Genomic Characterisation of Extra-intestinal Pathogenic *Escherichia coli* from Human Infection

A thesis submitted for the degree of

Doctor of Philosophy

by

Jessica McKinnon

Bachelor of Biotechnology (Honours)

Certificate of Authorship

I declare that this thesis by compilation is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Life Sciences at the University of Technology Sydney.

This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution.

This research is supported by an Australian Government Research Training Program.

Production Note:
Signature removed prior to publication.

Jessica McKinnon

May 2020

Peer Reviewed Publications

Published with First Authorship (* authors contributed equally)

Darling AE*, **McKinnon J***, Worden P, Santos J, Charles IG, Roy Chowdhury P and Djordjevic SP. (2014). A draft genome of *Escherichia coli* sequence type 127 strain 2009-46. *Gut Pathogens* 6(32).

McKinnon J*, Roy Chowdhury P* and Djordjevic SP. (2018). Genomic analysis of multidrug-resistant *Escherichia coli* ST58 causing urosepsis. *International Journal of Antimicrobial Agents* 52(3):430-435.

Roy Chowdhury P*, **McKinnon J***, Liu M and Djorodjevic SP. (2018). Multidrug resistant uropathogenic Escherichia coli ST405 with a novel, composite IS*26* transposon in a unique chromosomal location. *Frontiers* 8(9):3212.

McKinnon J, Roy Chowdhury P and Djordjevic SP. (2020) Molecular analysis of an IncF ColV-like plasmid lineage that carries a complex resistance locus with a trackable genetic signature. *Microb Drug Resist*.

Bogema DR*, **McKinnon J***, Liu M, Hitchick N, Miller N, Venturini C, Iredell J, Darling AE, Roy Chowdhury P, Djordjevic SP. (2020). Whole-genome analysis of extraintestinal *Escherichia coli* sequence type 73 from a single hospital over a 2 year period identified different circulating clonal groups. *Microb Genom* 6(1).

In Preparation

McKinnon J, Roy Chowdhury P and Djordjevic SP. (2019). Whole genome sequencing of 83 class 1 integron positive clinical isolates from urinary tract infection.

Co-Authorships

Melvold JA, Wyrsch ER, **McKinnon J**, Roy Chowdhury P, Charles IG and Djordjevic, SP. (2017). Identification of a novel qnrA allele, qnrA8, in environmental Shewanella algae. *Journal of Antimicrobial Chemotherapy* 72, 2949-2952.

Roy Chowdhury P, **McKinnon J**, Wyrsch E, Hammond JM, Charles IG and Djordjevic SP. (2014). Genomic interplay in bacterial communities: implications for growth promoting practices in animal husbandry. *Frontiers in Microbiology* 5.

Conference Presentations

- Molecular Dissection of Drug Resistance Regions in Clinical *Escherichia coli* Isolates from a Sydney Hospital

  New Horizons: Research and Education for Optimal Health
  Wednesday 20th November, 2013, Poster 3

- Comparative Genomics of Sequential *Escherichia coli* Strains 2009-49 and 2009-52: A Micro-Evolutionary Event in Real-Time?

  BacPath 13: Molecular Analysis of Bacterial Pathogens
  Tuesday 29th September, 2015, Poster Session 3

- Identification of a Unique Antibiotic Resistance and Virulence Locus from a Human Blood-Stream Infection

  International Conference on Plasmid Biology
  Poster presentation, 2016

- A uropathogenic *Escherichia coli* ST405 with a unique complex resistance region

  Mobile Genetic Element: Transposable Elements, Integrative and Plasmids
  Oral presentation number 8, 2017

Abstract

Tracking resistance genes based on specific structural features of class 1 integrons is an integral part of clinical epidemiology. A class 1 integron is a gene capture and expression unit, most frequently found in Gram-negative bacteria. They are known to be one of the greatest contributors to the spread of multi-drug resistance genes within clinical isolates.

Recent studies indicate that certain insertion elements target specific regions of class 1 integrons creating unique structures. Consequently, the resistance pool in such integrons goes unnoticed in standard molecular screening methodologies, although they are equally efficient in expressing and disseminating resistance genes. Examples of such class 1 integrons with atypical structures, now realised to have widely disseminated within human and animal *E. coli* populations worldwide, were generated by the insertion of a genetic element known as IS*26*. This project aimed to determine the presence of such structures, as well as characterise the diverse complex resistance loci (CRL) and virulence cargo of extraintestinal pathogenic *E. coli* samples in a specific Sydney hospital. This was achieved by identifying the virulence-associated and antimicrobial resistance gene pool harboured by these strains with a particular focus on insertion or deleterious molecular events, discerning the locations of CRL within genome or plasmid where possible and observing trends within clonal groups.

Targeted PCR, Sanger sequencing and next generation short and long read genome sequencing techniques in conjunction with bioinformatic analyses were used to characterise clinical *E. coli* samples from Sydney Adventist Hospital isolated between 2009 and 2011.

Our data suggests that antibiotic resistance is readily transferring between host populations via lateral gene transfer of mobile elements in Sydney through the observation of unique molecular signatures. Clonal groups were identified in the cohort which share virulence and antimicrobial resistance traits, in some cases at seemingly differing stages of evolution. Large scale studies such as this provide insight to the mechanisms and forces driving the dissemination of antimicrobial resistance.

Contents

Certificate of Authorship 2

Peer Reviewed Publications 3

Conference Presentations 4

Abstract 5

Abbreviations 12

Chapter 1: Background 14

1.1 Antibiotics and the Evolution of Resistance 15

1.2 Lateral Gene Transfer 18

1.2.1 The Concept 18

1.2.2 Homologous Recombination 20

1.2.3 Site-specific Recombination 21

1.3 Established *Escherichia coli* Lineages 21

1.3.1 Lateral Gene Transfer and Drug Resistance in *E. coli* 24

1.4 Mobile Genetic Elements 25

1.4.1 Plasmids 26

1.4.2 Transposons and Insertion Elements 29

1.4.3 Integrons 30

1.4.4 Genomic Islands 33

1.4.5 Complex Resistance Loci 33

1.5 Class 1 Integrons: Evolution and Zoonosis 37

1.6 Experimental Rationale and Approach 40

1.6.1 List of Methodologies and Software 41

Chapter 2: A draft genome of *Escherichia coli* sequence type 127 strain 2009-46 44

Abstract 46

Background 46

Results 46

Conclusion 46

Background 46

Methods 47

Genome sequencing 47

Assembly and annotation 48

Quality assurance 48

Initial findings 48

Antibiotic resistance profile 50

Future directions 51

Availability of supporting data 51

Competing interests 51

Authors' contributions 51

Acknowledgements 51

References 55

Chapter 3: Multidrug resistant uropathogenic *Escherichia coli* ST405 with a novel, composite IS*26* transposon in a unique chromosomal location 57

Abstract 59

1. Introduction 60

2. Materials and Methods 61

2. 1 Strains, isolation and culture conditions 61

2.2 DNA purification 61

2.3 Fosmid library construction, screening and PCR conditions 62

2.4 Amplicon and whole genomes sequencing 62

2.5 Bioinformatics 63

3. Results 63

3.1 Genomic analysis of strains 2009-27 and 2009-30 63

3.2 Plasmid analysis 64

3.3 Characterisation of the CRL in strain 2009-27 65

3.4 Prevalence of Tn*6242* 66

3.5 Comparative phylogenomics of ST405 strains 66

3.6 Evolution of Tn*6242* 67

3.7 Generation of laterally mobile translocatable units from Tn*6242* 68

4. Discussion 68

Acknowledgements 70

Author contributions 70

Conflict of Interest 70

Ethics statement 70

References 75

Chapter 4: Whole genome analysis of ExPEC ST73 from a single hospital over a 2-year period identified different circulating clonal groups 80

Abstract 82

Data Summary 82

Impact Statement 83

Introduction 84

Methods 85

Isolate source and culture conditions 85

Nucleic acid purification and whole genome sequencing 86

Genome assembly and gene presence 86

Archived sequence read selection 86
S1-PFGE analysis 87
SNP based phylogenetic analyses 87
Results 88
Assembly information and statistics 88
Public read high-throughput sequencing analysis 88
Virulence profiles of Sydney strains 89
Antibiotic resistance 90
Structure of class 1 integrons in ST73 strains from Sydney 91
Discussion 93
Author statements 95
Funding information 95
Acknowledgements 96
Ethical statement 96
Conflicts of interest 96
Abbreviations 96
References 103
Chapter 5: Genomic analysis of multidrug-resistant *Escherichia coli* ST58 causing urosepsis. 108
Abstract 110
1. Introduction 111
Materials and Methods 112
2.1. Bacterial strains 112
2.2. Calibrated dichotomous sensitivity (CDS) testing 112
2.3. Whole-genome sequencing and analyses 113
2.4. Single nucleotide polymorphisms (SNPs) 113

2.5. Nucleotide sequence accession nos. 114
3. Results and Discussion 114
3.1. Whole-genome sequencing and comparative genomic analyses 114
3.2. Virulence-associated genes (VAGs) in ST58 strains 115
3.3. Characterisation of the complex resistance locus in the Sydney ST58 strains 115
3.4. pSDJ2009-52F carries the Tn*1721*/Tn*21* hybrid transposon 116
4. Conclusion 117
Acknowledgments 118
Funding 118
Competing interests 118
Ethical approval 118
References 123
Chapter 6: Molecular analysis of an IncF ColV-like plasmid lineage that carries a complex resistance locus with a trackable genetic signature 125
Abstract 127
Introduction: 127
Methods: 129
Sequencing of pSDJ2009-52F 129
Single Nucleotide Polymorphism (SNP) Analysis 129
BEAST2 Analysis 129
Results: 130
Plasmid Information 130
Virulence and Antibiotic Resistance Gene Carriage 130
Comparative SNP Analyses 131
BEAST2 Analysis 132

Molecular Signatures 132
Discussion 133
CRediT Statement 134
Funding Statement 135
References 141
Chapter 7: General Discussion 144
7.1 Dominant Lineages in Clinical ExPEC and Trends in Antibiotic Resistance 145
7.2 IS*26*-mediated Insertion Signatures and Zoonosis 150
7.3.1 Limitations 156
References 158

Abbreviations

AIEC	adherent-invasive *E. coli*
APEC	avian pathogenic *E. coli*
bp	base pair
CDS	calibrated dichotomous sensitivity
CRL	complex resistance loci/locus
CS	conserved segment
DAEC	diffusely adhering *E. coli*
EAEC	enteroaggregative *E. coli*
EHEC	enterohemorrhagic *E. coli*
EIEC	enteroinvasive *E. coli*
EPEC	enteropathogenic *E. coli*
ESBL	extended spectrum β-lactamase
ETEC	enterotoxigenic *E. coli*
ExPEC	extra-intestinal pathogenic *E. coli*
GI	genomic island
GIT	gastro-intestinal tract
HC	hemorrhagic colitis
HUS	haemolytic uraemic syndrome
Inc	incompatibility
IPEC	intestinal pathogenic *E. coli*
IS	insertion sequence
kb	kilobase
LB	Luria-Bertani/lysogeny broth
LGT	lateral gene transfer
MDR	multi-drug resistant
MGE	mobile genetic element
MLST	multi-locus sequence typing

NMEC	neonatal meningitis *E. coli*
nt	nucleotide
PCR	polymerase chain reaction
PFGE	pulsed field gel electrophoresis
SMAC	sorbitol-MacConkey agar
SMRT	single molecule real time
SNP	single nucleotide polymorphism
SRA	sequence read archive
ST	sequence type
TU	translocatable unit
UPEC	uropathogenic *E. coli*
UTI	urinary tract infection
VAG	virulence-associated gene