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