levine, M. M. (1987). Experimental human infections with *Giardia lamblia*. Journal of Infectious Diseases, 156(6), 974-984.

Navarro-i-Martinez, L., da Silva, A. J., Bornay-Llinares, F. J., Moura, I. N., del Aguila, C., Oleaga, A., & Pieniazek, N. J. (2007). Detection and molecular characterization of *Cryptosporidium bovis*-like isolate from a newborn lamb in Spain. Journal of Parasitology, 93(6), 1536-1538.

Ng, J., Yang, R., McCarthy, S., Gordon, C., Hijjawi, N., Ryan, U. (2011). Molecular characterization of *Cryptosporidium* and *Giardia* in pre-weaned calves in Western Australia and New South Wales. Vet Parasitol 176, 145-150.

Ng-Hublin, J. S. Y., Combs, B., MacKenzie, B., & Ryan, U. (2013). Human Cryptosporidiosis Diagnosed in Western Australia: a Mixed Infection with *Cryptosporidium* meleagridis, the *Cryptosporidium* Mink Genotype, and an Unknown *Cryptosporidium* Species. Journal of Clinical Microbiology, 51(7), 2463–2465.

Nichols, R. A. B., Moore, J. E., & Smith, H. V. (2006). A rapid method for extracting oocyst DNA from *Cryptosporidium*-positive human faeces for outbreak investigations. Journal of Microbiological Methods, 65(3), 512–524.

Nielsen, E. M. (2002). Occurrence and strain diversity of thermophilic campylobacters in cattle of different age groups in dairy herds. Letters in Applied Microbiology, 35(1), 85-89.

Nielsen, E. M., Engberg, J., Fussing, V., Petersen, L., Brogren, C. H., On, S. L. W. (2000). Evaluation of phenotypic and genotypic methods for subtyping *Campylobacter jejuni* isolates from humans, poultry, and cattle. J Clin Microbiol, 38, 3800-3810.

Nieminski, E. C., Schaefer, F. W., & Ongerth, J. E. (1995). Comparison of 2 Methods for Detection of *Giardia* Cysts and *Cryptosporidium* Oocysts in Water. Applied and Environmental Microbiology, 61(5), 1714–1719.

Nime, F. A., Burek J. D., Page, D. L., Holscher, M. A. & Yardley, J. H. (1976). Gastroenterology, 70, 592-598.

NIWA. (2006). New Zealand River Environment Classification Database. Ministry for the Environment. Wellington.

NIWA. (2010). New Zealand River Environment Classification Database. Ministry for the Environment. Wellington.

NNDSS Annual Report Writing Group. (2012). Australia's notifiable disease status, 2010: annual report of the national notifiable diseases surveillance system. In M. Curran (Ed.), Communicable Diseases Intelligence (Vol. 36, pp. 26–30).

Nylen, G., Dunstan, F., Palmer, S. R., Andersson, Y., Bager, F., Cowden, J., ... & Molbak, K. (2002). The seasonal distribution of *Campylobacter* infection in nine European countries and New Zealand. Epidemiology and Infection, 128(03), 383-390.

NZFSA, 2010. The economic cost of foodborne disease in New Zealand. Applied Economics Pty Ltd. New Zealand Food Safety Authority.

O'Connor, D.R. (2002). Report of the Walkerton Inquiry: Part 1 – The Events of May 2000 and Related Issues. The Walkerton Inquiry, Toronto,504 pp.

O'Donoghue, P. J. (1995). *Cryptosporidium* and Cryptosporidiosis in Man and Animals. International Journal for Parasitology, 25(2), 139–195.

Oberhuber, G., Lauer, E., Stolte, M., & Borchard, F. (1991). Cryptosporidiosis of the Appendix-Vermiformis - a Case-Report. Zeitschrift Fur Gastroenterologie, 29(11), 606–608.

Obiri-Danso, K., Paul, N., Jones, K. (2001). The effects of UVB and temperature on the survival of natural populations and pure cultures of *Campylobacter jejuni*, Camp. *coli*, Camp. lari and urease-positive thermophilic campylobacters (UPTC) in surface waters. J Appl Microbiol, 90, 256-267.

O'handley, R. M., Cockwill, C., McAllister, T. A., Jelinski, M., Morck, D. W., & Olson, M. E. (1999). Duration of naturally acquired giardiosis and cryptosporidiosis in dairy calves and their association with diarrhea. Journal of the American Veterinary Medical Association, 214(3), 391-396.

O'handley, R. M., Olson, M. E., Fraser, D., Adams, P., & Thompson, R. C. A. (2000). Prevalence and genotypic characterisation of *Giardia* in dairy calves from Western Australia and Western Canada. Veterinary parasitology, 90(3), 193-200.

O'handley, R. M., & Olson, M. E. (2006). Giardiasis and cryptosporidiosis in ruminants. Veterinary Clinics of North America: Food Animal Practice, 22(3), 623-643.

Okun, D. (1996). From Cholera to Cancer to Cryptosporidiosis. J. Environ. Eng., 122:6(453), 453-458.

Olson, M. E., McAllister, T. A., Deselliers, L., Morck, D. W., Cheng, K. J., Buret, A. G., & Ceri, H. (1995). Effects of giardiasis on production in a domestic ruminant (lamb) model. American journal of veterinary research, 56(11), 1470-1474.

Olson, M. E., Guselle, N. J., O'Handley, R. M., Swift, M. L., McAllister, T. A., Jelinski, M. D., & Morck, D. W. (1997a). *Giardia* and *Cryptosporidium* in dairy calves in British Columbia. The Canadian Veterinary Journal, 38(11), 703.

Olson, M. E., Thorlakson, C. L., Deselliers, L., Morck, D. W., & McAllister, T. A. (1997b). *Giardia* and *Cryptosporidium* in Canadian farm animals. Veterinary parasitology, 68(4), 375-381.

Olson, M. E., O'Handley, R. M., Ralston, B. J., McAllister, T. A., & Thompson, R. A. (2004). Update on *Cryptosporidium* and *Giardia* infections in cattle. Trends in parasitology, 20(4), 185-191.

Ongerth, J. E., & Stibbs, H. H. (1987). Identification of *Cryptosporidium* Oocysts in River Water. Applied and Environmental Microbiology, 53(4), 672–676.

Ongerth, J. E., & Saaed, F. M. (2013). Distribution of *Cryptosporidium* oocysts and *Giardia* cysts in water above and below the normal limit of detection. Parasitology research, 112(2), 467-471.

Ontario. Ministry of the Attorney General, & O'Connor, D. R. (2002). Report of the Walkerton Inquiry: the Events of May 2000 and Related Issues: a Summary. Ontario Ministry of the Attorney General.

Oporto, B., Esteban, J. I., Aduriz, G., Juste, R. A., & Hurtado, A. (2007). Prevalence and strain diversity of thermophilic campylobacters in cattle, sheep and swine farms. Journal of applied microbiology, 103(4), 977-984.

Orr, M. B., Mackintosh, C. G., & Suttie, J. M. (1985). Cryptosporidiosis in Deer Calves. New Zealand Veterinary Journal, 33(9), 151–152.

Ortega-Mora, L. M., & Wright, S. E. (1994). Age-Related Resistance in Ovine Cryptosporidiosis - Patterns of Infection and Humoral Immune-Response. Infection and Immunity, 62(11), 5003–5009.

Osbjer, K., Tano, E., Chhayheng, L., Mac-Kwashie, A. O., Fernström, L. L., Ellström, P., ... & San, S. (2016). Detection of *Campylobacter* in human and animal field samples in Cambodia. APMIS.

Panciera, R. J., Thomassen, R. W., & Garner, F. M. (1971). Cryptosporidial infection in a calf. Veterinary Pathology Online, 8(5-6), 479-484.

Pao, S., Hagens, B. E., Kim, C., Wildeus, S., Ettinger, M. R., Wilson, M. D., ... & Kaseloo, P. (2014). Prevalence and molecular analyses of *Campylobacter jejuni* and Salmonella spp. in co-grazing small ruminants and wild-living birds. Livestock Science, 160, 163-171.

Paoletti, B., Giangaspero, A., Gatti, A., Iorio, R., Cembalo, D., Milillo, P., & Traversa, D. (2009). Immunoenzymatic analysis and genetic detection of *Cryptosporidium parvum* in lambs from Italy. Experimental Parasitology, 122(4), 349–352.

Paraud, C., Pors, I., & Chartier, C. (2010). Evaluation of oral tilmicosin efficacy against severe cryptosporidiosis in neonatal kids under field conditions. Veterinary Parasitology, 170(1-2), 149–152.

Pareek, C. S., Smoczynski, R., & Tretyn, A. (2011). Sequencing technologies and genome sequencing. Journal of applied genetics, 52(4), 413-435.

Pasupuleti, V., Escobedo, A. A., Deshpande, A., Thota, P., Roman, Y., & Hernandez, A. V. (2014). Efficacy of 5-nitroimidazoles for the treatment of giardiasis: a systematic review of randomized controlled trials. PLoS Negl Trop Dis, 8(3), e2733.

Patel, V., Patel, A. K., Parmar, N. R., Patel, A. B., Reddy, B., & Joshi, C. G. (2014). Characterization of the rumen microbiome of Indian Kankrej cattle (Bos indicus) adapted to different forage diet. Applied Microbiology and Biotechnology, 98(23), 9749–9761.

Payment, P., Plante, R., & Cejka, P. (2001). Removal of indicator bacteria, human enteric viruses, *Giardia* cysts, and *Cryptosporidium* oocysts at a large wastewater primary treatment facility. Canadian Journal of Microbiology, 47(3), 188-193.

Paz e Silva, F. M., Lopes, R. S., Bresciani, K. D. S., Amarante, A. F. T., & Araujo, J. P. (2014). High occurrence of *Cryptosporidium ubiquitum* and *Giardia duodenalis* genotype E in sheep from Brazil. Acta Parasitologica / Witold Stefański Institute of Parasitology, Warszawa, Poland, 59(1), 193–6.

Peng, M. M., Matos, O., Gatei, W., Das, P., Stantic-Pavlinic, M., Bern, C., Sulaiman, I.M., Glaberman, S., Lal, A.A., Xiao, L. H. (2001). A comparison of *Cryptosporidium* subgenotypes from several geographic regions. Journal of Eukaryotic Microbiology, 28s–31s.

Peng, X. Q., Tian, G. R., Ren, G. J., Yu, Z. Q., Lok, J. B., Zhang, L. X., ... & Zhao, G. H. (2016). Infection rate of *Giardia duodenalis*, *Cryptosporidium* spp. and Enterocytozoon bieneusi in cashmere, dairy and meat goats in China. Infection, Genetics and Evolution, 41, 26-31.

Pereira, T. D. G., Atwill, E. R., & Jones, T. (1999). Comparison of sensitivity of immunofluorescent microscopy to that of a combination of immunofluorescent microscopy and immunomagnetic separation for detection of *Cryptosporidium parvum* oocysts in adult bovine feces. Applied and Environmental Microbiology, 65(7), 3236–3239.

Pérez-Losada, M., Cabezas, P., Castro-Nallar, E., & Crandall, K. A. (2013). Pathogen typing in the genomics era: MLST and the future of molecular epidemiology. Infection, Genetics and Evolution, 16, 38-53.

Piccirillo, A., Giacomelli, M., Salata, C., Bettanello, S., De Canale, E., & Palù, G. (2014). Multilocus sequence typing of *Campylobacter jejuni* and *Campylobacter coli* from humans and chickens in North-Eastern Italy. New Microbiologica, 37(4), 557-562.

Plutzer, J., & Karanis, P. (2007). Genotype and subtype analyses of *Cryptosporidium* isolates from cattle in Hungary. Veterinary Parasitology, 146(3-4), 357–362.

Plutzer, J., Tako, M. H., Marialigeti, K., Torokne, A., & Karanis, P. (2007). First investigations into the prevalence of *Cryptosporidium* and *Giardia* spp. in Hungarian drinking water. Journal of Water and Health, 5(4), 573–584.

Power, M. L., Shanker, S. R., Sangster, N. C., & Veal, D. A. (2003). Evaluation of a combined immunomagnetic separation/flow cytometry technique for epidemiological investigations of *Cryptosporidium* in domestic and Australian native animals. Veterinary Parasitology, 112(1-2), 21–31.

Prattley, D., French, N. P., Kwan, E. & Pita, A. (2010). Gastro-Intestinal Protozoa Research and Services (Vol. 15). Palmerston North: Massey University.

Pritchard, G. C., Marshall, J. A., Giles, M., Chalmers, R. M., & Marshall, R. N. (2007). *Cryptosporidium parvum* infection in orphan lambs on a farm open to the public. Veterinary Record: Journal of the British Veterinary Association, 161(1).

Pritchard, G. C., Marshall, J. A., Giles, M., Mueller-Doblies, D., Sayers, A. R., Marshall, R. N., ... & Chalmers, R. M. (2008). *Cryptosporidium* species in lambs submitted for diagnostic postmortem examination in England and Wales. Veterinary Record, 163(23), 688-689.

Quilez, J., Sanchez-Acebo, C., Clavel, A., Del Cacho, E., & Lopez-Bernad, F. (1996). Comparison of an acid-fast stain and a monoclonal antibody-based immunofluorescence reagent for the detection of *Cryptosporidium* oocysts in faecal specimens from cattle and pigs. Veterinary parasitology, 67(1), 75-81.

R Core Team. (2013). A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ralston, B. J., Cockwill, C. L., Guselle, N. J., Van Herk, F. H., McAllister, T. A., & Olson, M. E. (2003). Prevalence of *Giardia* and *Cryptosporidium andersoni* and their effects on performance in feedlot beef cattle. Canadian Journal of Animal Science, 83(1), 153–159.

Ralston, B., Thompson, R. C. A., Pethick, D., McAllister, T. A., & Olson, M. E. (2010). *Cryptosporidium andersoni* in Western Australian feedlot cattle. Australian Veterinary Journal, 88(11), 458–460.

Ramette, A. (2007). Multivariate analyses in microbial ecology. Fems Microbiology Ecology, 62, 142-160.

Ramonaite, S., Rokaityte, A., Tamuleviciene, E., Malakauskas, A., Alter, T., & Malakauskas, M. (2013). Prevalence, quantitative load and genetic diversity of *Campylobacter* spp. in dairy cattle herds in Lithuania. Acta Veterinaria Scandinavica, 55.

Rapp, D., & Ross, C. (2012). Prevalence of six *Campylobacter* species in a New Zealand dairy goat herd. New Zealand Journal of Agricultural Research, 55(3), 235–240.

Rapp, D., Ross, C. M., Pleydell, E. J., & Muirhead, R. W. (2012). Differences in the fecal concentrations and genetic diversities of *Campylobacter jejuni* populations among

individual cows in two dairy herds. Applied and Environmental Microbiology, 78(21), 7564–7571.

Rapp, D., Ross, C. M., Cave, V., & Muirhead, R. W. (2014). Prevalence, concentration and genotypes of *Campylobacter jejuni* in faeces from dairy herds managed in farm systems with or without housing. Journal of applied microbiology, 116(4), 1035-1043.

Read, C., Walters, J., Robertson, I. D., & Thompson, R. C. A. (2002). Correlation between genotype of Giardia*duodenalis* and diarrhoea. International journal for parasitology, 32(2), 229-231.

Read, C. M., Monis, P. T., & Thompson, R. A. (2004). Discrimination of all genotypes of *Giardia duodenalis* at the glutamate dehydrogenase locus using PCR-RFLP. Infection, Genetics and Evolution, 4(2), 125-130.

Reed, C., Sturbaum, G. D., Hoover, P. J., Sterling, C. R. (2002). *Cryptosporidium parvum* mixed genotypes detected by PCR-restriction fragment length polymorphism analysis. Applied and Environmental Microbiology, 68, 427-429.

Reinoso, R., Becares, E., & Smith, H. V. (2008). Effect of various environmental factors on the viability of *Cryptosporidium parvum* oocysts. Journal of applied microbiology, 104(4), 980-986.

Reuter, M., Mallett, A., Pearson, B. M., van Vliet, A. H. M. (2010). Biofilm formation by *Campylobacter jejuni* is increased under aerobic conditions. Applied and Environmental Microbiology, 76, 2122-2128.

Revetta, R. P., Gomez-Alvarez, V., Gerke, T. L., Curioso, C., Santo Domingo, J. W., & Ashbolt, N. J. (2013). Establishment and early succession of bacterial communities in monochloramine-treated drinking water biofilms. FEMS Microbiology Ecology, 86(3), 404–414.

Rieux, A., Chartier, C., Pors, I., Delafosse, A., Paraud, C. (2013). Molecular characterization of *Cryptosporidium* isolates from high-excreting young dairy calves in dairy cattle herds in Western France. Parasitol Res 112, 3423-3431.

Rinaldi, L., Musella, V., Condoleo, R., Saralli, G., Veneziano, V., Bruni, G., ... & Cringoli, G. (2007). *Giardia* and *Cryptosporidium* in water buffaloes (Bubalus bubalis). Parasitology research, 100(5), 1113-1118.

Rind, E., & Pearce, J. (2010). The spatial distribution of campylobacteriosis in New Zealand, 1997–2005. Epidemiology and infection, 138(10), 1359-1371.

Robertson, L. J., Campbell, A. T., & Smith, H. V. (1992). Survival of *Cryptosporidium parvum* oocysts under various environmental pressures. Applied and Environmental Microbiology, 58(11), 3494-3500.

Robertson, L. J., & Gjerde, B. K. (2006). Fate of *Cryptosporidium* oocysts and *Giardia* cysts in the Norwegian aquatic environment over winter. Microbial ecology, 52(4), 597-602.

Robertson, L. J. (2009). *Giardia* and *Cryptosporidium* infections in sheep and goats: a review of the potential for transmission to humans via environmental contamination. Epidemiology and infection, 137(07), 913-921.

Robertson, L. J., Gjerde, B. K., & Hansen, E. F. (2010). The zoonotic potential of *Giardia* and *Cryptosporidium* in Norwegian sheep: a longitudinal investigation of 6 flocks of lambs. Veterinary parasitology, 171(1), 140-145.

Robertson, L. J., & Lim, Y. A. L. (2011). Waterborne and environmentally-borne giardiasis (pp. 29-69). Springer Vienna.

Robinson, G., Chalmers, R. M., Stapleton, C., Palmer, S. R., Watkins, J., Francis, C., & Kay, D. (2011). A whole water catchment approach to investigating the origin and distribution of *Cryptosporidium* species. Journal of applied microbiology, 111(3), 717-730.

Rochelle, P. A., De Leon, R., Johnson, A., Stewart, M. H., & Wolfe, R. L. (1999). Evaluation of immunomagnetic separation for recovery of infectious *Cryptosporidium parvum* oocysts from environmental samples. Applied and environmental microbiology, 65(2), 841-845.

Rodgers, M. R., Flanigan, D. J., & Jakubowski, W. (1995). Identification of Algae Which Interfere with the Detection of *Giardia* Cysts and *Cryptosporidium* Oocysts and a Method for Alleviating This Interference. Applied and Environmental Microbiology, 61(10), 3759–3763.

Roig, B., Delpla, I., Baurès, E., Jung, A. V, & Thomas, O. (2011). Analytical issues in monitoring drinking-water contamination related to short-term, heavy rainfall events. TrAC Trends in Analytical Chemistry, 30(8), 1243–1251.

Rollins, D. M., & Colwell, R. R. (1986). Viable but Nonculturable Stage of *Campylobacter-jejuni* and Its Role in Survival in the Natural Aquatic Environment. Applied and Environmental Microbiology, 52(3), 531–538.

Rose, J. B., Landeen, L. K., Riley, K. R., & Gerba, C. P. (1989). Evaluation of immunofluorescence techniques for detection of *Cryptosporidium* oocysts and *Giardia* cysts from environmental samples. Applied and environmental microbiology, 55(12), 3189-3196.

Rose, J. B., & Gerba, C. P. (1991). Assessing Potential Health Risks from Viruses and Parasites in Reclaimed Water in Arizona and Florida, USA. Water Science and Technology, 23(10-12), 2091–2098.

Rose, J. B. (1997). Environmental ecology of *Cryptosporidium* and public health implications. Annual Review of Public Health, 18, 135–161.

Rose, J. B., Huffman, D. E., & Gennaccaro, A. (2002). Risk and control of waterborne cryptosporidiosis. FEMS Microbiology Reviews, 26(2), 113-123.

Rosef, O., Gondrosen, B., Kapperud, G., & Underdal, B. (1983). Isolation and Characterization of *Campylobacter-jejuni* and *Campylobacter*-coli from Domestic and Wild Mammals in Norway. Applied and Environmental Microbiology, 46(4), 855–859.

Rosen, M. R., Hong, Y. S., Sheppard, D., Roberts, K., Viljevac, Z., Smaill, A., & Reeves, R. R. (2000). Final report for the IAEA urban aquifers RCA: Determining the effects of storm water infiltration on groundwater quality in an urban fractured rock aquifer, Auckland, New Zealand.

Ross, C. M., & Donnison, A. M. (2006). *Campylobacter jejuni* inactivation in New Zealand soils. Journal of Applied Microbiology, 101(5), 1188–1197.

Ruiz, A., Foronda, P., González, J. F., Guedes, A., Abreu-Acosta, N., Molina, J. M., & Valladares, B. (2008). Occurrence and genotype characterization of *Giardia duodenalis* in goat kids from the Canary Islands, Spain. Veterinary parasitology, 154(1), 137-141.

Ryan, U., Read, C., Hawkins, P., Warnecke, M., Swanson, P., Griffith, M., ... & Cox, P. (2005). Genotypes of *Cryptosporidium* from Sydney water catchment areas. Journal of Applied Microbiology, 98(5), 1221-1229.

Ryan, U., & Power, M. (2012). *Cryptosporidium* species in Australian wildlife and domestic animals. Parasitology, 139(13), 1673-1688.

Ryan, U., & Caccio, S. M. (2013). Zoonotic potential of *Giardia*. International Journal for Parasitology, 43(12-13), 943–956.

Ryan, U., & Hijjawi, N. (2015). New developments in *Cryptosporidium* research. International Journal for Parasitology, 45(6), 367–373.

Rzeżutka, A., & Kaupke, A. (2013). Occurrence and molecular identification of *Cryptosporidium* species isolated from cattle in Poland. Veterinary parasitology, 196(3), 301-306.

Sahin, O., Plummer, P., Sulaj, K., Pereira, S., Robbe-Austerman, S., Jordan, D., … Zhang, Q. (2007). An emerging strain of *Campylobacter jejuni* responsible for ovine abortion in the United States. Zoonoses and Public Health, 54, 80.

Sails, A. D., Swaminathan, B., & Fields, P. I. (2003). Utility of multilocus sequence typing as an epidemiological tool for investigation of outbreaks of gastroenteritis caused by *Campylobacter jejuni*. Journal of Clinical Microbiology, 41(10), 4733–4739.

Samuel, M. C., Vugia, D. J., Shallow, S., Marcus, R., Segler, S., McGivern, T., Kassenborg, H., Reilly, K., Kennedy, M., Angulo, F. (2004). Epidemiology of sporadic *Campylobacter* infection in the United States and declining trend in incidence, FoodNet 1996-1999. Clinical Infectious Diseases, 38, S165–S174.

Sanapareddy, N., Hamp, T. J., Gonzalez, L. C., Hilger, H. A., Fodor, A. A., & Clinton, S. M. (2009). Molecular diversity of a North Carolina wastewater treatment plant as revealed by pyrosequencing. Applied and Environmental Microbiology, 75(6), 1688–1696.

Sanders, S. Q., Boothe, D. H., Frank, J. F., & Arnold, J. W. (2007). Culture and detection of *Campylobacter jejuni* within mixed microbial populations of biofilms on stainless steel. Journal of Food Protection, 70(6), 1379–1385.

Sanford, S. E., & Josephson, G. K. A. (1982). Bovine Cryptosporidiosis - Clinical and Pathological Findings in 42 Scouring Neonatal Calves. Canadian Veterinary Journal-Revue Veterinaire Canadienne, 23(12), 343–347.

Sanger, F., Donelson, J. E., Coulson, A. R., Kössel, H. & Fischer, D. (1973). Use Of Dna Polymerase I Primed By A Synthetic Oligonucleotide To Determine A Nucleotide Sequence In Phage f1 DNA. Proceedings of the National Academy of Sciences, 70, 1209-1213.

Sanger, F., Nicklen, S. & Coulson, A. R. (1977). DNA Sequencing with Chain-Terminating Inhibitors. Proceedings of the National Academy of Sciences of the United States of America, 74, 5463-5467.

Santin, M., Trout, J. M., Xiao, L. H., Zhou, L., Greiner, E., & Fayer, R. (2004). Prevalence and age-related variation of *Cryptosporidium* species and genotypes in dairy calves. Veterinary Parasitology, 122(2), 103–117.

Santín, M., Trout, J. M., & Fayer, R. (2007). Prevalence and molecular characterization of *Cryptosporidium* and *Giardia* species and genotypes in sheep in Maryland. Veterinary parasitology, 146(1), 17-24.

Santin, M., Trout, J. M., & Fayer, R. (2008). A longitudinal study of cryptosporidiosis in dairy cattle from birth to 2 years of age. Veterinary Parasitology, 155(1-2), 15–23.

Santin, M. (2013). Clinical and subclinical infections with *Cryptosporidium* in animals. New Zealand Veterinary Journal, 61(1), 1–10.

Santin, M., & Fayer, R. (2015). Enterocytozoon bieneusi, *Giardia*, and *Cryptosporidium* Infecting White-tailed Deer. Journal of Eukaryotic Microbiology, 62(1), 34-43.

Sari, B., Arslan, M. O., Gicik, Y., Kara, M., & Tasci, G. T. (2009). The prevalence of *Cryptosporidium* species in diarrhoeic lambs in Kars province and potential risk factors. Tropical Animal Health and Production, 41(5), 819–826.

Sartory, D. P., Parton, A., Parton, A. C., Roberts, J., & Bergmann, K. (1998). Recovery of *Cryptosporidium* oocysts from small and large volume water samples using a compressed foam filter system. Letters in Applied Microbiology, 27(6), 318–322.

Sasakawa, C. (2009). Molecular mechanisms of bacterial infection via the gut (Vol. 337). Springer.

Sauch, J. F. (1985). Use of immunofluorescence and phase-contrast microscopy for detection and identification of *Giardia* cysts in water samples. Applied and environmental microbiology, 50(6), 1434-1438.

Savill, M. G., N, J. A., Ball, A., Klena, J. D., Scholes, P., Whyte, R. J., McCormick, R. E.& Jankovic, D. (2001). Enumeration of *Campylobacter* in New Zealand recreational and drinking waters. Journal of Applied Microbiology, 91(1), 38-46.

Schets, F. M., Van Wijnen, J. H., Schijven, J. F., Schoon, H., & de Roda Husman, A. M. (2008). Monitoring of waterborne pathogens in surface waters in Amsterdam, The Netherlands, and the potential health risk associated with exposure to *Cryptosporidium* and *Giardia* in these waters. Applied and environmental microbiology, 74(7), 2069-2078.

Scheutz, F., Nielsen, E. M., Frimodt-Moller, J., Boisen, N., Morabito, S., Tozzoli, R., ... & Caprioli, A. (2011). Characteristics of the enteroaggregative Shiga toxin/verotoxin-producing *Escherichia coli* O104: H4 strain causing the outbreak of haemolytic uraemic syndrome in Germany, May to June 2011. Euro Surveill, 16(24), 19889.

Schmutz, J., Wheeler, J., Grimwood, J., Dickson, M., Yang, J., Caoile, C., Bajorek, E., Black, S., Chan, Y. M., Denys, M., Escobar, J., Flowers, D., Fotopulos, D., Garcia, C., Gomez, M., Gonzales, E., Haydu, L., Lopez, F., Ramirez, L., Retterer, J., Rodriguez, A., Rogers, S., Salazar, A., Tsai, M. & Myers, R. M. (2004). Quality assessment of the human genome sequence. Nature, 429, 365-368.

Schnyder, M., Kohler, L., Hemphill, A., & Deplazes, P. (2009). Prophylactic and therapeutic efficacy of nitazoxanide against *Cryptosporidium parvum* in experimentally challenged neonatal calves. Veterinary Parasitology, 160(1-2), 149–154.

Schönberg-Norio, D., Takkinen, J., Hänninen, M. L., Katila, M. L., Kaukoranta, S. S., Mattila, L., Rautelin, H. (2004). Swimming and *Campylobacter* infections. Emerg Infect Dis, 10, 1474-1477.

Scott, C. A., Smith, H. V., Mtambo, M. M. A., & Gibbs, H. A. (1995). An epidemiological study of *Cryptosporidium parvum* in two herds of adult beef cattle. Veterinary Parasitology, 57(4), 277-288.

Scott, W. G., Scott, H. M., Lake, R. J., Baker, M. G. (2000). Economic cost to New Zealand of foodborne infectious disease. The New Zealand Medical Journal 113:281-284.

Sears, A., Baker, M. G., Wilson, N., Marshall, J., Mullner, P., Campbell, D. M., Lake, R. J.& French, N. P. (2011). Marked campylobacteriosis decline after interventions aimed at poultry, New Zealand. Emerging Infectious Diseases, 17(6), 1007-1015.

Sebald, M., & Veron, M. (1963, November). Base DNA content and classification of vibrios. In Annales de l'Institut Pasteur (Vol. 105, p. 897).

Selander, R. K., Caugant, D. A., Ochman, H., Musser, J. M., Gilmour, M. N., & Whittam, T. S. (1986). Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. Applied and Environmental Microbiology, 51(5), 873.

Sevinç, F., ŞİMŞEK, A., & Uslu, U. (2006). Massive *Cryptosporidium parvum* infection associated with an outbreak of diarrhoea in neonatal goat kids. Turkish Journal of Veterinary and Animal Sciences, 29(6), 1317-1320.

Shanks, O. C., McLellan, S., Huse, S. M., & Sogin, M. L. (2011). Characterization of microbial community structures in recreational waters and primary sources of faecal pollution with a next-generation sequencing approach. Environ Microbiol: Curr Technol Water Appl, 203–223.

Shepherd, K. M., & WynJones, A. P. (1996). An evaluation of methods for the simultaneous detection of *Cryptosporidium* oocysts and *Giardia* cysts from water. Applied and Environmental Microbiology, 62(4), 1317–1322.

Shikani, H., & Weiss, L. M. (2014). Human cryptosporidiosis: a clinical perspective. In *Cryptosporidium*: parasite and disease (pp. 383-421). Springer Vienna.

Shirley, D. A. T., Moonah, S. N., & Kotloff, K. L. (2012). Burden of disease from cryptosporidiosis. Current Opinion in Infectious Diseases, 25(5), 555–563.

Shokralla, S., Spall, J. L., Gibson, J. F., & Hajibabaei, M. (2012). Next-generation sequencing technologies for environmental DNA research. Molecular Ecology, 21(8), 1794–805.

Shrestha, R. D., Grinberg, A., Dukkipati, V. S. R., Pleydell, E. J., Prattley, D. J., & French, N. P. (2014). Infections with multiple *Cryptosporidium* species and new genetic variants in young dairy calves on a farm located within a drinking water catchment area in New Zealand. Veterinary Parasitology, 202(3-4), 287–291.

Sikkema-Raddatz, B., Johansson, L. F., De Boer, E. N., Almomani, R., Boven, L. G., Van Den Berg, M. P., Van Spaendonck-Zwarts, K. Y., Van Tintelen, J. P., Sijmons, R. H., Jongbloed, J. D. H. & Sinke, R. J. 2013. Targeted Next-Generation Sequencing can Replace Sanger Sequencing in Clinical Diagnostics. Human Mutation, 34, 1035-1042.

Silva, J., Leite, D., Fernandes, M., Mena, C., Gibbs, P. A., & Teixeira, P. (2011). *Campylobacter* spp. as a foodborne pathogen: a review. Frontiers in Microbiology, 2.

Silva, J., & Teixeira, P. (2015). Tackling *Campylobacter*: a Review. American Journal of Advanced Food Science and Technology, 3(2), 107-124.

Silverlås, C., Näslund, K., Björkman, C., & Mattsson, J. G. (2010). Molecular characterisation of *Cryptosporidium* isolates from Swedish dairy cattle in relation to age, diarrhoea and region. Veterinary parasitology, 169(3), 289-295.

Silverlas, C., & Blanco-Penedo, I. (2013). *Cryptosporidium* spp. in calves and cows from organic and conventional dairy herds. Epidemiology and Infection, 141(3), 529–539.

Simmons, G., Hope, V., Lewis, G., Whitmore, J., & Gao, W. (2001). Contamination of potable roof-collected rainwater in Auckland, New Zealand. Water Research, 35(6), 1518–1524.

Simon, C. & Daniel, R. 2011. Metagenomic analyses: past and future trends. Applied and Environmental Microbiology, 77, 1153-1161.

Singh, P., Foley, S. L., Nayak, R., & Kwon, Y. M. (2012). Multilocus sequence typing of Salmonella strains by high-throughput sequencing of selectively amplified target genes. Journal of microbiological methods, 88(1), 127-133.

Sinton, L., Hall, C., & Braithwaite, R. (2007). Sunlight inactivation of *Campylobacter jejuni* and Salmonella enterica, compared with *Escherichia coli*, in seawater and river water. Journal of Water and Health, 5(3), 357–365.

Sischo, W. M., Atwill, E. R., Lanyon, L. E., & George, J. (2000). Cryptosporidia on dairy farms and the role these farms may have in contaminating surface water supplies in the north-eastern United States. Preventive Veterinary Medicine, 43(4), 253–267.

Slapeta, J. (2012). Letter: The name *Cryptosporidium tyzzeri* Ren, Zhao, Zhang, Ning, Jian, Wang, Lv, Wang, Arrowood and Xiao, 2012 is permanently invalid. Experimental Parasitology, 130(3), 306–307.

Slapeta, J. (2013). Cryptosporidiosis and *Cryptosporidium* species in animals and humans: A thirty colour rainbow? International Journal for Parasitology, 43(12-13), 957–970.

Slavin, D., 1955. *Cryptosporidium* meleagridis (sp. nov.). J. Comp. Pathol. 65, 262±6.

Smith, J. L. (1995). Arthritis, Guillain-Barré syndrome, and other sequelae of *Campylobacter jejuni* enteritis. Journal of Food Protection®, 58(10), 1153-1170.

Smith, H. V., Campbell, B. M.& Grimason, A. M. (2003). Comments on research note. Water Research, 37(10), 2525-2527.

Smith, H. V, Nichols, R. A. B., Mallon, M., Macleod, A., Tait, A., Reilly, W. J., Browning, L. M., Gray, D., Reid, S. W. J., Wastling, J. M. (2005). Natural *Cryptosporidium hominis* infections in Scottish cattle. Veterinary Record, 156(22), 710–711.

Smith, H. V, Caccio, S. M., Tait, A., McLauchlin, J., & Thompson, R. C. A. (2006). Tools for investigating the environmental transmission of *Cryptosporidium* and *Giardia* infections in humans. Trends in Parasitology, 22(4), 160–167.

Smith, H. V, & Nichols, R. A. B. (2010). *Cryptosporidium*: Detection in water and food. Experimental Parasitology, 124(1), 61–79.

Snel, S. J., Baker, M. G., & Venugopal, K. (2009a). The epidemiology of cryptosporidiosis in New Zealand, 1997-2006. The New Zealand Medical Journal (Online), 122(1290).

Snel, S. J., Baker, M. G., & Venugopal, K. (2009b). The epidemiology of giardiasis in New Zealand, 1997-2006. The New Zealand Medical Journal (Online), 122(1290).

Snel, S. J., Baker, M. G., Kamalesh, V., French, N., & Learmonth, J. (2009). A tale of two parasites: the comparative epidemiology of cryptosporidiosis and giardiasis. Epidemiology and infection, 137(11), 1641-1650.

Snelling, W. J., Moore, J. E., McKenna, J. P., Lecky, D. M., & Dooley, J. S. G. (2006). Bacterial-protozoa interactions; an update on the role these phenomena play towards human illness. Microbes and Infection, 8(2), 578–587.

Soave, R., Danner, R. L., Honig, C. L., Ma, P., Hart, C. C., Nash, T., & Roberts, R. B. (1984). Cryptosporidiosis in Homosexual Men. Annals of Internal Medicine, 100(4), 504–511.

Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., ... & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere". Proceedings of the National Academy of Sciences, 103(32), 12115-12120.

Soltane, R., Guyot, K., Dei-Cas, E., & Ayadi, A. (2007). Prevalence of *Cryptosporidium* spp. (Eucoccidiorida : Cryptosporiidae) in seven species of farm animals in Tunisia. Parasite-Journal De La Societe Francaise De Parasitologie, 14(4), 335–338.

Spencer, E., List, R., More, F., This, L., Fisheries, U. S., Aquaculture, C., … Health, A. (2011). Environmental and public health implications of wastewater quality. African Journal of Biotechnology, 10(13), 2379–2387.

Spencer, S. E., Marshall, J., Pirie, R., Campbell, D., Baker, M. G., & French, N. P. (2012). The spatial and temporal determinants of campylobacteriosis notifications in New Zealand, 2001–2007. Epidemiology and infection, 140(09), 1663-1677.

Sprong, H., Caccio, S. M., van der Giessen, J. W. B., & Partners, Z. N. &. (2009). Identification of Zoonotic Genotypes of *Giardia duodenalis*. Plos Neglected Tropical Diseases, 3(12).

Sproston, E. L., Ogden, I. D., MacRae, M., Dallas, J. F., Sheppard, S. K., Cody, A. J., … Strachan, N. J. C. (2011). Temporal variation and host association in the *Campylobacter* population in a longitudinal ruminant farm study. Applied and Environmental Microbiology, 77(18), 6579–86.

St Jean, G., Y. Couture, P. Dubreuil, and J. L. Frechette. "Diagnosis of *Giardia* infection in 14 calves." Journal of the American Veterinary Medical Association 191, no. 7 (1987): 831-832.

Staley, C., Unno, T., Gould, T. J., Jarvis, B., Phillips, J., Cotner, J. B. and Sadowsky, M. J. (2013), Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. Journal of Applied Microbiology, 115: 1147–1158.

Stanley, K. N., Wallace, J. S., Currie, J. E., Diggle, P. J., & Jones, K. (1998). Seasonal variation of thermophilic Campylobacters in lambs at slaughter. Journal of Applied Microbiology, 84(6), 1111–1116.

Stanley, K., & Jones, K. (2003). Cattle and sheep farms as reservoirs of *Campylobacter*. J Appl Microbiol 94:104s-113s.

Stark, D., Barratt, J. L. N., van Hal, S., Marriott, D., Harkness, J., & Ellis, J. T. (2009). Clinical Significance of Enteric Protozoa in the Immunosuppressed Human Population. Clinical Microbiology Reviews, 22(4), 634.

Starkey, S. R., Wade, S. E., Schaaf, S., & Mohammed, H. O. (2005). Incidence of *Cryptosporidium parvum* in the dairy cattle population in a New York City Watershed. Veterinary Parasitology, 131(3-4), 197–205.

Steele, M., McNab, B., Fruhner, L., DeGrandis, S., Woodward, D., & Odumeru, J. A. (1998). Epidemiological typing of *Campylobacter* isolates from meat processing plants by

pulsed-field gel electrophoresis, fatty acid profile typing, serotyping, and biotyping. Applied and Environmental Microbiology, 64(7), 2346–2349.

Stelzer, E. A., Strickler, K. M., & Schill, W. B. (2012). Interlaboratory comparison of three microbial source tracking quantitative polymerase chain reaction (qPCR) assays from fecal-source and environmental samples (No. 2012-5087). US Geological Survey.

Sterling, C. R., & Arrowood, M. J. (1986). Detection of *Cryptosporidium* Sp Infections Using a Direct Immunofluorescent Assay. Pediatric Infectious Disease Journal, 5(1), S139–S142.

Stevenson, M., Nunes, T., Sanchez, J., & Thornton, R. (2013). epiR: functions for analysing epidemiological data, R package version 0.9-32. 2011.

Stone, D., Davis, M., Baker, K., Besser, T., Roopnarine, R., & Sharma, R. (2013). MLST Genotypes and Antibiotic Resistance of *Campylobacter* spp. Isolated from Poultry in Grenada. Biomed Research International.

Stott, R., Davies-Colley, R., Nagels, J., Donnison, A., Ross, C., & Muirhead, R. (2011). Differential behaviour of *Escherichia* coil and *Campylobacter* spp. in a stream draining dairy pasture. Journal of Water and Health, 9(1), 59–69

Strachan, N. J. C., Gormley, F. J., Rotariu, O., Ogden, I. D., Miller, G., Dunn, G. M., … Forbes, K. J. (2009). Attribution of *Campylobacter* Infections in Northeast Scotland to Specific Sources by Use of Multilocus Sequence Typing. Journal of Infectious Diseases, 199(8), 1205–1208.

Stucki, U., Frey, J., Nicolet, J., & Burnens, A. P. (1995). Identification of *Campylobacter-jejuni* on the Basis of a Species-Specific Gene That Encodes a Membrane-Protein. Journal of Clinical Microbiology, 33(4), 855–859.

Sturdee, A. P., Bodley-Tickell, A. T., Archer, A., & Chalmers, R. M. (2003). Long-term study of *Cryptosporidium* prevalence on a lowland farm in the United Kingdom. Veterinary Parasitology, 116(2), 97–113.

Su, C., Lei, L., Duan, Y., Zhang, K. Q., & Yang, J. (2012). Culture-independent methods for studying environmental microorganisms: methods, application, and perspective. Applied microbiology and biotechnology, 93(3), 993-1003.

Sulaiman, I. M., Morgan, U. M., Thompson, R. C. A., Lal, A. A., & Xiao, L. H. (2000). Phylogenetic relationships of *Cryptosporidium* parasites based on the 70-kilodalton heat shock protein (HSP70) gene. Applied and Environmental Microbiology, 66(6), 2385–2391.

Sulaiman, I. M., Lal, A. A., & Xiao, L. H. (2001). A population genetic study of the *Cryptosporidium parvum* human genotype parasites. Journal of Eukaryotic Microbiology, 24s–27s.

Sulaiman, I. M., Lal, A. A., & Xiao, L. H. (2002). Molecular phylogeny and evolutionary relationships of *Cryptosporidium* parasites at the actin locus. Journal of Parasitology, 88(2), 388–394.

Suresh, K. G., Toranzos, G. A., Fayer, R., Nissaparton, V., Olveda, R., Ashbolt, N., & Gannon, V. (2012). Assessing the importance of zoonotic waterborne pathogens. Animal Waste, Water Quality and Human Health, 17-72.

Sweeny, J. P. A., Ryan, U. M., Robertson, I. D., & Jacobson, C. (2011). *Cryptosporidium* and *Giardia* associated with reduced lamb carcase productivity. Veterinary Parasitology, 182(2-4), 127–139.

Tam, C. C., Rodrigues, L. C., Viviani, L., Dodds, J. P., Evans, M. R., Hunter, P. R., ... & O'Brien, S. J. (2012). Longitudinal study of infectious intestinal disease in the UK (IID2 study): incidence in the community and presenting to general practice. Gut, 61(1), 69-77.

Tan, B., Ng, C., Nshimyimana, J. P., Loh, L. L., Gin, K. Y.-H., & Thompson, J. R. (2015). Next-generation sequencing (NGS) for assessment of microbial water quality: current progress, challenges, and future opportunities. Frontiers in Microbiology, 6.

Tanriverdi, S., Arslan, M. O., Akiyoshi, D. E., Tzipori, S., & Widmer, G. (2003). Identification of genotypically mixed *Cryptosporidium parvum* populations in humans and calves. Molecular and Biochemical Parasitology, 130(1), 13–22.

Tanriverdi, S., & Widmer, G. (2006). Differential evolution of repetitive sequences in *Cryptosporidium parvum* and *Cryptosporidium hominis*. Infection Genetics and Evolution, 6(2), 113–122.

Tanriverdi, S., Grinberg, A., Chalmers, R. M., Hunter, P. R., Petrovic, Z., Akiyoshi, D. E., … Widmer, G. (2008). Inferences about the Global Population Structures of *Cryptosporidium parvum* and *Cryptosporidium hominis*. Applied and Environmental Microbiology, 74(23), 7227–7234.

Teh, K. H., Flint, S., & French, N. (2010). Biofilm formation by *Campylobacter jejuni* in controlled mixed-microbial populations. International Journal of Food Microbiology, 143(3), 118–124.

Tholozan, J. L., Cappelier, J. M., Tissier, J. P., Delattre, G., & Federighi, M. (1999). Physiological characterization of viable-but-nonculturable *Campylobacter jejuni* cells. Applied and Environmental Microbiology, 65(3), 1110–1116.

Thomas, C., Gibson, H., Hill, D. J., & Mabey, M. (1999). *Campylobacter* epidemiology: an aquatic perspective. Journal of Applied Microbiology, 85, 168s–177s.

Thomas, T., Gilbert, J. & Meyer, F. (2012). Metagenomics-a guide from sampling to data analysis. Microb Inform Exp, 2.

Thompson, H. P., Dooley, J. S. G., Kenny, J., McCoy, M., Lowery, C. J., Moore, J. E., & Xiao, L. (2007). Genotypes and subtypes of *Cryptosporidium* spp. in neonatal calves in Northern Ireland. Parasitology Research, 100(3), 619–624.

Thompson, R. C. A., & Monis, P. (2012). *Giardia*- From Genome to Proteome. Advances in Parasitology, Vol 78, 78, 57–95.

Thomsen, P. F., & Willerslev, E. (2015). Environmental DNA–an emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation, 183, 4-18.

Thurston-Enriquez, J. A., Gilley, J. E., & Eghball, B. (2005). Microbial quality of runoff following land application of cattle manure and swine slurry. Journal of water and health, 3(2), 157-171.

Till, D., McBride, G., Ball, A., Taylor, K., & Pyle, E. (2008). Large-scale freshwater microbiological study: rationale, results and risks. Journal of Water and Health, 6(4), 443–460.

Topp, E., Scott, A., Lapen, D. R., Lyautey, E., & Duriez, P. (2009). Livestock waste treatment systems for reducing environmental exposure to hazardous enteric pathogens: some considerations. Bioresource technology, 100(22), 5395-5398.

Tran, C. P., Bode, R. W., Smith, A. J., & Kleppel, G. S. (2010). Land-use proximity as a basis for assessing stream water quality in New York State (USA). Ecological Indicators, 10(3), 727–733.

Traub, R., Wade, S., Read, C., Thompson, A., & Mohammed, H. (2005). Molecular characterization of potentially zoonotic isolates of *Giardia duodenalis* in horses. Veterinary Parasitology, 130(3-4), 317–321.

Tringe, S. G., & Rubin, E. M. (2005). Metagenomics: DNA sequencing of environmental samples. Nature Reviews Genetics, 6(11), 805–814.

Trotz-Williams, L. A., Jarvie, B. D., Peregrine, A. S., Duffield, T. F., & Leslie, K. E. (2011). Efficacy of halofuginone lactate in the prevention of cryptosporidiosis in dairy calves. Veterinary Record, 168(19).

Trout, J. M., Santín, M., Greiner, E., & Fayer, R. (2004). Prevalence of *Giardia duodenalis* genotypes in pre-weaned dairy calves. Veterinary parasitology, 124(3), 179-186.

Trout, J. M., Santín, M., Greiner, E., & Fayer, R. (2005). Prevalence and genotypes of *Giardia duodenalis* in post-weaned dairy calves. Veterinary parasitology, 130(3), 177-183.

Trout, J. M., Santín, M., Greiner, E. C., & Fayer, R. (2006). Prevalence and genotypes of *Giardia duodenalis* in 1–2 year old dairy cattle. Veterinary parasitology, 140(3), 217-222.

Trout, J. M., Santín, M., & Fayer, R. (2007). Prevalence of *Giardia duodenalis* genotypes in adult dairy cows. Veterinary parasitology, 147(3), 205-209.

Tzanidakis, N., Sotiraki, S., Claerebout, E., Ehsan, A., Voutzourakis, N., Kostopoulou, D., ... & Geurden, T. (2014). Occurrence and molecular characterization of *Giardia duodenalis* and *Cryptosporidium* spp. in sheep and goats reared under dairy husbandry systems in Greece. Parasite, 21, 45.

Tzipori, S., Angus, K. W., Campbell, I., & Gray, E. W. (1980). *Cryptosporidium* - Evidence for a Single-Species Genus. Infection and Immunity, 30(3), 884–886.

Tzipori, S., Angus, K. W., Campbell, I., & Sherwood, D. (1981). Diarrhea in young red deer associated with infection with *Cryptosporidium*. Journal of Infectious Diseases, 144(2), 170-175.

Tzipori, S., Smith, M., Birch, C., Barnes, G., & Bishop, R. (1983a). Cryptosporidiosis in Hospital Patients with Gastroenteritis. American Journal of Tropical Medicine and Hygiene, 32(5), 931–934.

Tzipori, S., Smith, M., Halpin, C., Angus, K. W., Sherwood, D., & Campbell, I. (1983b). Experimental Cryptosporidiosis in Calves - Clinical Manifestations and Pathological Findings. Veterinary Record, 112(6), 116–120.

Tzipori, S., & Ward, H. (2002). Cryptosporidiosis: biology, pathogenesis and disease. Microbes and Infection, 4(10), 1047-1058.

Uaboi-Egbenni, P. O., Bessong, P. O., Samie, A., & Obi, C. L. (2011). Prevalence and antimicrobial susceptibility profiles of *Campylobacter jejuni* and *coli* isolated from diarrheic and non-diarrheic goat faeces in Venda region, South Africa. African Journal of Biotechnology, 10(64), 14116–14124.

Uehlinger, F. D., Greenwood, S. J., O'Handley, R., McClure, J. T., Coklin, T., Dixon, B. R., De Boer, M., Zwiers, H.& Barkema, H. W. (2011). Prevalence and genotypes of *Giardia duodenalis* in dairy and beef cattle in farms around Charlottetown, Prince Edward Island, Canada. Canadian Veterinary Journal, 52(9), 967-972.

Uga, S., Matsuo, J., Kono, E., Kimura, K., Inoue, M., Rai, S. K., & Ono, K. (2000). Prevalence of *Cryptosporidium parvum* infection and pattern of oocyst shedding in calves in Japan. Veterinary Parasitology, 94(1-2), 27–32.

USEPA. (2005). Method 1623: *Cryptosporidium* and *Giardia* in Water by Filtration/IMS/FA (PDF) December 2005 update (EPA 815-R-05-002). Office of Water (4607) Washington DC: U.S. Environmental Protection Agency.

USEPA. (2013). Method 1623.1: *Cryptosporidium* and *Giardia* in Water by Filtration/IMS/FA (PDF) December 2012. Office of Water (4607) Washington DC: U.S. Environmental Protection Agency.

Van Donkersgoed, J., Janzen, E. D., Chirino-Trejo, M., Berry, C., Clark, E. G., & Haines, D. M. (1990). *Campylobacter jejuni* abortions in two beef cattle herds in Saskatchewan. The Canadian Veterinary Journal, 31(5), 373.

Van Dyke, M. I., Morton, V. K., McLellan, N. L., & Huck, P. M. (2010). The occurrence of *Campylobacter* in river water and waterfowl within a watershed in southern Ontario, Canada. Journal of Applied Microbiology, 109, 1053–1066.

Venter, J. C., Remington, K., Heidelberg, J. F., Halpern, A. L., Rusch, D., Eisen, J. A., … Smith, H. O. (2004). Environmental Genome Shotgun Sequencing of the Sargasso Sea. Science , 304 (5667), 66–74.

Ventura, G., Cauda, R., Larocca, L. M., Riccioni, M. E., Tumbarello, M., & Lucia, M. B. (1997). Gastric cryptosporidiosis complicating HIV infection: Case report and review of the literature. European Journal of Gastroenterology & Hepatology, 9(3), 307–310.

Vereen, E., Lowrance, R. R., Cole, D. J., & Lipp, E. K. (2007). Distribution and ecology of campylobacters in coastal plain streams (Georgia, United States of America). Applied and environmental microbiology, 73(5), 1395-1403.

Wagenaar, J. A., van Bergen, M. A. P., Blaser, M. J., Tauxe, R. V, Newell, D. G., & van Putten, J. P. M. (2014). *Campylobacter* fetus Infections in Humans: Exposure and Disease. Clinical Infectious Diseases, 58(11), 1579–1586.

Walker, C. L. F., Rudan, I., Liu, L., Nair, H., Theodoratou, E., Bhutta, Z. A., ... & Black, R. E. (2013). Global burden of childhood pneumonia and diarrhoea. The Lancet, 381(9875), 1405-1416.

Wang, Y., Feng, Y., Cui, B., Jian, F., Ning, C., Wang, R., ... & Xiao, L. (2010). Cervine genotype is the major *Cryptosporidium* genotype in sheep in China. Parasitology research, 106(2), 341-347.

Wang, R. J., Wang, H. L., Sun, Y. R., Zhang, L. X., Jian, F. C., Qi, M., Ning, C., Xiao, L. H. (2011). Characteristics of *Cryptosporidium* Transmission in Preweaned Dairy Cattle in Henan, China. Journal of Clinical Microbiology, 49(3), 1077–1082.

Wang, J., McLenachan, P. A., Biggs, P. J., Winder, L. H., Schoenfeld, B. I. K., Narayan, V. V, … Lockhart, P. J. (2013). Environmental bio-monitoring with high-throughput sequencing. Briefings in Bioinformatics , 14 (5), 575–588.

Wang, R., Li, G., Cui, B., Huang, J., Cui, Z., Zhang, S., ... & Wang, M. (2014). Prevalence, molecular characterization and zoonotic potential of *Cryptosporidium* spp. in goats in Henan and Chongqing, China. Experimental parasitology, 142, 11-16.

Watkins, R. R., & Eckmann, L. (2014). Treatment of giardiasis: current status and future directions. Current infectious disease reports, 16(2), 1-8.

Weber, R., Bryan, R. T., & Juranek, D. D. (1992). Improved Stool Concentration Procedure for Detection of *Cryptosporidium* Oocysts in Fecal Specimens. Journal of Clinical Microbiology, 30(11), 2869–2873.

Webster, K. A., Smith, H. V, Giles, M., Dawson, L., & Robertson, L. J. (1996). Detection of *Cryptosporidium parvum* oocysts in faeces: Comparison of conventional coproscopical methods and the polymerase chain reaction. Veterinary Parasitology, 61(1-2), 5–13.

Weijtens, M. J. B. M., Reinders, R. D., Urlings, H. A. P., & Van der Plas, J. (1999). *Campylobacter* infections in fattening pigs; excretion pattern and genetic diversity. Journal of Applied Microbiology, 86(1), 63–70.

Weisburger, W. R., Hutcheon, D. F., Yardley, J. H., Roche, J. C., Hillis, W. D., & Charache, P. (1979). Cryptosporidiosis in an immunosuppressed renal-transplant recipient with IgA deficiency. American journal of clinical pathology, 72(3), 473-478.

Wells, S. J., Dee, S., & Godden, S. (2002). Biosecurity for gastrointestinal diseases of adult dairy cattle. Veterinary Clinics of North America: Food Animal Practice, 18(1), 35-55.

Wells, B., Shaw, H., Hotchkiss, E., Gilray, J., Ayton, R., Green, J., ... & Innes, E. (2015). Prevalence, species identification and genotyping *Cryptosporidium* from livestock and deer in a catchment in the Cairngorms with a history of a contaminated public water supply. Parasit Vectors, 8(1), 66.

Wesley, I. V, Wells, S. J., Harmon, K. M., Green, A., Schroeder-Tucker, L., Glover, M., & Siddique, I. (2000). Fecal shedding of *Campylobacter* and Arcobacter spp. in dairy cattle. Applied and Environmental Microbiology, 66(5), 1994–2000.

WHO. (2003). Guidelines for safe recreational water environments: Coastal and fresh waters (Vol. 1). World Health Organization.

WHO. (2004). Waterborne Zoonoses: Identification, Causes and Control. In J. A. Cotruvo, A. Dufour, G. Rees, J. Bartram, R. Carr, D. O. Cliver, … V. Gannon (Eds.), . London, UK: IWA Publishing.

WHO. (2011). Guidelines for drinking-water quality - 4th ed. World Health Organization. ISBN 978 92 4 1548151.

WHO. (2012). "Progress on Drinking-Water and Sanitation–2012 Update" launched on 6 March 2012.

Wicki, M., Svoboda, P., & Tanner, M. (2009). Occurrence of *Giardia lamblia* in recreational streams in Basel-Landschaft, Switzerland. Environmental research, 109(5), 524-527.

Widmer, G. (2004). Population genetics of *Cryptosporidium parvum*. Trends in Parasitology, 20(1), 3–6.

Wielinga, C., Ryan, U., Thompson, R. C. A., & Monis, P. (2011). Multi-locus analysis of *Giardia duodenalis* intra-Assemblage B substitution patterns in cloned culture isolates suggests sub-Assemblage B analyses will require multi-locus genotyping with conserved and variable genes. International Journal for Parasitology, 41(5), 495–503.

Wilcock, R. J. (2006). Assessing the relative importance of faecal pollution sources in rural catchments. Environment Waikato.

Wilkes, G., Edge, T., Gannon, V., Jokinen, C., Lyautey, E., Medeiros, D., Neumann, N., Ruecker, N., Topp, E. & Lapen, D. R. (2009). Seasonal relationships among indicator bacteria, pathogenic bacteria, *Cryptosporidium* oocysts, *Giardia* cysts, and hydrological indices for surface waters within an agricultural landscape. Water Research, 43(8), 2209–2223.

Wilson D. J., Gabriel E., Leatherbarrow A. J. H., Cheesbrough J., Gee S., Bolton E., Fox A., Fearnhead P., Hart C. A., Dinggle P. J. (2008). Tracing the source of campylobacteriosis. Plos Genetics 4.

Wilson N. (2005). A Systematic Review of the Aetiology of Human Campylobacteriosis in New Zealand: Report to the Food Safety Authority of New Zealand.

Winkworth, C. L., Matthaei, C. D., Townsend, C. R. (2008). Prevalence of *Giardia* and *Cryptosporidium* spp in calves from a region in New Zealand experiencing intensification of dairying. New Zeal Vet J 56, 15-20.

Wong T., On S., & Michie H. (2006). *Campylobacter* in New Zealand: reservoirs, sources and the labyrinth of transmission routes. NZJEH 29:1-6.

Woodward L., O'Brien J., & Pearce B. (2005). *Campylobacter* and poultry. What we achieved-where we will go: 45.

Wu, J., Gao, W., Johnson, R. H., Zhang, W., & Meldrum, D. R. (2013). Integrated metagenomic and metatranscriptomic analyses of microbial communities in the meso-and bathypelagic realm of North Pacific Ocean. Marine drugs, 11(10), 3777-3801.

Xiao, L., & Herd, R. P. (1994). Infection patterns of *Cryptosporidium* and *Giardia* in calves. Veterinary Parasitology, 55(3), 257–262.

Xiao, L. H., Limor, J., Morgan, U. M., Sulaiman, I. M., Thompson, R. A. C., & Lal, A. A. (2000). Sequence differences in the diagnostic target region of the oocyst wall protein gene of *Cryptosporidium* parasites. Applied and Environmental Microbiology, 66(12), 5499–5502.

Xiao, L., Limor, J., Bern, C., Lal, A. A., & Epidemic Working Group. (2001). Tracking *Cryptosporidium parvum* by sequence analysis of small double-stranded RNA. Emerging infectious diseases, 7(1), 141.

Xiao, L. H., Bern, C., Limor, J., Sulaiman, I., Roberts, J., Checkley, W., … Lal, A. A. (2001a). Identification of 5 types of *Cryptosporidium* parasites in children in Lima, Peru. Journal of Infectious Diseases, 183(3), 492–497.

Xiao, L. H., Limor, J., Bern, C., & Lal, A. A. (2001b). Tracking *Cryptosporidium parvum* by sequence analysis of small double-stranded RNA. Emerging Infectious Diseases, 7(1), 141–145.

Xiao, L. H., Singh, A., Limor, J., Graczyk, T. K., Gradus, S., & Lal, A. (2001c). Molecular characterization of *Cryptosporidium* oocysts in samples of raw surface water and wastewater. Applied and Environmental Microbiology, 67(3), 1097–1101.

Xiao, L. H., Fayer, R., Ryan, U., & Upton, S. J. (2004). *Cryptosporidium* taxonomy: Recent advances and implications for public health. Clinical Microbiology Reviews, 17(1), 72.

Xiao, L.H., Zhou, L., Santin, M., Yang, W.L., Fayer, R. (2007). Distribution of *Cryptosporidium parvum* subtypes in calves in eastern United States. Parasitol Res 100, 701-706.

Xiao, L.H. (2010). Molecular epidemiology of cryptosporidiosis: An update. Exp Parasitol 124, 80-89.

Yamazaki-Matsune, W., Taguchi, M., Seto, K., Kawahara, R., Kawatsu, K., Kumeda, Y., … Tsukamoto, T. (2007). Development of a multiplex PCR assay for identification of *Campylobacter coli*, *Campylobacter fetus*, *Campylobacter hyointestinalis* subsp *hyointestinalis*, *Campylobacter jejuni*, *Campylobacter lari* and *Campylobacter upsaliensis*. Journal of Medical Microbiology, 56(11), 1467–1473.

Yason, J. A. D., & Rivera, W. L. (2007). Genotyping of *Giardia duodenalis* isolates among residents of slum area in Manila, Philippines. Parasitology research, 101(3), 681-687.

Yang, R. C., Jacobson, C., Gordon, C.& Ryan, U. (2009). Prevalence and molecular characterisation of *Cryptosporidium* and *Giardia* species in pre-weaned sheep in Australia. Veterinary Parasitology, 161(1-2), 19-24.

Yang, R. C., Jacobson, C., Gardner, G., Carmichael, I., Campbell, A. J. D., Ng-Hublin, J., & Ryan, U. (2014). Longitudinal prevalence, oocyst shedding and molecular characterisation of *Cryptosporidium* species in sheep across four states in Australia. Veterinary Parasitology, 200(1-2), 50–58.

Yang, R., Paparini, A., Monis, P., & Ryan, U. (2014). Comparison of next-generation droplet digital PCR (ddPCR) with quantitative PCR (qPCR) for enumeration of *Cryptosporidium* oocysts in faecal samples. International journal for parasitology, 44(14), 1105-1113.

Yang, R., Elankumaran, Y., Hijjawi, N., & Ryan, U. (2015). Validation of cell-free culture using scanning electron microscopy (SEM) and gene expression studies. Experimental parasitology, 153, 55-62.

Yoder, J. S., Beach, M. J., & Centers for Disease Control and Prevention (CDC). (2007). Giardiasis surveillance—United States, 2003–2005. MMWR Surveill Summ, 56(7), 11-18.

Yoder, Jonathan S., Ryan M. Wallace, Sarah A. Collier, Michael J. Beach, Michele C. Hlavsa, and Centers for Disease Control and Prevention (CDC). (2012). "Cryptosporidiosis surveillance—United States, 2009–2010." MMWR Surveill Summ, 61(5), 1-12.

Zhang, W., Zhang, X., Wang, R., Liu, A., Shen, Y., Ling, H., ... & Zhang, L. (2012). Genetic characterizations of *Giardia duodenalis* in sheep and goats in Heilongjiang Province, China and possibility of zoonotic transmission. PLoS Negl Trop Dis, 6(9), e1826.

Ziemer, C. J., Bonner, J. M., Cole, D., Vinjé, J., Constantini, V., Goyal, S., ... & Saif, L. J. (2010). Fate and transport of zoonotic, bacterial, viral, and parasitic pathogens during swine manure treatment, storage, and land application. Journal of animal science, 88(13), E84-E94.

Ziprin, R. L., & Harvey, R. B. (2004). Inability of cecal microflora to promote reversion of viable nonculturable *Campylobacter jejuni*. Avian Diseases, 48(3), 647–650.


@Rima Shrestha

----------**********----------