UTS UNIVERSITY OF TECHNOLOGY SYDNEY

Phylogenomic analysis of *Escherichia coli* from dogs with clinical symptoms of uropathogenic disease: An Australian study

by Paarthiphan Elankumaran, BAppSc(Hons)

Thesis submitted in fulfilment of the requirements for
the degree of
Doctor of Philosophy

under the supervision of Steven Djordjevic and Cameron Reid

Australian Institute for Microbiology and Infection
School of Life Sciences
Faculty of Science
University of Technology Sydney

October 2022

Certificate of Original Authorship

I, Paarthiphan Elankumaran declare that this thesis, is 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. This thesis is wholly my own work unless otherwise referenced 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 the Australian Government Research Training Program.

Signature: Production Note:
Signature removed prior to publication.

Date: 02.10.2022

Acknowledgement

First, I would like to thank my supervisor Steven Djordjevic for providing me with this excellent opportunity to conduct my PhD research at AIMI, UTS. I am highly indebted for your mentorship and guidance over the past years without which I could not have produced this thesis. It has been an honour to work with you and learn from you.

I take great pride in thanking my co-supervisor Cameron Reid who has been the real guide in my years long journey of PhD research. I highly appreciate your effort in training me with the bioinformatics pipelines as well as the laboratory techniques. Your constant guidance, help and motivation were critical for me in completing my research project.

I would also like to thank all the other members of the Djordjevic research group especially Piklu, Max, Veronica and Ethan for helping me with my laboratory work and providing me with progress reviews as members of my internal review panel.

I also like to extend my gratitude to our collaborators from the University of Melbourne, Glenn and Marc for providing me with the primary collection of *E. coli* strains for my research project.

It is also a pleasure for me to thank Australian Center for Genomic Epidemiological Microbiology (AusGEM) for supporting me financially by providing me with the PhD scholarship stipend.

Finally, I would like to extend my love and gratitude for members of my family especially my late grandparents and siblings without whom I wouldn't be the person I am today.

Statement

This is a thesis by compilation. The Chapters 4, 5 and 6 constitute results Chapters. Chapters 4 and 6 are published in peer reviewed journals while Chapter 5 is currently submitted and under review in a peer reviewed journal and will be published soon. The published manuscripts are attached as individual Chapters as they appeared in the journals.

List of Publications

First author manuscripts

- **Genomic and Temporal Trends in Canine ExPEC Reflect Those of Human ExPEC**

  Paarthiphan Elankumaran[a], Max L. Cummins[a], Glenn F. Browning[b], Marc S. Marenda[b], Cameron J. Reid[a], Steven P.Djordjevic[a]

  [a]Australian Institute for Microbiology and Infection, School of Life Sciences, Faculty of Science, University of Technology Sydney, Ultimo, New South Wales, Australia

  [b]Asia-Pacific Centre for Animal Health, Department of Veterinary Biosciences, Melbourne Veterinary School, Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville and Werribee, Victoria, Australia

  Microbiology Spectrum; Vol. 10, No. 3; 8 June 2022.

  https://journals.asm.org/doi/10.1128/spectrum.01291-22

  (As Chapter 4; Published)

- **Identification of genes influencing the evolution of *Escherichia coli* ST372 in dogs and humans**

  Paarthiphan Elankumuran[1], Glenn F. Browning[2], Marc S. Marenda[2], Amanda Kidsley[3], Marwan Osman[4,5], Marisa Haenni[6], James R. Johnson[7], Darren J. Trott[3], Cameron J. Reid[1], Steven P. Djordjevic[1]

  [1]Australian Institute for Microbiology and Infection, School of Life Sciences, Faculty of Science, University of Technology Sydney, Ultimo, NSW, Australia

[2]Asia-Pacific Centre for Animal Health, Department of Veterinary Biosciences, Melbourne Veterinary School, Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville and Werribee, Victoria, Australia

[3]Australian Centre for Antimicrobial Resistance Ecology, School of Animal and Veterinary Sciences, University of Adelaide, Roseworthy, Australia

[4]Laboratoire Microbiologie Santé et Environnement, Doctoral School of Sciences and Technology, Faculty of Public Health, Lebanese University, Tripoli, Lebanon

[5]Department of Public and Ecosystem Health, College of Veterinary Medicine, Cornell University, Ithaca, NY, USA

[6]ANSES, Université de Lyon, Unité Antibiorésistance et Virulence Bactériennes, Lyon, France

[7]Minneapolis VA Medical Center, Minneapolis, MN, USA

(As Chapter 5; Submitted)

- **Close genetic linkage between human and companion animal extraintestinal pathogenic *Escherichia coli* ST127**

Paarthiphan Elankumaran[a], Glenn F. Browning[b], Marc S. Marenda[b], Cameron J. Reid[a], Steven P. Djordjevic[a]

[a]iThree Institute, School of Life Sciences, Faculty of Science, University of Technology Sydney, Ultimo, NSW, Australia

[b]Asia-Pacific Centre for Animal Health, Department of Veterinary Biosciences, Melbourne Veterinary School, Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville and Werribee, Victoria, Australia

Current Research in Microbial Sciences; Vol. 3; 23 January 2022.

https://www.sciencedirect.com/science/article/pii/S2666517422000037?via%3Dihub

(As chapter 6; Published)

Co-author manuscripts

- **Genomic comparisons of *Escherichia coli* ST131 from Australia**

Dmitriy Li[1], Ethan R. Wyrsch[1], Paarthiphan Elankumaran[1], Monika Dolejska[2,3,4,5], Marc S. Marenda[6], Glenn F. Browning[6], Rhys N. Bushell[6], Jessica McKinnon[1], Piklu Roy Chowdhury[1], Nola Hitchick[7], Natalie Miller[7], Erica Donner[8], Barbara Drigo[8], Dave Baker[9], Ian G. Charles[9], Timothy Kudinha[10], Veronica M. Jarocki[1], * and Steven Philip Djordjevic[1], *

[1]iThree Institute, University of Technology Sydney, Ultimo, NSW, Australia

[2]CEITEC VETUNI, University of Veterinary Sciences Brno, Brno, Czech Republic

[3]Department of Biology and Wildlife Disease, Faculty of Veterinary Hygiene and Ecology, University of Veterinary Sciences Brno, Czech Republic

[4]Biomedical Center, Charles University, Czech Republic

[5]Department of Clinical Microbiology and Immunology, Institute of Laboratory Medicine, The University Hospital Brno, Brno, Czech Republic

[6]Department of Veterinary Biosciences, Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Victoria, Australia

[7]San Pathology, Sydney Adventist Hospital, Wahroonga, NSW 2076, Australia

[8]Future Industries Institute, University of South Australia, Adelaide, South Australia, Australia

[9]Quadram Institute, Norwich, UK

[10]Central West Pathology Laboratory, Charles Sturt University, Orange, NSW, 2800, Australia

Microbial Genomics; Vol. 7, Issue 12; 15 December 2021.

https://www.microbiologyresearch.org/content/journal/mgen/10.1099/mgen.0.00072

(Published)

- **Dominance of pandemic extraintestinal pathogenic *Escherichia coli* sequence types ST73, ST95, ST127 and ST131 in historic human urine isolates: a genomic analysis of antimicrobial resistance and virulence linked to F plasmids.**

Dmitriy Li[1,2], Paarthiphan Elankumaran[1,2], Timothy Kudinha[3], Amanda Kidsley[4], Sam Abraham[5], Darren Trott[4], Veronica M. Jarocki[1,2], Steven P. Djordjevic[1,2]

[1]Australian Institute for Microbiology & Infection, University of Technology Sydney, Ultimo, NSW, Australia

[2]Australian Centre for Genomic Epidemiological Microbiology, University of Technology Sydney, NSW, Australia

[3]Central West Pathology Laboratory, Charles Sturt University, Orange, NSW, Australia

[4]School of Animal and Veterinary Science, The University of Adelaide, Adelaide, South Australia, Australia

[5]Antimicrobial Resistance and Infectious Disease Laboratory, Murdoch University, Perth, South Australia, Australia

(Submitted)

- **Phylogenomic analysis of a global cohort of *Escherichia coli* ST38: evidence of interspecies and environmental transmission**

Piklu Roy Chowdhury[1,2], Priyanka Hastak[1,2], Ethan Wyrsch[1], Dmitriy Li[1], Paarthiphan Elankumaran[1], Monika Dolejska[3], Glen F. Browning[4], Mark Marenda[4], Thomas Gottlieb[5,6], Elaine Cheong[5,6], John Merlino[5,6], Garry S. A. Myers[1] and Steven P. Djordjevic[1,2]

[1]The Ithree Institute, University of Technology Sydney, City Campus, Ultimo, NSW 2007, Australia.

[2]Australian Centre for Genomic Epidemiological Microbiology, University of Technology Sydney, PO Box 123, Broadway, NSW 2007, Australia

[3]Department of Biology and Wildlife Disease, Faculty of Veterinary Hygiene and Ecology, University of Veterinary Sciences Brno, Czech Republic

[4]Asia-Pacific Centre for Animal Health, Department of Veterinary Biosciences, Melbourne Veterinary School, Faculty of Veterinary and Agricultural Sciences, University of Melbourne, Parkville and Werribee, Victoria, Australia

[5]Department of Microbiology and Infectious Diseases, Concord Hospital and NSW Health Pathology, Hospital Road, Concord 2139, NSW, Australia.

[6]Faculty of Medicine, University of Sydney, NSW Australia.

(Submitted)

Table of Contents

Abbreviations

µM - Micrometre

AIEC - Adherent Invasive *E. coli*

APVMA - Australian Pesticides and Veterinary Medicines Authority

ARG - Antibiotic Resistant Gene

bp - Base pair

BSI - Blood Stream Infections

CRISPR - Clustered Regularly Interspaced Short Palindromic Repeats

DDBJ - DNA Data Bank of Japan

DNA - Deoxyribosonucleic Acid

E. coli - *Escherichia coli*

EAEC - Enteroaggregative *E. coli*

EBI - European Bioinformatics Institute

EHEC - Enterohemorrhagic *E. coli*

EIEC - Enteroinvasive *E. coli*

EPEC - Enteropathogenic *E. coli*

EPS - Extracellular Polymeric Substances

ESBL - Extended Spectrum Beta Lactamase

ETEC - Enterotoxigenic *E. coli*

ExPEC - Extraintestinal Pathogenic *E. coli*

GTA - Gene Transfer Agent

GWAS - Genome Wide Association Study

HGT - Horizontal Gene Transfer

IPEC - Intraintestinal Pathogenic *E. coli*

IS - Insertion Sequence

ITR - Inverted Terminal Repeat

LB - Luria Broth

LINE - Long Interspaced Nuclear Elements

LTR - Long Terminal Repeat

MDR - Multiple Drug Resistance

MGE - Mobile Genetic Element

MLST - Multi Locus Sequence Typing

MRSA - Methicillin resistant *Staphylococcus aureus*

MRSP - Methicillin resistant *Staphylococcus pseudintermedius*

MVC - Melbourne Veterinary Collection

NCBI - National Center for Biotechnology Information

PAI - Pathogenicity Associated Island

rMLST - Ribosomal MLST

RNA - Ribonucleic Acid

RST - Replicon Sequence Type

SINE - Short Interspaced Nuclear Element

SNP - Single Nucleotide Polymorphism

ST - Sequence Type

TE - Transposable Element

UN - United Nations

UPEC - Uropathogenic *E. coli*

UTI - Urinary Tract Infections

VAG - Virulence Associated Gene

wgMLST - Whole Genome MLST

WGS - Whole Genome Sequencing

WHO - World Health Organization

Abstract

Australia is one of the countries with the highest number of dog ownership in the world with about 40% households owning at least one dog. Most dog owners consider their dog to be part of their family and often report high levels of physical contact with them in addition to sharing food from the same household. However, there is a wide array of pathogens which could be transmitted from dogs to humans and vice versa that have the potential to cause serious infections in both. Extraintestinal pathogenic *E. coli* (ExPEC) is the leading cause of urinary tract infections (UTIs) in dogs and is the most common gram-negative bacterial pathogen in humans. Hence, ExPEC represent a significant burden in healthcare and veterinary settings. Companion animals and humans are known to share ExPEC but the extent of sequence types (STs) that cause extraintestinal diseases in dogs is not well understood.

Whole genome sequencing (WGS) based analyses of *Escherichia coli* of canine origin from Australia are very limited and most of the past research were based on molecular methodologies such as polymerase chain reaction (PCR). Here, we carried out the largest WGS based phylogenomic analyses of *E. coli* isolated from dogs from Australia. The primary study collection named "MVC" comprised 377 *E. coli* isolates originating from dogs with clinical symptoms of UTI collected over a 11-year period from 2007 to 2017 by Clinical Microbiology Laboratory of the Melbourne Veterinary School, University of Melbourne, Australia. Most of the isolates originated from the urinary tract (219; 58.1%) while significant numbers of isolates from the general infections (72; 19.1%), soft tissues (34; 9%) and gastrointestinal tract (51; 13.5%) were included in the collection.

Analysis of whole genome sequences of 377 canine *E. coli* isolates identified a total of 53 STs with 18 of them represented by at least five sequences. The five most predominant STs were ST372 (69, 18.3%), ST73 (31, 8.2%), ST127 (22, 5.8%), ST80 (19, 5.0%) and ST58 (14, 3.7%). Apart from ST372, all of these are prominent human ExPEC STs. Other common ExPEC STs identified included ST12, ST131, ST95, ST141, ST963, ST1193, ST88 and ST38. Phylogroup B2 was the most dominant phylogroup (225, 59.7%) followed by the phylogroups B1 (51, 13.5%), D (30, 8.0%), and A (21, 5.6%). Carriage of antibiotic resistance genes (ARGs) was low while the collection possessed an extensive array of virulence associated genes (VAGs). Important F virulence plasmids such as ColV (56, 14.9%) and pUTI89-like (29, 7,7%) plasmids were identified. A total of 51 (13.53%) isolates possessed the class 1 integron integrase gene *intI1*, a well-known marker for antibiotic resistance in gram-negative bacteria. Virulence gene profiles, antimicrobial resistance carriage, and trends in plasmid carriage for specific STs were generally reflective of those seen in humans. Many of the prominent STs were observed repetitively over an 11-year time span, indicating their persistence in the dogs in the community, which is most likely driven by household sharing of *E. coli* between humans and their pets. The case of ST372 as a dominant canine lineage observed sporadically in humans is flagged for further investigation.

ST372 are widely reported as the major *E. coli* sequence type in dogs globally. They are also a sporadic cause of extra-intestinal infections in humans. Despite this, it is unknown whether ST372 strains from dogs and humans represent shared or distinct populations. Furthermore, little is known about genomic traits that might explain the prominence of ST372 in dogs or presence in humans. To address this, a variety of bioinformatics analyses were applied to a global collection of 407 ST372 *E. coli* whole genome sequences to characterise their epidemiological features, population structure, and associated accessory genomes.

It was confirmed that dogs are the dominant host of ST372 and that clusters within the population structure exhibit distinctive O:H types. Human adapted strains possessed O45:H31 and O18:H31 serotypes and dog adapted strains possessed O83:H31 serotype. One phylogenetic cluster, 'cluster M', comprised almost half of the sequences and showed the divergence of two human-restricted clades that carried different O:H types to the remainder of the cluster. There is evidence to support transmission of ST372 between dogs and humans within different clusters of the phylogeny, including multiple acquisitions of the *pdu* propanediol utilisation operon have occurred in clusters dominated by isolates of canine source, possibly linked to diet, whereas loss of the *pdu* operon and acquisition of K antigen virulence genes characterised human-restricted lineages.

E. coli ST127, a recently emerged global pathogen noted for high virulence gene carriage, is a leading cause of UTIs and blood stream infections (BSIs). ST127 is frequently isolated from humans and companion animals; however, it is unclear if they are distinct or related populations of ST127. We performed a phylogenomic analysis of 299 *E. coli* ST127 of diverse epidemiological origin to characterise their population structure, genetic determinants of virulence, antimicrobial resistance, and repertoire of mobile genetic elements (MGEs) with a focus on plasmids.

The core gene phylogeny was divided into 13 clusters, the largest of which (BAP4) contained the majority of human and companion animal origin isolates. This dominant cluster displayed genetic differences to the remainder of the phylogeny, most notably alternative gene alleles encoding important virulence factors including lipid A, flagella, and K capsule. Furthermore, numerous close genetic linkages (<30 SNPs) between human and companion animal isolates

were observed within the cluster. Carriage of ARGs in the collection was limited, but virulence gene carriage was extensive. We found evidence of pUTI89-like virulence plasmid carriage in over a third of isolates, localised to four of the major phylogenetic clusters. Our study supports global scale repetitive transfer of *E. coli* ST127 lineages between humans and companion animals, particularly within the dominant BAP4 cluster.

The knowledge gained through this research study will form a basis for future more elaborate studies related to canine *E. coli* from Australia and around the world. It will also help in the understanding of the role of canine ExPEC and the development of strategies based on "one health" perspective of human health. Some of the limitations of this research study included restricted sampling of the primary study collection (MVC), disproportionate epidemiological stratification of the sequences sourced from public databases towards human and veterinary samples with antibiotic resistance from developed countries and the difficulty in tracking MGEs due to using short read sequences. These issues need to be rectified in future studies along with focusing on the analyses of sequences from other source niches of interest and around the world.