

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.

Phenotype-genotype relationships of *Escherichia coli* O157 and O26  
isolates from New Zealand

The crest of Massey University, featuring a shield with a sunburst and a book, surrounded by a laurel wreath.

A thesis presented in partial fulfilment of  
the requirements for the degree of  
Master of Science  
in  
Microbiology  
at Massey University,  
Manawatū, New Zealand.

Tessy Thankachen George

2017

mEpiLab

Institute of Veterinary, Animal and Biomedical Sciences

Massey University



## Abstract

Shiga toxin-producing *Escherichia coli* (STEC) is a zoonotic pathogen responsible for causing severe manifestations of gastroenteritis in humans worldwide. STEC is transmitted to humans through the consumption of compromised food and water, or direct animal contact. *Escherichia coli* (*E. coli*) O157 and *E. coli* O26 are considered to be among the important STEC serogroups worldwide due to their association with outbreaks and clinical cases of haemolytic uraemic syndrome and haemorrhagic colitis. This study is concerned with the application of phenotypic microarray technology and high-throughput sequencing technology to characterise *E. coli* O157 and *E. coli* O26 New Zealand isolates.

In this study, 190 phenotypes of carbon sources utilisation were studied in isolates belonging to *E. coli* O157 and O26 serogroups. The isolates were divided into groups based on the fermentation of sorbitol and rhamnose in *E. coli* O157 and O26, respectively. All the isolates respired on approximately 40% of the carbon sources. The sorbitol positive *E. coli* O157 isolates respired on more carbon sources compared to the sorbitol negative *E. coli* O157, rhamnose negative and rhamnose positive *E. coli* O26 isolates.

Genomic analysis showed that *E. coli* O157 isolates had shorter genomes compared to those of the *E. coli* O26 isolates. The core genome comparisons revealed differences between and within the *E. coli* O157 and O26 serogroups. Clustering of *E. coli* O157 and O26 isolates based on sorbitol and rhamnose fermentation, respectively, was observed.

The results obtained from this study illustrated that phenotypic and genotypic differences existed within the *E. coli* O157 isolates. The findings of the current study also demonstrated that while the *E. coli* O26 isolates had similar phenotypic characteristics, genotypic differences existed within the isolates.

## Acknowledgements

I sincerely thank my supervisors Dr Anne Midwinter, Associate Professor Patrick Biggs, Dr Julie Collins-Emerson and Professor Nigel French for their valuable suggestions and guidance throughout my study.

Thank you to Massey University and the Institute of Veterinary, Animal and Biomedical Sciences for providing me financial support and funds towards my research. Thank you to Kathy Hamilton, Professor Kevin Stafford and Debbie Hill for their administrative support.

I sincerely thank Dr Ben Phiri for proof-reading my thesis. I would like to thank Dr Jonathan Marshall for his helpful tips with the software R and Simon Verschaffelt for his technical support. Thank you to the staff and postgraduate students at mEpiLab, Massey University, I enjoyed working alongside all of you. I would like to thank Sameer Siddiqui, Rima Shrestha, Saroj Yadav, Punja Prasertsri, Antoine Nohra, José Solis-Ramirez, Kruno Bojanić, Doris Adeyinka and Shirli Notcovich for your kindness and friendship.

Most importantly, I would like to thank my family for their love, support and encouragement.

Let all that I am praise the Lord;  
May I never forget the good things He does for me.  
Psalm 103:2

## Abbreviations

%	Percentage
°C	Degree Celsius
IIA <sup>glc</sup>	Glucose-specific IIA protein
Acetyl-CoA	Acetyl coenzyme A
ADJ	Adjusted
AE	Attaching and effacing lesion
AFLP	Amplified fragment length polymorphism
<i>argW</i>	Stx-encoding bacteriophage insertion site
ATP	Adenosine triphosphate
BAC	Bacterial artificial chromosome
BD	Becton Dickinson and Company
BGI	Beijing Genomics Institute
BLAST	Basic Local Alignment Search Tool
bp	Base pairs
BWA	Burrows-Wheeler Alignment
cAMP	Cyclic adenosine monophosphate
CDC	Centers for Disease Control and Prevention
CO <sub>2</sub>	Carbon dioxide
COGs	Clusters of Orthologous Groups of proteins
CRP	cAMP receptor protein
CTA	Cystine trypticase agar
CT-RMAC	Rhamnose MacConkey medium containing cefixime and tellurite
CT-SMAC	Sorbitol MacConkey medium containing cefixime and tellurite
DAEC	Diffusely adhering <i>E. coli</i>
DNA	Deoxyribonucleic acid
<i>dsdA</i>	D-serine deaminase or dehydratase
<i>dsdC</i>	LysR-type transcriptional regulator
<i>dsdCXA</i>	D-serine tolerance locus
<i>dsdX</i>	D-serine transporter
EII	Enzymes II
<i>eae</i>	<i>E. coli</i> attaching and effacing

EAEC	Enteroaggregative <i>E. coli</i>
EAHEC	Enteroaggregative haemorrhagic <i>E. coli</i>
EDP	Entner-Doudoroff pathway
EHEC	Enterohaemorrhagic <i>E. coli</i>
<i>ehxA</i>	Haemolysin
EIA	Enzyme immunoassay
EIEC	Enteroinvasive <i>E. coli</i>
EPEC	Enteropathogenic <i>E. coli</i>
ERL	Enteric Reference Laboratory
<i>espA</i>	<i>E. coli</i> secreted protein
<i>espB</i>	<i>E. coli</i> secreted protein
<i>espD</i>	<i>E. coli</i> secreted protein
<i>espP</i>	Protease
ESR	Institute of Environment Science and Research Limited
ETEC	Enterotoxigenic <i>E. coli</i>
<i>E. coli</i>	<i>Escherichia coli</i>
<i>E. fergusonii</i>	<i>Escherichia fergusonii</i>
<i>fliC</i>	Flagellin
<i>fliC<sub>H11</sub></i>	Flagellin H11
<i>fyuA</i>	High-pathogenicity island gene
g	Grams
Gb	Gigabases
GUD	beta-glucuronidase
H	Flagellar antigen
h	Hours
HC	Haemorrhagic colitis
HGT	Horizontal gene transfer
HPI	High-pathogenicity island
HPr	Heat-stable protein
HUS	Haemolytic uraemic syndrome
<i>irp2</i>	High-pathogenicity island gene
K	Capsular antigen
KEGG	Kyoto Encyclopedia of Genes and Genomes

LEE	Locus of enterocyte effacement
LGT	Lateral gene transfer
M	Molar
Mb	Megabases
MDa	Megadalton
MEE	Multilocus enzyme electrophoresis
μl	Microlitres
ml	Millilitres
Min	Minutes
MLST	Multilocus sequence typing
MPI	Ministry of Primary Industries
NADH	Nicotinamide adenine dinucleotide
NCBI	National Centre for Biotechnology Information
ng	Nanograms
NI	North Island
<i>nle</i>	Non-LEE effector
NMD	National Microbiological Database
NZ	New Zealand
NZGL	New Zealand Genomics Limited
O	Somatic antigen
PAI	Pathogenicity island
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEP	Phosphoenolpyruvate
PFGE	Pulsed field gel electrophoresis
pH	Potential of hydrogen
PM	Phenotypic microarray
PPP	Pentose phosphate pathway
PTS	Phosphotransferase system
QC	Quality control
RFLP	Restriction fragment length polymorphism
RMAC	Rhamnose MacConkey medium
RN	Rhamnose non-fermenting or negative

RNA	Ribonucleic acid
RP	Rhamnose fermenting or positive
SBI	Shiga toxin-encoding bacteriophage insertion
<i>sbcB</i>	Stx-encoding bacteriophage insertion site
SBS	Synthesis by synthesis
SI	South Island
SMAC	Sorbitol MacConkey medium
SN	Sorbitol non-fermenting or negative
SP	Sorbitol fermenting or positive
STEC	Shiga toxin-producing <i>E. coli</i>
Stx, <i>stx</i>	Shiga toxin
<i>stx1</i>	Shiga toxin 1
<i>stx2</i>	Shiga toxin 2
<i>stx2c</i>	Subtype Shiga toxin 2
T3SS	Type III secretion system
TTC	2,3,5-triphenyl tetrazolium chloride
USA	United States of America
UTI	Urinary tract infection
v	Version
VTEC	Verotoxin producing <i>E. coli</i>
<i>wrbA</i>	Stx-encoding bacteriophage insertion site
<i>yehV</i>	Stx-encoding bacteriophage insertion site
<i>yhaJ</i>	LysR-type transcriptional regulator
<i>yhaO</i>	Inner membrane transporter
<i>yhaOMKJ</i>	D-serine sensory locus

# Contents

Abstract	iii
Acknowledgements	iv
Abbreviations	v
Chapter One	
General introduction	1
1.1 Background	1
1.2 Shiga toxin-producing <i>Escherichia coli</i> in New Zealand	1
1.3 Thesis objectives and structure	2
Chapter Two	
Literature review	
2.1 Introduction	4
2.2 <i>Escherichia coli</i>	5
2.2.1 Shiga toxin producing <i>Escherichia coli</i>	5
2.2.2 Clinical spectrum of STEC	6
2.2.3 Virulence characteristics of STEC	7
2.2.4 Shiga toxin-encoding bacteriophage insertion typing	9
2.2.5 Detection, isolation and identification of STEC	9
2.2.6 Prevalence of STEC	11
2.2.7 Epidemiology of STEC outbreaks	12
2.2.8 Epidemiology of STEC outbreaks in New Zealand	14
2.3 Impact of compromised meat on the meat industry	15
2.4 Evolution of <i>Escherichia coli</i>	16
2.5 Carbohydrate metabolism in bacteria	19
2.6 Phenotype and genotype	22
2.7 Phenotypic microarray technology	23
2.8 Whole genome sequencing	24
2.8.1 Illumina technology	25

## Chapter Three

### Materials and methods

3.1	Bacterial isolates	29
3.2	Shiga toxin-encoding bacteriophage insertion typing	29
3.3	Phenotypic microarray technology	34
3.3.1	Preparation of phenotypic microarray plates	34
3.3.2	Data recording and analysis of kinetic data	35
3.4	Confirmatory tests for phenotypes	36
3.4.1	Cystine trypticase agar medium with sorbitol	36
3.4.2	MacConkey agar with added carbon sources	37
3.4.2.1	MacConkey agar with D-serine or p-hydroxyphenylacetic acid or glycolic acid or glyoxylic acid	37
3.4.2.2	MacConkey agar with propionic acid	38
3.4.2.3	MacConkey control agar	39
3.5	Genomic DNA preparation	39
3.6	Genome sequencing	39
3.6.1	Library preparation method	39
3.6.2	Illumina MiSeq runs	40
3.7	Genome assembly and contig selection	41
3.8	Genome clustering and reference genomes	41
3.8.1	Clustering and core genome tree construction	41
3.9	Phenotypic-genotypic analysis	42
3.9.1	Clusters of Orthologous Groups of proteins analysis	42

## Chapter Four

### Results

4.1	Bacterial isolates	47
4.2	Shiga toxin-encoding bacteriophage insertion typing	47
4.3	Phenotypic microarray technology	49
4.3.1	Carbon utilisation profiles of <i>E. coli</i> O157 isolates	49
4.3.2	Carbon utilisation profiles of <i>E. coli</i> O26 isolates	51
4.3.3	Reproducibility of phenotypic microarray tests	56
4.4	Confirmatory tests for phenotypes	56
4.4.1	Cystine trypticase agar medium with sorbitol	56

4.4.2	MacConkey agar with added carbon sources	57
4.5	Genome sequencing	63
4.5.1	Core genome comparisons	63
4.5.1.1	<i>E. coli</i> O157 and <i>E. coli</i> O26 New Zealand isolates	63
4.5.1.2	<i>E. coli</i> O157 New Zealand isolates and <i>E. coli</i> O157 reference isolates	64
4.5.1.3	<i>E. coli</i> O26 New Zealand isolates and <i>E. coli</i> O26 reference isolates	64
4.5.1.4	<i>E. coli</i> O157 and O26 New Zealand isolates and <i>E. fergusonii</i> reference isolates	64
4.5.2	Clusters of Orthologous Groups of proteins analysis	65
 Chapter Five		
Discussion		
5.1	Introduction	75
5.2	Shiga toxin-encoding bacteriophage insertion typing	75
5.3	Phenotypic microarray technology and analysis	76
5.4	Genome sequencing and analysis	78
5.5	Limitations of the research	81
5.6	Conclusion	82
References		83
 Appendix		
A1	Heatmap of <i>E. coli</i> O157 respiration changes on PM1 including duplicate and replicate assays	106
A2	Heatmap of <i>E. coli</i> O157 respiration changes on PM2A including duplicate and replicate assays	107
A3	Heatmap of <i>E. coli</i> O26 respiration changes on PM2A	108

## List of Tables

3.1	Features of <i>E. coli</i> O157 and O26 New Zealand isolates	31
3.2	Features of <i>E. coli</i> O157 New Zealand isolates from previous study	31
3.3	Primers used for PCR in SBI typing	32
3.4	SBI typing summary of <i>E. coli</i> isolates	33
3.5	Summary of genome characteristics of <i>E. coli</i> O157, O26 and <i>E. fergusonii</i> reference isolates obtained from NCBI database	44
3.6	Classified functions of COGs	46
4.1	SBI types, sorbitol and rhamnose characterisation of <i>E. coli</i> O157 and O26 New Zealand isolates	48
4.2	Percentage of substrate utilisation of sorbitol negative <i>E. coli</i> O157 isolates in PM1 and 2A	55
4.3	Percentage of substrate utilisation of sorbitol positive <i>E. coli</i> O157 isolates in PM1 and 2A	55
4.4	Percentage of substrate utilisation of rhamnose negative <i>E. coli</i> O26 isolates in PM1 and 2A	55
4.5	Percentage of substrate utilisation of rhamnose positive <i>E. coli</i> O26 isolates in PM1 and 2A	55
4.6	Summary of the confirmatory test results and their respective PM1 results	62
4.7	Summary of genome characteristics of <i>E. coli</i> O157 and O26 New Zealand isolates	67
4.8	COG analysis of carbon metabolism in <i>E. coli</i> O157 isolates	73

## List of Figures

2.1	<i>Escherichia coli</i> ( <i>E. coli</i> )	13
2.2	Routes of transmission of <i>E. coli</i>	13
2.3	Meat carcasses in an abattoir	18
2.4	Proposed evolutionary model of <i>E. coli</i> O157:H7	18
2.5	Major pathways of carbohydrate metabolism in prokaryotes	20
2.6	Citric acid cycle and relate anaplerotic pathways	21
2.7	Respiration pathways coupled to cell physiology	27
2.8	Phenotypic microarray (PM) technology	27
2.9	Illumina workflow	28
4.1	Isolate EcCa519a in PM1	52
4.2	Heatmap of <i>E. coli</i> O157 respiration changes on PM1	53
4.3	Heatmap of <i>E. coli</i> O26 respiration changes on PM1	54
4.4	Reproducibility of phenotypic microarray tests	59
4.5	<i>E. coli</i> O157 New Zealand isolates in CTA medium with sorbitol	60
4.6	<i>E. coli</i> O157 New Zealand isolates on MacConkey media with added carbon sources	61
4.7	Cumulative frequency of trimmed read lengths for isolate EcCa24a	66
4.8	NeighborNet tree of <i>E. coli</i> O157 and <i>E. coli</i> O26 New Zealand isolates	68
4.9	NeighborNet tree of <i>E. coli</i> O157 New Zealand isolates and <i>E. coli</i> O157 reference isolates	69
4.10	NeighborNet tree of <i>E. coli</i> O26 New Zealand isolates and <i>E. coli</i> O26 reference isolates	70
4.11	NeighborNet tree of <i>E. coli</i> O157, O26 New Zealand isolates and <i>E. fergusonii</i> reference isolates	71
4.12	NeighborNet tree of <i>E. coli</i> O157 and O26 New Zealand isolates following COG analysis	72
4.13	Number of differences in proteins/domains associated with carbon metabolism among <i>E. coli</i> O157 and O26 New Zealand isolates	74
A1	Heatmap of <i>E. coli</i> O157 respiration changes on PM1 including duplicate and replicate assays	106

A2	Heatmap of <i>E. coli</i> O157 respiration changes on PM2A including duplicate and replicate assays	107
A3	Heatmap of <i>E. coli</i> O26 respiration changes on PM2A	108