

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Molecular epidemiology of Shiga toxin-producing
Escherichia coli (STEC) O157 and non-O157 STEC in
calves in the North Island of New Zealand**

A thesis submitted in partial fulfilment of the requirements for
the degree of Doctor of Philosophy

Massey University, Palmerston North

New Zealand

by

Hamid Irshad

February 2013

Abstract

Shiga toxin-producing *Escherichia coli* (STEC) are common inhabitants of the ruminant gastrointestinal tract and have emerged as important zoonotic pathogens of global public health concern. Four studies were carried out in the North Island of New Zealand to understand the epidemiology of STEC O157 and non-O157 STEC in calves.

The aim of the first study was to determine population structure, transmission dynamics and spatial relationship between genotypes of two zoonotic pathogens, STEC and *Campylobacter jejuni*, found on farms in a defined catchment area in the Waikato region. Pooled faecal samples (n=72) obtained from calves grazing in the catchment were analysed by RT-PCR for *E. coli* O26, O103, O111, O145 and O157. The number of samples positive for O26 (30/72) was high compared to O103 (7/72), O145 (3/72) and O157 (2/72) while no samples were positive for O111. Using immuno-magnetic separation 18 O26, two O103 and a single O145 isolate were recovered from RT-PCR positive samples. Fifty-three *C. jejuni* isolates were recovered from 72 pooled faecal samples. *E. coli* O26 and *C. jejuni* isolates were genotyped using pulsed field gel electrophoresis (PFGE) and multilocus sequence typing respectively. All O26 isolates could be divided into three clusters on the basis of PFGE using *Xba*I. These results indicated that *E. coli* O26 isolates recovered from calves on the same farm were generally more similar than isolates recovered from different farms in the catchment. There were a total of 13 different sequence types (STs) of *C. jejuni* isolated from the cattle and most of the molecular variation (approximately 75%) resided between animals within farms. This study suggested a high level of within farm transmission, and limited between-farm transmission indicating direct contact between animals or sharing of personnel or equipment could be important routes of transmission and good biosecurity measures may be helpful in reducing localised transmission.

The aim of the second study was to determine the distribution of virulence genes (*stx1*, *stx2*, *eae* and *ehxA*) in *E. coli* isolates from dairy calves less than a week of age (bobby calves). Sampling was carried out by systematic collection of recto-anal mucosal swab (RAMS) samples from bobby calves slaughtered at two abattoirs in the North Island of New Zealand. The samples were inoculated onto tryptone bile X-glucuronide (TBX) and sorbitol MacConkey agar (CT-SMAC). Blue and white colonies (one each) and purple and grey colonies (one each) were selected at random from TBX and CT-SMAC plates respectively. In total 975 *E. coli* isolates obtained from the two media were analysed by multiplex PCR to detect *stx1*, *stx2*, *eae* and *ehxA* genes. The most common combination of virulence markers were *eae*, *ehxA* (n=35) followed by *eae* (n=9). Only eight STEC were identified of which four were *stx2*, *eae*, *ehxA*-positive (3 O157:H7 and 1 O68:H24), three were *stx1*, *eae*, *ehxA* (2 O26 and 1 O71:HR) and there was one *stx2*-only isolate (ONT:HNM). All the isolates could be divided into 15 clusters with >70% similarity using *Xba*I by PFGE. These findings indicate that STEC of public health significance such as O157 and O26 are present in bobby calves and may represent an important source of human infection in New Zealand.

In the third study samples obtained in the second study were processed to determine the occurrence and spatial distribution of *E. coli* O26, O103, O111 and O145 in bobby calves in the North Island of New Zealand. The association of IgG concentration, weight, sex and breed with occurrence of O26, O103, O111 and O145 as determined by direct RT-PCR on RAMS enrichments was also investigated. Using RT-PCR 134/299 (44.8%) of RAMS were positive for O26, 68/299 (22.7%) for O103 and 47/299 (15.7%) for O145. No RAMS samples were RT-PCR positive with O111-specific primers. The success of isolation by culture of *E.*

E. coli O26 (49/134 isolates) was higher from RT-PCR positive samples as compared to O103 (4/68 isolates) and O145 (5/47 isolates). Using a logistic regression model no association was observed between PCR prevalence and the variables IgG, sex or breed of the calves. However, calves positive for O26 were more likely to be positive for O103 and vice versa (P=0.01) and similar association was found between calves positive for O145 and O103 (P=0.03). O26 isolates could be grouped into four clusters (A-D) of >70% similarity using *Xba*I PFGE. K function analysis did not indicate any evidence of spatial clustering of farms positive for O26, O103 or O145. This study indicated that O26 is more prevalent in bobby calves as compared to O103 and O145 and colostrum feeding may not be helpful in reducing the carriage of *E. coli* O26, O103 and O145.

In the fourth study *E. coli* isolates (n=137) obtained from previous studies were genotyped using PFGE and allelic profiling (based on 8 genes). The *eae* and *ehxA* genes of *E. coli* isolates were subtyped using PCR-restriction fragment length polymorphism (RFLP) analysis. Endonuclease digestion of *ehxA* PCR products with *Taq*I from 121 *ehxA* positive *E. coli* isolates resulted in six *ehxA* subtypes (A-F). Endonuclease digestion with *Hha*I or *Hae*III enzymes of the 129 *eae* positive isolates indicated that 82 were β followed by ϵ (n=34), γ (n=11) and δ/κ (n=2). An association between *eae* subtype and *E. coli* serogroup was also observed. All O26 isolates were subtype β , all O103 isolates were subtype ϵ and all O157 isolates were γ . *E. coli* isolates were also analysed for plasmid-associated alleles *espP*, *etpD*, *katP* and α -*hly* genes. Of 137 isolates 93 (67.8%) were positive for *espP*, 32 for *etpD* (23.3%), 76 for *katP* (55.4%) and nine for the α -*hly* (6.5%) gene. The genotyping ability of allelic profiling was compared with PFGE profiling using PERMANOVA and multidimensional scaling and the results indicated that isolates having similar allelic profiles had similar PFGE profiles and tended to cluster together. The results also indicated that *eae* and *ehxA* profiles could explain a high population of the variance in PFGE profiles. These results may provide basis for the development of new genotyping method for *E. coli* isolates, having high discriminatory power and being easier to perform. The *E. coli* serogroups and *eae* and *ehxA* subtypes observed in these serogroups have also been observed from human *E. coli* isolates indicating the public health significance of these isolates.

These studies provide useful information about epidemiology of O157 and non-O157 in calves in the North Island of New Zealand and indicate that calves may be an asymptomatic reservoir of STEC and a possible source of infection for humans. Better understanding of the population structure and transmission of STEC would help in devising appropriate control strategies. The implementation of these control measures could reduce the prevalence of STEC in calves and thus reduce transmission to humans.

Contents

1	Chapter 1	13
1.1	Literature review	13
1.2	Pathotypes of <i>E. coli</i>	14
1.3	History of Shiga toxin-producing <i>E. coli</i>	15
1.4	Epidemiology of STEC	16
1.4.1	Prevalence of STEC in cattle	16
1.4.2	Prevalence of STEC in New Zealand cattle	17
1.4.3	Prevalence of <i>E. coli</i> O26, O103, O111, O145 and O157 in cattle	17
1.4.4	Transmission of STEC between cattle	18
1.4.5	Seasonality	19
1.4.6	Risk Factors	19
1.5	Control of STEC in cattle	21
1.5.1	Vaccination	22
1.5.2	Probiotics	22
1.5.3	Bacteriophages	22
1.6	Epidemiology of STEC in humans	23
1.6.1	Epidemiology of STEC in humans in New Zealand	25
1.6.2	Incubation period	26
1.6.3	Dose response	26
1.6.4	Clinical manifestation	26
1.6.5	Vehicles of STEC transmission	27
1.7	Mechanism of virulence and pathogenicity	27
1.7.1	<i>E. coli</i> attachment and effacement gene	27
1.7.2	Shiga toxins	28
1.7.3	Plasmid Associated virulence factors	29
1.8	Importance of STEC for the New Zealand economy	29
1.9	Molecular epidemiology	30
1.9.1	Molecular epidemiology of <i>E. coli</i> O157	30
1.9.2	Molecular epidemiology of non-O157 STEC	31
1.9.3	Molecular Evolution	32
1.10	<i>Campylobacter</i>	34
1.10.1	Cattle as a reservoir of <i>Campylobacter</i>	34
1.10.2	Role of water in <i>Campylobacter</i> transmission	36
1.11	Conclusion	36
2	Chapter 2	44
2.1	Abstract	44
2.2	Introduction	45
2.3	Materials and methods	46
2.3.1	Isolation and characterisation of STEC	46

2.3.2	Isolation and characterisation of <i>C. jejuni</i>	48
2.3.3	Statistical analysis	49
2.4	Results	50
2.4.1	Population structure and characterisation of STEC	50
2.4.2	Population structure and characterisation of <i>C. jejuni</i>	50
2.5	Discussion	51
3	Chapter 3	64
3.1	Abstract	64
3.2	Introduction	65
3.3	Materials and methods	66
3.4	Results	68
3.5	Discussion	69
4	Chapter 4	78
4.1	Abstract	78
4.2	Introduction	79
4.3	Materials and methods	80
4.3.1	Statistical analysis	82
4.4	Results	84
4.5	Discussion	85
5	Chapter 5	98
5.1	Abstract	98
5.2	Introduction	99
5.3	Materials and methods	100
5.4	Results	104
5.5	Discussion	107
6	Chapter 6	122
7	Appendix	128
7.1	Appendix 1	128
7.2	Appendix 2	135

List of Tables

1.1	Classification of Shiga toxin-producing <i>Escherichia coli</i> serotypes into seropathotypes (Karmali et al., 2003)	41
1.2	The prevalence of Shiga toxin-producing <i>Escherichia coli</i> in cattle in different countries	42
2.1	Virulence factors associated with O157 and non-O157 serogroups obtained from 72 composite faecal samples. No samples were positive for serogroup O111 by RT-PCR.	56
2.2	The distribution of <i>C. jejuni</i> sequence types (STs) isolated from cattle and water from catchment stream. The STs indicated by "W" are associated with wildlife in New Zealand.	57
2.3	Global Analysis of Molecular Variance considering the farm structure outlined in Figure 2.1a.	58
2.4	Global Analysis of Molecular Variance considering the farm structure outlined in Figure 2.1b.	59
3.1	Distribution of virulence genes (<i>stx1</i> , <i>stx2</i> , <i>eae</i> , <i>ehxA</i>) in <i>E. coli</i> isolates obtained from recto-anal mucosal swabs (n=299) of bobby calves on sorbitol MacConkey agar (CT-SMAC) and tryptone bile X-glucuronide plates. Sorbitol fermenting (SF) and non-sorbitol fermenting (NSF) colonies were obtained on CT-SMAC plates while β -glucuronidase positive (β GP) and β -glucuronidase negative (β GN) colonies were obtained on TBX plates.	74
3.2	The proportion of samples positive for <i>E. coli</i> isolates containing virulence genes (<i>stx1</i> , <i>stx2</i> , <i>eae</i> , <i>ehxA</i>) from 299 recto-anal mucosal swabs of bobby calves from three regions of the North Island of New Zealand.	75
4.1	Primers used for PCR in <i>stx</i> -encoding bacteriophage insertion (SBI) typing.	89
4.2	The univariable and multivariable relationships between the likelihood a calf is positive by RT-PCR for <i>E. coli</i> O26 and five explanatory variables. The reference category for breed was Friesian	90
4.3	The univariable and multivariable relationships between the likelihood a calf is positive by RT-PCR for <i>E. coli</i> O103 and five explanatory variables. The reference category for breed was Friesian	91
4.4	The univariable and multivariable relationships between the likelihood a calf is positive by RT-PCR for <i>E. coli</i> O145 and five explanatory variables. The reference category for breed was Friesian	92
5.1	Relationship between serogroups and <i>ehxA</i> subtypes in <i>E. coli</i> (n=121) isolates recovered from calves in the North Island of New Zealand.	111
5.2	Relationship between <i>ehxA</i> subtypes and zone of haemolysis on sheep blood agar plates in <i>E. coli</i> (n=121) isolates recovered from calves in the North Island of New Zealand.	112

5.3	Relationship between serogroups and zone of haemolysis on sheep blood agar plates in <i>E. coli</i> (n=137) isolates recovered from calves in the North Island of New Zealand.	113
5.4	Relationship between <i>eae</i> subtypes and virulence profile of <i>E. coli</i> (n=129) isolates recovered from calves in the North Island of New Zealand.	114
5.5	Relationship between serogroups and <i>eae</i> subtypes in <i>E. coli</i> (n=129) isolates recovered from calves in the North Island of New Zealand.	115
5.6	The relationship between various combinations of plasmid-associated genes (<i>ehxA</i> , <i>espP</i> , <i>etpD</i> and <i>katP</i>) and serotypes of <i>E. coli</i> isolates (n=137) recovered from bobby calves in the North Island of New Zealand.	116
5.7	PERMANOVA analysis results indicating the contribution of allelic profile (8 genes) in the variation of PFGE profiles of <i>E. coli</i> isolates obtained from calves in the North Island of New Zealand. Variable serogroup was not included in the model to assess the contribution of <i>stx2</i> and <i>etpD</i> towards the model.	117
5.8	PERMANOVA analysis results indicating the contribution of allelic profile (8 genes) in the variation of PFGE profiles of <i>E. coli</i> isolates obtained from calves in the North Island of New Zealand. Variables <i>stx2</i> and <i>etpD</i> were not included in the model to assess the contribution of variable serogroup towards the model. . . .	118
7.1	PERMANOVA analysis results indicating the contribution of allelic profile (8 genes) in the variation of PFGE profiles of <i>E. coli</i> isolates obtained from bobby calves (excluding catchment isolates) in the North Island of New Zealand.	136
7.2	PERMANOVA analysis results indicating the contribution of allelic profile (8 genes) in the variation of PFGE profiles of <i>E. coli</i> isolates obtained from bobby calves (excluding catchment isolates) in the North Island of New Zealand. Variable serogroup was not included in the model.	137
7.3	PERMANOVA analysis results indicating the contribution of allelic profile (8 genes) in the variation of PFGE profiles of <i>E. coli</i> isolates obtained from bobby calves (excluding catchment isolates) in the North Island of New Zealand. Variables <i>stx2</i> and <i>etpD</i> were not included in the model.	138

List of Figures

1.1	Notified cases of Shiga toxin-producing <i>Escherichia coli</i> in New Zealand from 1997-2011 (Environmental Science and Research, 2011)	38
1.2	Proposed evolutionary model for emergence of <i>E. coli</i> O157:H7 complex based on mutation in <i>uidA</i> , <i>stx</i> producing, sorbitol and β -glucuronidase phenotypes, and multilocus enzyme electrophoretic profiles of <i>E. coli</i> O157:H7 and its relatives. Phenotypes of ancestor A1-A6 are shown; changes predicted to have occurred are in bold. Representative isolates are given below each electrophoretic type (ET). Strain with tracts of ancestor A3 (shaded circle) has not been reported (Feng et al., 1998)	39
1.3	Evolutionary scenario for EHEC 1 pathogens. <i>E. coli</i> O55:H7 belongs to the most ancestral subgroup of the EHEC 1 clade (subgroup A). The gray sphere depicts a probably extinct intermediate between O55:H7 and <i>E. coli</i> expressing the O157 lipopolysaccharide, with subgroup B consisting of the sorbitol fermenting (SF) O157:H- and subgroup C consisting of <i>E. coli</i> O157:H7. Critical intraclade events are noted. Intracluster ovals represent genomically sequenced strains (dark-pink), strains used for SNP consensus sampling (green), inferred founders (pale-pink), and postulated organisms that are immediate progenitors to the next cluster or subgroup (white). Screened strains were assigned to the main branch if they had each of the three signature SNPs among the 111 shared SNPs (38/48), and the minor branch if they lacked this SNPs (10/48). Distances not drawn to scale (Leopold et al., 2009)	40
2.1	Five distinct groups marked A-E were evident (Fig 2.1a). The road network is shown in bold black and the stream network in blue lines. Figure 2.1b shows an alternative grouping of farms according to whether they were adjacent to the main stream (X=green) or away from the main stream (Y=orange)	60
2.2	Clustering (UPGMA dendrogram) of the PFGE profiles of 18 O26 isolates from 17 composite faecal samples taken from 11 farms in catchment. Each farm is colour coded and the catchment region is indicated. Note two isolates from sample EcCa38 are included in the analysis. The last four columns indicate the presence of the <i>stx</i> , <i>eae</i> and <i>ehxA</i> virulence genes. The last lane is the <i>Salmonella</i> serotype Braenderup reference standard (H9812).	61
2.3	Genotype (sequence types = STs) richness in cattle and water samples. Rarefaction curves indicating the estimated number of unique STs in cattle (red) and water (blue) for varying sample sizes, with 95% confidence intervals.	62

2.4	A Minimum Spanning Tree showing the multilocus sequence types (STs) isolated from cattle (pink) and water (blue) in the catchment stream. Each circle represents a ST; the size of the circle is proportional to the number of isolates and the area of each circle represents the relative frequency of each source. Isolates are grouped according to their genetic relatedness; short solid lines between STs indicate they vary at just one of the seven MLST loci (i.e. they are Single Locus Variants or SLVs).	63
3.1	Map showing the distribution of <i>E. coli</i> isolates positive for virulence genes (<i>stx1</i> , <i>stx2</i> , <i>eae</i> , <i>ehxA</i>) in the North Island of New Zealand. The data are aggregated to 5x5km grid cells. If a single animal was positive the grid cell was coloured green. If no animals sampled were positive the cell was coloured beige.	76
3.2	Clustering using UPGMA and Dice co-efficient of the PFGE profiles of <i>E. coli</i> isolates positive for <i>stx1</i> , <i>stx2</i> , <i>eae</i> and <i>ehxA</i> from calves with >70% similarity cut off using <i>XbaI</i> . The last lane is the <i>Salmonella</i> serotype Braenderup reference standard (H9812).	77
4.1	Box plot showing the relationship between IgG and RT-PCR results for O26, O103 and O145	93
4.2	Box plot showing the relationship between weight and RT-PCR results for O26, O103 and O145	94
4.3	Maps showing the distribution of the farm of origin of <i>E. coli</i> O26, O103 and O145-positive calves in North Island, New Zealand. The data are aggregated to 5x5 km grid cells. Only cells containing farms that submitted calves to the abattoir that were subsequently sampled are displayed.	95
4.4	K function analysis of <i>E. coli</i> O26, O103 and O145-positive (cases) and <i>E. coli</i> O26, O103 and O145-negative (control) farms. The black line represents the difference between the two K functions and dashed lines represent the upper and lower simulation envelopes.	96
4.5	Clustering (UPGMA dendrogram) of the PFGE profiles of <i>E. coli</i> O26 isolates recovered from bobby calves from the North Island of New Zealand with >70% similarity cut off using <i>XbaI</i> . The last lane is the <i>Salmonella</i> serotype Braenderup reference standard (H9812).	97
5.1	Clustering (UPGMA dendrogram) of the PFGE profiles of <i>E. coli</i> isolates from calves with >70% similarity cut off using <i>XbaI</i> . The column PFGE group refers to the cluster of <i>E. coli</i> isolates. The last lane is the <i>Salmonella</i> serotype Braenderup reference standard (H9812). The column BG (β glucuronidase) shows the BG activity of isolates. The column a-hly (α -haemolysin) shows the presence (indicated by 1) or absence (indicated by 0) of α -hly gene and ND (not done) means isolate was not analysed for presence or absence of α -hly gene	119

5.2	Hierarchical clustering analysis of <i>eae</i> and <i>ehxA</i> positive <i>E. coli</i> isolates obtained from calves of the North Island of New Zealand based on the data from eight PCR(-RFLP) analysis: <i>stx1</i> , <i>stx2</i> , <i>eae</i> , <i>ehxA</i> , <i>espP</i> , <i>etpD</i> , <i>katP</i> and α - <i>hly</i> genes. Superscripts: a, β -glucuronidase; b, sorbitol fermenting; c, non-rhamnose fermenting; d, not done; e, untypable; f, rhamnose fermenting; g, non-sorbitol fermenting	120
5.3	Multi dimensional scaling plot showing association between PFGE and allelic profiles (8 genes) of 137 <i>E. coli</i> isolates obtained from calves from the North Island of New Zealand	121
7.1	Two dimensional plot showing the results of multiple correspondence analysis model of RT-PCR results (O26, O103, O145 and O157), weight, sex, breed and concentration of IgG of samples obtained from 299 bobby calves in the North Island of New Zealand. The variable PCR was used as supplementary variable (Δ)	129
7.2	Two dimensional plot showing the results of multiple correspondence analysis model of RT-PCR results (O26, O103, O145 and O157), weight, sex, breed and concentration of IgG of samples obtained from 299 bobby calves in the North Island of New Zealand. The variable region was used as supplementary variable (Δ)	130
7.3	Two dimensional plot showing the results of multiple correspondence analysis model of RT-PCR results (O26, O103, O145 and O157), weight, sex, breed and concentration of IgG of samples obtained from 299 bobby calves in the North Island of New Zealand. The variable breed was used as supplementary variable (Δ)	131
7.4	Two dimensional plot showing the results of multiple correspondence analysis model of RT-PCR results (O26, O103, O145 and O157), weight, sex, breed and concentration of IgG of samples obtained from 299 bobby calves in the North Island of New Zealand. The variable sex was used as supplementary variable (Δ)	132
7.5	Two dimensional plot showing the results of multiple correspondence analysis model of RT-PCR results (O26, O103, O145 and O157), weight, sex, breed and concentration of IgG of samples obtained from 299 bobby calves in the North Island of New Zealand. The variable weight was used as supplementary variable (Δ)	133
7.6	Two dimensional plot showing the results of multiple correspondence analysis model of RT-PCR results (O26, O103, O145 and O157), weight, sex, breed and concentration of IgG of samples obtained from 299 bobby calves in the North Island of New Zealand. The variable IgG was used as supplementary variable (Δ)	134

Acknowledgements

All the praises and thanks are for Almighty Allah the most gracious and merciful, the source of all knowledge and wisdom, who has blessed me with audacity and endowed me with potential and ability for completing this manuscript.

I wish to express my profound gratitude to my supervisor, Prof. Dr. Nigel P. French, for his keen interest, inspiring guidance, and valuable suggestions in planning, conducting and writing this manuscript.

I am extremely thankful and greatly indebted to Dr. Adrian Cookson, for his kind advice and constant encouragement throughout the course of study.

I am also extremely thankful to Dr. Deborah J. Prattley, for providing assistance in organising sampling schedule and guidance during the study.

I am also thankful to Dr. Grant Hotter for his valuable suggestions in the start of my study but unfortunately he died in 2009.

Thanks are also due to Dr. Anne Midwinter for providing help and guidance in the lab work throughout my study.

I am also thankful to Sara Vaughan for providing assistance throughout sampling and driving all the way to Waitotara and Eltham for sampling early in the morning.

I am also thankful to Errol Kwan for providing me help and guidance in the lab work.

I am also thankful to Rebecca Pattison for introducing me with PFGE.

I am also thankful to Dr. Jonathan Marshall for providing help in analysing the results.

I am also thankful to Dr. Julie Collins-Emerson and Dr. Eve Pleydell for their valuable suggestions and encouragement throughout my study.

I am also thankful to my fellow students especially Patricia Jaros, Barbara Binney, Anjia Friedrich, Ben Phiri and Rima. They had been very kind to me in the moments of stress.

I am also thankful to my friends Aamir, Islah, Faisal Rana, Zahid, Zaka, Ibrar Qureshi and Saleh Umair for their help and support throughout my study.

I am also thankful to Dr. Zulfiqar Butt and Babar Bhai for introducing me to Palmerston North. They always provided valuable suggestions and support in difficult times.

With sincere thanks, I am indebted to my affectionate parents, my wife, brothers, sisters and all my family members as they always remembered me in their prayers and provided me moral assistance for completing my study.

Last but not least I am also thankful to Higher Education Commission Pakistan for providing me financial assistance that made my study and stay reality in New Zealand.