# Genomic Epidemiology of *Escherichia coli* in Human Blood-Stream Infections

A thesis submitted for the degree of

Doctor of Philosophy

by

Priyanka Shirish Hastak

MSc. (Biotechnology)

# Certificate of Original Authorship

I, Priyanka Shirish Hastak, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the Faculty of Science, School of Life Sciences at the University of Technology Sydney is wholly my own work unless otherwise referenced or acknowledged.

I certify that all information sources and literature used are indicated in this thesis.

This document has not been submitted to any other academic institution.

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

Production Note:
Signature removed prior to publication.

**Priyanka Shirish Hastak**

**Date: 3rd March 2021**

# Acknowledgement:

I would first like to thank my supervisors, Associate Professor Garry Myers, Dr. Piklu Roy Chowdhury and Professor Steven Djordjevic for their guidance, training and support during my candidature. I would like to specially thank them for giving me the opportunity to work on such an excellent research project. It has been an honour to work with them and learn from them in the last 3 and half years. Also, I would like to thank the AusGem (Australian centre for genomic epidemiology microbiology) and Concord Repatriate General Hospital, Sydney for providing me with a great collection of *Escherichia coli* strains that I had the privilege to work with during my candidature.

I would like to thank the University of Technology Sydney and the i3 Institute for supporting me financially with the International Research Scholarship and the i3 Institute Research Scholarship, as well as the Australian Government Research Training Program.

Thank you to all the members of Djordjevic lab, past and present. I would particularly like to thank Ethan Wyrsch, Cameron Reid and Max Cummins for showing me how to work with a number of bioinformatics tools. Also thank you Veronica Jarocki, Dmitriy Li, Tiziana Zingali and Elisa Massella, for being great friends and colleagues that have made my candidature enjoyable and fun.

Thank you to the academic staff and technical staff at the i3 Institute, that includes Professor Liz Harry for her support and encouragement. I am glad and proud to be a part of the i3 Institute that has a friendly environment for young scientists to work on cutting edge research in infectious disease.

Lastly but most importantly, I would like to thank my family and friends. My loving and amazing parents for encouraging and believing in me, always. My sister, Radhika who has supported me spiritually throughout my thesis. Thank you to my aunt and uncle in Sydney for being my home, away from home. Thank you to all my friends back in India and here in Sydney for being my support team during my candidature. Love you all.

Statement:

This thesis is by compilation. The first and second result chapters are publications. The publications are listed in the section below. The last result chapter will be submitted for publishing shortly. The figures, tables and supplementary data is provided at the end of each chapter or as separate files in folder submitted with the thesis.

# Publications arising from this thesis

**1. Hastak, P**.; Cummins, M.L.; Gottlieb, T.; Cheong, E.; Merlino, J.; Myers, G.S.A.; Djordjevic, S.P.; Chowdhury, P.R.Genomic profiling of Escherichia coli isolates from bacteraemia patients: A 3-year cohort study of isolates collected at a Sydney teaching hospital. Microb. Genom. 2020, 6,e000371. https://doi.org/10.1099/mgen.0.000371

CRediT Author statement:

**Priyanka Hastak:** Methodology, Writing- Original Draft, Sample Collection, Writing- Reviewing and Editing, Data Curation, Visualisation. **Piklu Roy Chowdhury**: Conceptualization, Methodology, Writing- Reviewing and Editing, Validation, Data Curation, Project administration, Supervision, Funding acquisition, Resources. **Garry Myers**: Supervision, Validation, Time Management, Methodology, Project Administration, Writing- Review and Editing. **Steven Djordjevic:** Supervision, Validation, Time Management, Writing- Review and Editing, Funding acquisiton. **Tom Gottleib:** Project administration, Fund acquistion, Resources. **John Merlino**: Methodology, Resources. **Elaine Cheong**: Project administration, Resources. **Max Cummins**: Methodology, Software, Visualisation.

**2. Hastak, P**., Fourment, M., Darling, A., Gottleib., T, Cheong, E., Merlino, J., Myers, G., Djordjevic, S.P., and Roy Chowdhury, P. *Escherichia coli* ST8196 is a novel, locally evolved, and extensively drug resistant pathogenic lineage within the ST131 clonal complex. Emerg Microbes Infect. 2020 Dec;9(1):1780-1792. doi: 10.1080/22221751.2020.1797541.
CRediT Author statement:

**Priyanka Hastak:** Methodology, Writing- Original Draft, Writing- Reviewing and Editing, Data Curation, Visualisation. **Piklu Roy Chowdhury**: Conceptualization, Methodology, Writing- Reviewing and Editing, Validation, Data Curation, Project administration, Supervision, Funding acquisition, Resources. **Garry Myers**: Supervision, Validation, Time Management, Methodology, Project Administration, Writing- Review and Editing. **Steven Djordjevic:** Supervision, Validation, Time Management, Writing- Review and Editing, Funding acquisiton. **Aaron Darling**: Conceptualization, Validation, Methodology, Software.

**Mathieu Fourment**: Methodology, Software, Visualization. **Tom Gottleib:** Project administration, Resources. **John Merlino**: Methodology, Resources. **Elaine Cheong**: Project administration, Resources.

**Contents presented in chapter 6 will be included in a manuscript which is in preliminary draft stages.**

**Hastak, P.,** Gottlieb, T., Cheong, J., Myers, G., Djordjevic, S.P., and Roy Chowdhury, P.

Comparative genomic analysis of ST38 *Escherichia coli* from human, animal and environmental settings reported in Australia and globally.

# Table of Contents

# List of abbreviations:

| | |
|---|---|
| AMR | Antimicrobial Resistance |
| ARG | Antimicrobial Resistance Gene |
| MGE | Mobile Genetic Element |
| IS | Insertion Sequence |
| DNA | Deoxyribonucleic acid |
| cc | Clonal complex |
| bp | Base pair |
| CRL | Complex Resistance Locus |
| BSI | Blood-stream infections |
| ESBL | Extended-spectrum beta-lactamase |
| ETEC | Enterotoxigenic *Escherichia coli* |
| ExPEC | Extraintestinal Pathogenic *Escherichia coli* |
| UTI | Urinary Tract Infection |
| LCBs | Locally Colinear Blocks |
| LGT | Lateral Gene Transfer |
| MDR | Multiple-Drug (Antibiotic) Resistance |
| MLST | Multi-locus Sequence Typing |
| ST | Sequence Type |
| SNP | Single Nucleotide Polymorphism |
| WGS | Whole Genome Sequence |
| HO | Hospital Onset |
| CO | Community Onset |
| TE | Transposable Elements |

| | |
|---|---|
| GI | Genomic Island |
| MLST | Multi-Locus Sequence Typing |
| CS | Conserved Segment |
| ICU | Intensive Care Unit |
| VAG | Virulence Associated Genes |

# Abstract:

Blood-stream infections (BSI) are associated with high mortality and morbidity world-wide. The most common aetiological agent of these infections is *Escherichia coli*. The rise of multi-drug resistant (MDR) *E. coli* is a major concern in clinical medicine and needs to be monitored for implementation of infection control strategies. The misuse of antimicrobials in clinical treatment, and as growth promotors in food-producing animals, is a key component of the rapid evolution of MDR *E. coli*. MDR organisms spread antimicrobial resistance and virulence factors by lateral gene transfer via mobile genetic elements (MGEs). Despite this, there is limited knowledge of the origins and underlying mechanisms of BSI MDR *E. coli* infections in Australia.

We conducted genomic epidemiological analyses of *E. coli* from human blood-cultures collected from a Sydney teaching hospital. The collection was dominated by *E. coli* causing community onset BSI, with clones carrying a plethora of virulence and antimicrobial resistance genes. Clinical class 1 integrons associated with IS*26* were identified in a number of sequence types (ST), indicating that they are important drivers of evolution and spread of clinically important antimicrobial resistance genes. We identified IncFII-IncFIB plasmids in the majority of our collection. A number of STs harboured ColV like IncFII plasmids, carry a number of important virulence traits, were also identified. We observed a novel ST, ST8196, that clustered with globally disseminated *E. coli* ST131 isolates. We also identified an emerging ST, ST38, that exhibited resistance to a broad range of clinical beta-lactamases that have been identified in a number of pandemic *E. coli* isolates.

Our study provides evidence that *E. coli* BSIs in Australia have a reservoir of antimicrobial resistance and virulence determinants that are potentially circulating not only in the community and hospital settings but across agricultural settings. These genes are not only present to clinical settings but can circulate within different ecosystems potentially via plasmids or other MGEs. Our findings also reveal that there are specific antimicrobial resistance genes such as ESBLs that are found in dominant *E. coli* clones in both humans and food producing animals. This is likely to occur due the overuse of antimicrobials in human and animal settings that is driving the increase in antimicrobial resistance among bacteria that cause diseases. It is therefore

essential to rigorously monitor these infections to help manage the global problem of antimicrobial resistance, and to reduce or prevent disease outbreaks.