Comparative genomics of *Escherichia coli* causing bloodstream infections in the Hospital for Tropical Diseases, Ho Chi Minh city, Vietnam

Tu Thi Phuong Le

Wellcome Trust Sanger Institute

University of Cambridge

Darwin College

This dissertation is submitted for the degree of Master of Philosophy in Biological Science

September 2017

Declaration

I hereby declare that:

This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text and bibliography. My dissertation is not substantially the same as any that I have submitted for a degree or diploma or other qualification at any other University. I further state that no part of my dissertation has already been or is being concurrently submitted for any such degree, diploma or other qualification.

I confirmed that my dissertation does not exceed the limit of length of 20,000 words prescribed in the Special Regulations of the MPhil examination for which I am a candidate.

Acknowledgements

This thesis was possible because of the collective efforts of many people, from the patient beds to lab benches and computer clusters, trying to deconvolute one of the oldest and cheekiest bacteria that cause big issues not only in a single referral hospital in Vietnam, but also across the globe.

This has been a tough but wonderful journey. I was extremely grateful to work on this project and have seen myself grow academically, intellectually, and personally in the past year in Cambridge.

First of all, I would like to thank my main supervisor Prof. Nicholas Robert Thomson for his unbelievable patience and guidance during the research. His continuous encouragement and support during difficult times is something that I always greatly appreciated.

I also would like to thank my former supervisor Prof. Stephen Baker from Oxford University Clinical Research Unit (OUCRU) for coming up with the idea for this project during the application process to the MPhil program as it turned out to be a very interesting project.

I would like to express my gratitude to my thesis committee members including Julian Rayner, Stephen Bentley, Estee Torok, and Kate Baker, who have provided valuable feedback and criticism.

I would like to thank members of Team 216 at the Sanger Institute for their tremendous daily support, as well as for providing me with a remarkably happy and friendly place to work. In no particular order, these include Angèle Bénard, Mathew Beale and Déborah Kreyenbihler. With no prior background on working with whole genome sequencing data, I therefore could not have completed the project without the guidance and patience of Alison Mather (Cambridge University), Daryl Domman, Eva Heinz and Alex Wailan (Sanger Institute), who never hesitating to spend time teaching me when I was stuck. I also would like to thank Ngure Kagia, another MSc student, who helped me to feel comfortable using Linux and computer clusters at the very beginning of my time at Sanger. Finally, I am indebted to Matthew Dorman for giving very constructive helpful feedback during my writing process.

The blood-derived *Escherichia coli* strains from this six-year period is the effort of all Microbiology staff at Hospital for Tropical Disease (Ho Chi Minh city, Vietnam), especially Nguyen Huu Hien and Tran Thi Ngan who carried out the blood culturing, identification, antimicrobial phenotype testing and sample archiving. The clinical metadata extracted from electronic hospital record database for this study was kindly provided by Nguyen Huu Hien, Dr Nguyen Phu Huong Lan and Dr Nguyen Van Vinh Chau. The second set of blood and stool isolates, as well as its associated clinical data, came from a big study on hospital acquired infections in the Intensive Care Unit lead by Dr Duong Bich Thuy, Mr James Campell, Dr Louise Thwaites, Prof Stephen Baker and Prof Guy Thwaites. I also would like to thank Nguyen Hoang Thu Trang, To Thi Thanh Huyen and Vong Vinh Phat for assisting me during the DNA extraction stage. I also would like to thank Dr Ho Dang Trung Nghia for discussions about the context of management and treatment of BSI patients in the hospital. My work relied on enormous sequencing

data of nearly 700 bacterial isolates, therefore I am indebted to the DNA core sequencing team and also the Pathogen Informatics team at the Sanger Institute for sample management, sequencing, and providing many analysis pipelines and bioinformatics supports. Many scripts and software I used for WGS analysis in this thesis were written by Simon Harris and Andrew Page.

Apart from working, I would like to give special thanks to Benjamin Bai, Simon Scutts, other friends and colleagues at Sanger and Darwin college badminton club for keeping me sane through many fun and joyful badminton matches.

Finally, this study could not have made possible without the support of the Wellcome Trust, who generously funded not only my university fees and living expenses, but also the research costs in both institutions in two nations, Sanger (UK) and OUCRU (Vietnam). I hope my study will shed a light on how the cheeky *E. coli* cause disease in Vietnamese patients as well as acquire drug resistance over time. Future work will be required to understand how to prevent and reduce the burden of disease in Vietnamese patients.

To my parents and little sister,

for their endless and unconditional love.

Table of Contents

Declaration……………………………………………………………………………………....i

Acknowledgements………………………………………………………………………………...ii

Dedication..……………..………………………......………………………………...…………...iv

Contents……………………………………………………………………………………………...v

Figures……...………………………………………………………………………………...…..vii

Tables………………………………………………………………………………………………viii

Abstract……………………………………………………………………………………………..ix

Chapter 1: Introduction.. vii
1.1. The history, taxonomy, and characteristics of *Escherichia coli*..1
1.1.1. Introduction..1
1.1.2. Commensal *E. coli* 1
1.1.3. The *E. coli* genome..2
1.1.4. Pathogenic *E. coli*...2
1.2. *E. coli* associated with bloodstream infections (BSIs), sepsis and bacteraemia 3
1.3. Hospital for Tropical Disease (HTD), Vietnam...4
1.4. Antimicrobial resistance in *E. coli*...4
1.5. AMR *E. coli* in Vietnam..5
1.5.1 ESBL – producing *E. coli* and carbepenemase- producing *E. coli*...............................6
1.5.2 Carbapenemase-producing *E. coli* 7
1.6. Mobile genetic elements (MGEs)..7
1.7. Reservoirs of *E. coli* and transmission mechanism..8
1.8. Aims and objectives of the study 8

Chapter 2: Comparative Genomics of BSIs caused by *E. coli* in the Hospital for Tropical Diseases in Vietnam 10
2.1. Introduction...10
2.2. Methods...11
2.2.1. Strain collection 11
2.2.2 Ethics...11
2.2.3 Bacterial isolation 11
2.2.4. Metadata...12
2.2.5. Whole genome sequencing...12
2.2.6. Accessions 12
2.2.7. Genome assembly and annotation...12
2.2.8. Quality control of sequencing data..13
2.2.9. *In silico* MLST and phylogroup classification 13
2.2.10. Construction of core genome phylogeny 13
2.3. Results..14
2.3.1. Sequencing and QC 14
2.3.2. Basic patient demographic and clinical outcome of hospitalised patients with BSIs. 15
2.3.3. The genomic diversity and population structure of BSI and carriage *E. coli* from Vietnamese patients attending HTD...16
2.4. Discussion..25

Chapter 3: The accessory genome of *Escherichia coli* of patients attending HTD. 27
3.1. Introduction 27
3.2. Methods 30
3.2.1. *In silico* virulence, AMR, and plasmid replicon type gene detection 30
3.2.2. Antimicrobial phenotype testing 30
3.2.3. Statistical analysis 30
3.3. Results 31
3.3.1. The distribution of virulence factors between blood and rectal swab isolates 31
3.3.2. The distribution of virulence factors between different BAPS lineages/ phylogroups/STs 33
3.3.3. Antimicrobial resistance phenotype between blood/rectal swab isolates 41
3.3.4. Antimicrobial resistance genes distribution in blood/ rectal swab 42
3.4. Discussion 50
Chapter 4: Understanding the relationships between BSI and carriage *E. coli* isolates in patients attending HTD, Vietnam. 53
4.1. Introduction 53
4.2. Methods 54
4.2.1 Study participants 54
4.2.2. Isolate collection 54
4.2.3. Phylogenetic analysis 54
4.2.4. AMR genes, replicon types and virulence genes identification 54
4.3. Results 54
4.3.1. Patient and sample recruitment 54
4.3.2. Genomic diversity among rectal swab and blood derived *E. coli* isolates on admission to HTD ICU 55
4.3.2. Longitudinal diversity 59
4.4. Discussion 63
Chapter 5 Final discussion 65

Figures

Figure 2.1. Sequencing and QC work flow for the samples included for whole genome sequencing (WGS) in this study **14**

Figure 2.2: Comorbidity and outcome in patients with *E. coli* BSIs...**16**

Figure 2.3: The pan genome of 687 *Escherichia coli* **18**

Figure 2.4: Population structure of 665 *E. coli* isolates analysed in this study.......................**19**

Figure 2.5: The phylogenies of *E. coli* taken from HTD, including reference strains, reported alongside clinical metadata for *E. coli* causing BSI...**22**

Figure 2.6: STs distribution of blood and rectal swab isolates...**23**

Figure 2.7: A core genome tree visualised in Phandango for blood and rectal swab isolates............**24**

Figure 3.1: Volcano plot for odds ratio versus $-\log_{10}$ p-value of gene distribution between invasive and carriage isolates...**33**

Figure 3.2: Phylogeny of *E. coli* lineages and virulence genes found in each isolate. **34**

Figure 3.3: The distribution of AMR resistance between blood and carriage isolate......................**42**

Figure 3.4: Total number of resistance genes and replicon types..**43**

Figure 3.5: AMR genotypes and phenotypes across all *E. coli* isolates.................................**44**

Figure 3.6: The genetic environment of *mcr-1* on a chromosomal contig................................**45**

Figure 4.1: Phylogenetic diversity among rectal swab and blood derived *E. coli* isolates from HTD **57**

Figure 4.2: Pairwise SNP distances within and between patients based on the variant calls from the core gene alignment..**58**

Figure 4.3: Twenty longitudinal isolates and from seven patients **59**

Figure 4.4: Graph shows phylogenetic relationship between isolates in patient P3 (03-283)..........**63**

Tables

Table 2.1: Clinical characteristic of *E. coli* BSIs patients……………………..………………………15

Table 3.1: Common virulence factors identified in ExPEC…………………………………………...28

Table 3.2: The percentage of isolates carrying specific virulence genes in the 12 most dominant sequence types of *E. coli* found in this study……………………………………….…………..36

Table 3.3: *bla* CTX-M distribution among different STs ………...……………………………...46

Table 3.4: Plasmid replicon type identified in *E. coli* collection………..…………………………...47

Abstract

Escherichia coli (*E. coli*) is a versatile bacterium, with the capability to act not only as a commensal coloniser but also as a pathogen that can cause invasive disease. Currently, *E. coli* is the leading cause of bloodstream infections in both developed and developing countries, accounting for 25-30% of bacteraemia cases globally. *E. coli* bacteraemia is further exacerbated by the emergence of antimicrobial resistance, particularly extended spectrum β-lactamases (ESBLs) and carbapenemases in Gram-negative bacteria. Understanding the nature and diversity of *E. coli* causing bloodstream infections is crucial for the enhancement of infection control measures, to minimise the further emergence of antimicrobial resistant isolates, and for the reduction of morbidity and mortality.

We used whole-genome sequencing to analyse the population structure of 506 invasive and 159 carriage *E. coli* isolates collected at The Hospital for Tropical Diseases, Vietnam during 2010-2015, and to look for genetic signatures that differentiate them from one another. We found substantial diversity therein, both in the total number of sequence types present, and in the number of resistant genes carried. Among blood isolates, ST131, ST95, ST69, ST1193 and ST73 were dominant STs, while from the rectal swab (carriage) isolates, ST131, ST1193 and ST648 were most prevalent. All of these STs from blood and rectal swab samples remained dominant over the entire study period, the exception being ST1193. Interestingly ST1193 was not present in Vietnam prior to 2011, but once introduced, it quickly emerged and replaced other more drug sensitive *E. coli* clones. From longitudinal sampling and paired *E. coli* isolated from blood and rectal swab from the same patients, we demonstrated that majority of *E. coli* infections (57%) were acquired from the patients' own gut microbiota. This study shows that several genetic factors, including genes that mediate adhesion (fimbriae and pili), iron acquisition (siderophores), immune evasion (capsule synthesis) and toxin elaboration (haemolysin and cytotoxic necrotizing factor 1) are significantly more common in invasive *E. coli* than in carriage strains. We were also able to begin to explain anomalies in the patterns of antimicrobial resistance of blood and rectal swab *E. coli* isolates using the population structure we defined here.

Taken together, we have generated a genetic framework for future studies focusing on *E. coli* BSI at HTD. These data also demonstrate that a combination of both virulence genes and antimicrobial resistance genes are essential for the success of certain lineages in causing invasive disease in Vietnamese patients. Perhaps more importantly we also show that patients who develop bacteraemia in hospital did so largely through infection with isolates already present in their intestinal tract and so were predominantly community acquired invasive infections rather than hospital acquired infections, contrary to our original hypothesis.

Chapter 1: Introduction

1.1. The history, taxonomy, and characteristics of *Escherichia coli*

1.1.1. Introduction

Escherichia coli (*E. coli*) is a bacterial genus and species, belonging to the family *Enterobacteriaceae*. Originally called "Bacterium coli commune" in 1885, it was subsequently renamed *E. coli* after Theodor Escherich, a German paediatric physician, who investigated and characterised the bacteria colonising infant guts, as well as their role in infant digestion [1].

We now know that *E. coli* is a ubiquitous member of microbiota in the lower intestinal tract of reptiles, humans, as well as most other warm-blooded animals [2]. A typical human gut can carry up to 10^9 *E. coli* cells [3]. This compares to the 3.0 x 10^{13} human cells making up the body and the estimated 3.8 x 10^{13} number of bacterial cells that colonise the human body [4]. In the context of the digestive tract, *E. coli* represents about 0.1 - 5% of the gut microbiota; although it is ubiquitous, it represents only a minor component of the gut bacterial population [5, 6]. Since *E. coli* grows rapidly, has relatively simple nutritional requirements, and is highly genetically amenable, it quickly became the main Gram-negative prokaryotic model system [2].

E. coli is not only a benign resident of the human gut, but can also contribute positively to human health. For example, *E. coli* produces Vitamin K and serves as a barrier layer in the gut, helping to outcompete and to outgrow pathogenic bacteria which attempt to colonise that niche [7]. *E. coli* is also known to inhabit niches outside of the warm-blooded gut, including water, sediments and soils [8, 9], it is not restricted to colonising humans or warm-blooded animals. Moreover, along with this ability to colonise a diverse set of niches, *E. coli* is also a diverse organism genetically. The ability of *E. coli* to survive and to colonise many niches, both biological and abiotic environments, as well as to behave as both a commensal and pathogenic bacterium, makes this species an excellent candidate for the study of bacterial evolution and adaptation [10].

1.1.2. Commensal *E. coli*

The population structure of commensal *E. coli* has been reviewed extensively by Tenaillon *et al* [3]. Along with others gut microbiome, *E. coli* is one of the first bacterial species to colonized infant's gastrointestinal tract after natural birth delivery and might originate from the maternal faecal flora [11]. The population of *E. coli* in the human gut is diverse: at any given time point, each person carries one predominant *E. coli* strain/genotypes which accounts for more than 50% of the total *E. coli* in the gut, and simultaneously harbours several other strains/genotypes at different percentages [12]. Longitudinal sampling of stool isolates revealed that there are frequently more than two genotypes of *E. coli* colonising the gut, defined as resident (stay for years) or transient colonisers (only present for a few

days or weeks). *E. coli* are classified into four major phylogroups A, B1, B2 and D [13] and populations of these phylogroups are spatiotemporally diverse. Different human populations are dominated by different phylogroups, and dominant phylogroups in a population may vary over time [14]. For example, the guts of native people in French Guyana are composed of more "commensal" *E. coli* phylogroups A and D than people from an urban city like Paris [15]. There was also a shift from phylogroup A to B2 in Parisians between the 1980s and 2000s. Such differences are explained mainly by diet and hygiene [16, 17].

1.1.3. The *E. coli* genome

Bacterial genomes can be sub-divided into core and accessory regions. Core genes are families of genes that are found in all members of a particular species, whereas accessory genes are present in some, but not all, genomes of that species. Together, the core and accessory regions of genes constitute the pan genome. From comparative genomic studies that have analysed the core and pan genomes of *E. coli*, we know that the core genome of *E. coli* consists of nearly 2,000 genes. However, the size of the predicted core depends heavily on the number of genomes and the cutoffs used to define orthologous genes [18, 19].

The pan genome defined by comparative genomic analysis of the first 20 fully sequenced *E. coli* genomes, which represent all *E. coli* pathotypes, contains around 18,000 genes [18]. The number of genes in the *E. coli* pan genome increases linearly with the number of newly added *E. coli* genomes [20]. Compared to more host restricted prokaryotes such as *Chlamydia trachomatis* or *Corynebacterium pseudotuberculosis*, *E. coli is* thought to have an "open" pan genome, whereby accessory genes might provide significant benefits to the *E. coli* host such as substrate utilisation and stress tolerance in harsh environments [21]. A recent analysis of 2,085 *E. coli* genomes estimate that there are 3,188 core gene families across *Escherichia coli* species and approximately 89,000 unique gene families [22, 23].

1.1.4. Pathogenic *E. coli*

Despite being part of the normal microbiota of healthy individuals and a useful laboratory workhorse, variants of *E. coli* are also known to be pathogens which are estimated to cause million infections *per* year [24, 25].

E. coli is one of the four leading causes of moderate to severe diarrheal disease in children under five years of age [26, 27]. Globally, it is the main cause of urinary tract infections (UTIs) and blood stream infections (BSIs) in the elderly, and of meningitis in neonates [28]. These pathogenic groups can be further subdivided into intestinal pathogenic *E. coli* (InPEC) and extraintestinal pathogenic *E. coli* (ExPEC). Of the InPEC, there are five major groups, known as pathotypes, defined based on the specific disease with which they are associated: enteroaggregative *E. coli* (EAEC), enterohaemorrhagic *E. coli* (EHEC), enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC) and diffusively adherent *E. coli* (DAEC). Four main pathotypes exist for ExPEC: uropathogenic

E. coli (UPEC), neonatal meningitis *E. coli* (NMEC), avian-pathogenic *E. coli* (APEC) and sepsis-associated *E. coli* (SePEC). The intestinal *E. coli* pathotype is based on the clinical manifestation of the disease and the virulence genes which they carry [29]. EPEC harbours loci of enterocyte effacement (LEE) and bundle-forming pilus gene (*bfp*), whereas EHEC encodes LEE together with genes encoding the Shiga toxin (*stx1*, *stx2*, or both together). ETEC expresses the LT and ST enterotoxins. EIEC possess *ipaH* genes located on pINV plasmids that enable this pathotype to invade host cells. The genetic determinants of the APEC pathotype remain unclear [30].

Although there are many different *E. coli* pathotypes, the subject of this study are those pathotypes associated with bloodstream infections, sepsis and bacteraemia. These pathotypes will now be considered in more detail.

1.2. *E. coli* associated with bloodstream infections (BSIs), sepsis and bacteraemia

Prior to the year 2000, it is estimated that 10 million people were affected every year by *E. coli* BSIs, with mortality rates of 44 deaths per 100,000 individuals [31]. *E. coli* bacteraemia is defined by the presence of *E. coli* in the blood, whereas sepsis is defined as the host systemic inflammatory response to the *E. coli* bloodstream infection, which leads to multiple organ failure [32]. *E. coli* can enter the blood as a consequence of severe primary infections such as pneumonia or urinary infection [33]. Other routes of infection include the direct inoculation of the bacteria into the blood *via* such routes as chronic urinary catheterisation, mechanical ventilation or prostate biopsy. Bacteraemia can also emerge from the gut through a process called bacterial translocation, whereby bacteria or bacterial products are translocated from the gut to mesenteric lymph nodes, from where they disseminate and cause disease systematically [34, 35]. In leukaemia patients, the bowel appears to be a portal for *E. coli* translocation and repeated bacteraemia episodes, whereby the same genotype has been isolated from blood and stool [36]. Data from mice which have undergone chemotherapy, or from mice bearing tumours, confirm this bacterial translocation hypothesis, although the mechanism underlying this phenomenon remains unknown [37].

In developed countries such as the United States of America, Canada, and the United Kingdom, the annual incidence of invasive *E. coli* bloodstream infections ranges between 30 [38] to 63 cases *per* 100,000 [39]. This increases with age, especially in people over 45 years old, with incidence in this senior population (>45 years old) being up to 150 cases *per* 100,000 [40]. There has also been a general increase in the number of reported of *E. coli* bacteraemia cases [41] in the 21st century; for instance, from 1999 - 2011 and after 2006, the number of reported *E. coli* bacteraemia cases in UK increased from 3.4 to 5.7 *per* 10,000 bedstays [42].

Unfortunately, there are limited data available regarding the incidence and risk factors of *E. coli* bacteremia in developing world. In developing countries within Africa and South Asia, bacteraemia is commonly associated with HIV-positive patients who are co-infected with invasive non-typhoid

Salmonella or *Salmonella* Typhi infections. However, among non-*Salmonella* enterobacterial infections, *E. coli* is still the leading cause of BSIs both in adult and children [43-45] affecting both HIV and non-HIV patients equally [46]. In Cambodia, surveys carried out from 2007 to 2010 showed that *E. coli* accounted for nearly 30% of BSIs cases [47]. In Vietnam, retrospective blood culture data from 1994 to 2008 showed the decline of *S.* Typhi and *S.* Paratyphi A, but an increase in the number of fungi, non-typhoid *Salmonella* and *E. coli* infections that was concurrent with the HIV epidemic [48]. *E. coli* and *Klebsiella pneumoniae* were recently reported as the two most common cause of community and hospital acquired blood stream infection in the North of Vietnam [49]. A recent sepsis study in Southeast Asia encompassing 13 public hospitals in Indonesia, Thailand and Vietnam showed that *E. coli* is one of the most common cause of sepsis in this region [45].

It is hypothesised that those *E. coli* bacteraemia cases which are not related to invasive medical procedures often begin as infections of the urinary tract, gastrointestinal tract, or the lung, and subsequently progress into an invasive blood stream infection [50]. The general risk factors for progressing to a blood stream infection include dialysis, solid organ transplantation and neoplastic disease [38]. In the adult population, aging is an important factor that modulates the risk of developing bacteraemia. Other comorbid medical conditions that alter the immune system, such as infection with HIV, diabetes, cancer, and cirrhosis, also increase the risk of nosocomial bacteraemia [51]. Risk factors for *E. coli* colonisation and infection in neonates include premature birth, low birth weight, prolonged mechanical ventilation, length of hospital stay and antibiotic usage [52].

1.3. Hospital for Tropical Disease (HTD), Vietnam

The retrospective study described in this thesis was conducted at Hospital for Tropical Diseases, a 550-bed tertiary referral hospital for infectious disease in Vietnam. The hospital consists of seven wards, including one intensive care unit (ICU) that treats both adult and children patients and admits 38,000 patients annually. The hospital mainly serves patients resident in Ho Chi Minh City and the southern part of Vietnam, with a catchment population of 42 million people [53]. Patients without infectious diseases, such as cancer, tuberculosis, and non-HIV immunosuppression are referred to other hospitals.

1.4. Antimicrobial resistance in *E. coli*

An antibiotic is defined as a substance able to inhibit the growth of or kill a microbe [54]. Since the discovery of the first antibiotic, penicillin, in 1927 and its initial widespread use as a therapeutic agent in the 1940s, a wide range of antibiotic compounds have saved millions of lives and helped to control infectious disease, as well as to prevent infections during invasive procedures such as surgery and routine biopsies. Whilst the term 'antibiotic' was originally used to describe naturally occurring compounds, the term used now is 'antimicrobial', in order to encompass semi-synthetic antibiotics [55].

Since the first use of antimicrobials, the phenomenon of bacteria becoming non-susceptible or resistant to these compounds has been documented [54]. Bacterial antimicrobial resistance (AMR) describes the

process by which a micro-organism shows a reduced or non-susceptible phenotype to an antimicrobial compound that previously inhibited or killed it. This is achieved through multiple different strategies such as drug inactivation or modification, alteration of the drug target site, alteration of the metabolic pathway affected by the drug, and reduction of the rate of drug accumulation within the cell *via* reduced uptake or active efflux [56]. Bacterial AMR phenotypes can be due to intrinsic or acquired mechanisms of resistance. Intrinsic resistance is the innate ability of a bacterium to be resistant to antimicrobial compounds due to natural structure or functional characteristics, thereby allowing bacteria to tolerate particular drugs. For example, *Klebsiella* species are naturally resistant to penicillin since they carry core chromosomal genes that encode enzymes which degrade ampicillin as it enters the cell. Acquired resistance however is when drug sensitive bacteria become resistant through point mutations, by acquiring genes from other bacteria through horizontal gene transfer or a combination of these two mechanisms. An example of acquired resistance through the generation of point mutations include mutations within genes encoding the drug target site – these include point mutations in chromosomal genes such as DNA gyrase or the 23S rRNA genes which confer resistance to fluoroquinolones and aminoglycosides, respectively. Examples of acquired resistance through horizontal gene transfer of mobile genetics elements (MGEs) from another bacterium include plasmids that encode genes that confer resistance to carbapenems and most other classes of antimicrobial [56](see Section 1.6).

Although point mutations in chromosomal genes are vertically inherited, and their dissemination throughout a bacterial population requires the spread and expansion of a resistant bacterial clone, MGEs are particularly able to disseminate rapidly and broadly, not only between members of the same species, but also more broadly within bacterial families. For example, *mcr*-1, the gene that confers colistin resistance, was first identified in *E. coli*, but is now found in *Salmonella*, *Klebsiella*, *Enterobacter* [57] and has also been shown to be capable of transfer outside of the *Enterobacteriaceae* into *Pseudomonas aeruginosa* in the laboratory [58].

1.5. AMR *E. coli* in Vietnam

For critically ill patients suffering BSIs, especially in patients in intensive care units, oral or intravenous antibiotics are the only interventions suitable to combat these infections. However, modern treatment options for *E. coli* are very limited, due to the high level of AMR in this species. Multi drug resistant *E. coli* (MDR) is defined as bacteria with reduced susceptibility to at least one antimicrobial agent in three or more antimicrobial classes. In the Hospital for Tropical Diseases (HTD), Vietnam, according to the hospital treatment guideline, patients are treated empirically with fluoroquinolone or third generation cephalosporin if they have a suspected BSI. The reality in HTD, as well as in many other hospital settings, is that blood culture results and antimicrobial susceptibility test data are generally not available until 48 hours after testing. If a patient fails to respond to these antimicrobial treatments because they are infected with a MDR *E. coli*, for example, it is sometimes too late to switch to another antibiotic therapy regime [59]. Such patients commonly develop septic shock leading to a poor

prognosis [60]. Furthermore, successful treatment of patients infected with MDR *E. coli* increases treatment costs because patients need to be treated with newer classes of antibiotic. Infection with MDR *E. coli* BSIs also often recur, and significantly increase length of hospital stay [61]. It is therefore important to identify the prevalence of bacterial infections that show AMR to guide better treatment and patient management. Currently in HTD, all patients with ESBL-producing *E. coli* will be treated with carbapenem or with piperacillin-tazobactam as alternative therapy (personal communication).

The increasing of MDR *E. coli* is not restricted to HTD in Vietnam. The same trend has been observed in other hospitals in low to middle income countries [62, 63]. Indeed, *E. coli* has now been listed by the World Health Organization (WHO) as one of twelve bacteria that pose the greatest threat to human health, highlighting the urgent development of new antibiotic therapies [64]. Since 3rd – generation cephalosporin was used as an empirical treatment for this study we will only focus on the prevalence of ESBL – producing *E. coli* collected between 2010 and 2015 from patients attending HTD in Vietnam.

1.5.1 ESBL – producing *E. coli* and carbepenemase- producing *E. coli*

Extended – spectrum β-lactam (ESBL) is a term that refers to a class of enzyme capable of hydrolysing β-lactam antibiotics. β-lactam antibiotics are a class of broad spectrum antibiotics containing a β-lactam ring, including penicillins, cephalosporins, monobactam and carbapenems. These antibiotics bind to bacterial penicillin binding proteins, making the bacterium unable to synthesise cell walls correctly, leading to a failure of the bacterial cell to divide. Third generation cephalosporins have a broad spectrum of activity against all Gram-negative species; these are favoured by clinicians for use in treating a variety of infections [65, 66]. β-lactamase enzymes, however, can be inhibited by inhibitors such as clavulanic acid or tazobactam.

Prior to the 1990s, class A β-lactamase enzymes, including TEM and SHV enzymes, were typically responsible for conferring an ESBL-producing phenotype upon the *Enterobacteriaceae*. However, the epidemiology of ESBL-producing *E. coli* changed dramatically since the acquisition of CTX-M genes by this species. These genes have been proposed to have originated from *Kluyvera* spp., and are encoded mainly on plasmids [67]. The genetic context of CTX-M-producing *E. coli* is complicated; CTX-M genes are often co-inherited with other resistance genes as part of Insertion Sequences (IS), fragments of DNA capable of being mobilised around bacterial chromosomes and plasmids. IS elements such as IS*Ecp1*, IS*CR1* and IS*26* often coexist with $bla_{\text{CTX-M}}$ [68].

Currently, over 150 $bla_{\text{CTX-M}}$ genes have been described and they are classified into five groups based on their amino acid sequence similarity: CTX-M-1, 2, 8, 9 and 25. CTX-M proteins share >94% amino acid identity within each group and <= 90% similarity between various groups.

1.5.2 Carbapenemase-producing *E. coli*

Carbapenems are also members of the β-lactam class of antibiotics, and include imipenem, ertapenem, meropenem and doripenem. Carbapenems are often considered to be "antibiotics of last resort". Despite this, resistance to carbapenems conferred by carbapenemase enzymes is spreading worldwide [69]. Different type of carbapenemases include bla_{NDM}, bla_{KPC}, bla_{OXA}, bla_{IMP} genes. However, since the majority of *E. coli* remain sensitive to this drug within our hospital, we will not focus on these genes in detail in this thesis.

1.6. Mobile genetic elements (MGEs)

Mobile genetic elements (MGEs) can mobilise resistance genes horizontally within a bacterial population

Horizontal gene transfer (HGT) is the process by which genetic material can be transferred between different strains of a bacterial species (intra-transfer), or between closely related bacteria of different species (inter-transfer). Unlike naturally occurring mutations, which take time to accumulate in a genome and to accumulate in a population during repeated cellular division, bacteria can potentially acquire multiple resistance genes, and therefore phenotypes, in a single acquisition event. The dissemination of drug resistance within a population of bacteria can be potentiated by HGT, such that resistance genes can be spread both laterally and vertically throughout a population.

There are three mechanisms for HGT - transformation, conjugation and transduction.

Transformation is the process by which competent bacteria take up exogenous DNA and thereby incorporate novel genetic material into the genome. This may take place under stressful conditions, or under conditions that affect bacterial growth. In the environment, extracellular DNA is constantly excreted from live bacteria and is released from dead bacteria, serving as a pool of genetic information. Naked chromosomal DNA or plasmids released into soil and water can remain intact for hours or days, whereas DNA that is released into the blood by invasive bacteria will be degraded within several minutes by DNAse enzymes that are components of human serum. However, this relatively short time is sufficient for transformation to occur.

Conjugation is the transfer of genetic material, often plasmids or transposons, through direct contact between bacteria cells using a pili as a bridge. Each plasmid has their own replication system, those with the same replication system cannot coexist stably within the same cell, and are considered to be "incompatible". This incompatibility led to the development of the incompatibility typing scheme (commonly called Inc-typing) for plasmids [70]. Plasmids encode their own replication machinery, separate to that of the chromosome, and also can be maintained vertically using toxin-antitoxin systems [70]. Transposons which contain IS elements and can be mobilised *via* self-encoded transposase enzymes, however, can be incorporated into the host chromosome or plasmids through transposition.

Transduction is the process by which bacteria acquire new DNA upon infection by bacteriophages. Bacteriophages are viruses that infect bacteria and incorporate themselves into the host chromosome as a prophage during the lysogenic stage of their life cycle. A prophage can switch from a lysogenic state to a virulent lytic state, and in doing so may co-package host DNA together with the phage genome. Upon cell lysis, this host DNA can be transduced to a new host cell upon infection by a viral particle, and thus facilitate gene transfer between different bacteria [71].

1.7. Reservoirs of *E. coli* and transmission mechanism

The ability of *E. coli* to colonise and survive in different hosts and environments makes it difficult to identify the source of *E. coli* infections. In the past, *E. coli* infections such as those caused by ExPEC have been described as sporadic infections caused by bacteria that originate from the host's gastrointestinal (GI) tract [72]. In addition, *E. coli* is able to colonise the GI tract of warm blooded animals; therefore, many investigations sought the origin of transmission through the food chain and farm animals. Unfortunately, these studies used low resolution genotyping methods such as pulse-field gel electrophoresis (PFGE), multi locus variable - number tandom repeat analysis (MLVA) and multi locus sequence typing (MLST) [73, 74]. Although detection of the same MLST genotype and same AMR phenotype can occur between animal and human isolates, the conclusions drawn from these studies are potentially misleading and do not provide enough resolution to infer *E. coli* transmission. For example, in a study by Been and colleagues, they show that at the whole genome sequence level, isolates from chicken and humans were not identical and they actually share the same mobile genetic element (plasmid) rather than representing the spill over of same clonal strains between two populations [75].

A recent study in Sweden also supports the notion that the clonal spread of *E. coli* from animals to human was unlikely. This study by Börjesson *et al.* on ESBL-producing *E. coli* from nearly 5,000 samples collected from food, farm animals, healthy humans and BSI patients found that there was limited clonal spread of these *E. coli* strains from animals and food to humans, but *E. coli* from poultry can serve as a reservoir for drug resistance genes and for plasmids that can move to *E. coli* types commonly found in humans [76]. Nevertheless, whether *E. coli* from animals still represents a source for human infections requires further investigation even if it remains controversial [77]. Outside human and animal hosts, *E. coli*, especially strains within phylogroup B2 and D found in coastal sediments, could represent a potential reservoir for commensal and pathogenic *E. coli* [78].

1.8. Aims and objectives of the study

There are no data available regarding the genetics of those *E. coli* that cause bloodstream infections (BSIs) in the Vietnamese population. We aimed to sequence 690 bacterial samples to obtain a high-resolution snapshot of the landscape of *E. coli* causing BSIs. We also screened for virulence genes and AMR genes to determine if these non-core genes are associated with invasive disease.

Research questions

1. Is there a specific lineage or sequence type that is circulating and causing invasive disease in Vietnamese patients, or disease outcome dependent upon host factors alone, such that all *E. coli* are capable of causing invasive disease? (*i.e.*, is there any specific lineage associated with disease outcome?)
2. How is the gene content of invasive *E. coli* different to that of carriage *E. coli* isolates? Are there significant genetic markers that differentiate these invasive bacteria such as AMR genes, colonisation factors, toxins, siderophores, capsule type, *etc*.?

Chapter 2: Comparative Genomics of BSIs caused by *E. coli* in the Hospital for Tropical Diseases in Vietnam

2.1. Introduction

Hospital acquired infections (HAIs) are a major threat to patient safety, which in locations with poor surveillance and infection control can lead to high rates of infection and significant mortality. The problem of HAIs is further exacerbated by the emergence of antimicrobial resistance, particularly extended spectrum β-lactamases (ESBLs) and carbapenemases in Gram-negative bacteria.

Healthcare settings across Asia are under increasing pressure to control the spread of multidrug-resistant (MDR) HAIs and to reduce the overall levels of MDR in bacteria. Recent work has specifically highlighted nosocomial bloodstream infections (BSIs) as a major cause of mortality in high dependency units. Notably, two outbreaks of *Klebsiella pneumoniae* in 2012 occurred in the neonatal intensive care unit (NICU) in a tertiary hospital in Nepal, which had a mortality rate of 75%. This study found, using whole genome sequencing, that there were regular peaks of BSIs across the hospital, associated with MDR HAIs that recurred even after deep cleaning of the hospital. This upsurge in MDR BSIs with high mortality is a growing trend echoed in other healthcare settings across Asia [79]. These peaks in infection are driven by an increase in the rate of isolation (from blood) of bacteria such as *Escherichia coli* that are resistant to most available antimicrobials. Currently, in 2015, 24% of all BSIs in Hospital for Tropical Disease (HTD), Ho Chi Minh city, Vietnam are attributable to *E. coli* (unpublished data). Understanding the nature of HAIs across Asia is required to enhance infection control measures, to optimise therapeutic approaches, to minimise the further emergence of antimicrobial resistance, and to reduce morbidity and mortality due to BSIs caused by MDR HAIs.

The hypothesis:

We hypothesise that bacterial isolates causing BSIs are more closely related to one other than to isolates taken from the gut of patients attending our hospital. This is because many BSI *E. coli* are hospital acquired isolates which are likely to be members of a restricted number of genetic lineages, and to carry genetic loci (such as genes encoding antimicrobial resistance and siderophores) which are more strongly associated with an invasive phenotype in comparison to non-invasive phenotypes, including UTIs and respiratory infections.

Specific Aim:

The aim of this chapter was to determine the genetic epidemiology of *E. coli* causing BSIs in the Hospital for Tropical Diseases (HTD) in Ho Chi Minh city, Vietnam. To address this hypothesis, we set out to compare the genomes of *E. coli* associated with BSIs to those of carriage *E. coli* isolates taken from patients attending this hospital. For a limited number of samples, it was possible to collect bacterial

isolates from both the rectal swab samples and bloodstream infections of the same patient for comparison.

2.2. Methods

2.2.1. Strain collection

Our strain collection consisted of 690 *E. coli* isolates derived from two different studies called 15EN and 06EI. Of these 496 represented a longitudinal collection of isolates that were cultured from patients with BSIs admitted to Hospital for Tropical Diseases (HTD) in Vietnam between 2010 and 2014 (15EN). These isolates were selected randomly from the approximately 100 isolates collected per year during this period. Also included were 94 randomly selected *E. coli* isolates taken from rectal swabs from patients included in the 06EI study looking for HAIs in the intensive care unit (ICU) in HTD during 2015. All of the above randomly selected samples were collected from patients that had been sampled on or after admission. All isolates are described in detail in Supplementary Tables S1 and S2.

The remaining 100 isolates included in this study were derived from 31 patients with matched rectal swab (69) and BSI (31) isolates. These patients were enrolled in an extension of the 15EN study during 2015 at HTD. These matched samples are included in the general analysis below but will also be described in more detail in Chapter 4. All rectal isolates were considered to be carriage isolates as they were not linked to any clinically recorded signs of enteric disease.

2.2.2 Ethics

Both studies were approved by the ethical committee at Hospital for Tropical Disease (HTD) and by the University of Oxford Tropical Research Ethics Committee (OxTREC).

2.2.3 Bacterial isolation

All microbiological culture work was performed in laboratories at HTD and OUCRU. Routine care at HTD dictates that patients admitted with febrile illness and with a suspected bacteraemia will have their blood cultured as part of routine diagnosis of the hospital. Briefly, for these samples, 5 mL of patient's blood was incubated at 37 ˚C in a BACTEC bottle. Positive BACTEC bottles were then sub-cultured onto selective MacConkey agar (MC) for further microbiology identification either by matrix assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI TOF, Bruker) or using API20E strips (BioMerieux, France) following the manufacturer's guidelines. Commensal and carriage *E. coli* were cultured from rectal swabs of patients participating in the 06EI study. Swabs were taken twice weekly from admission day until discharge. Microbiology identification procedure was as for blood isolates, except that for each rectal swab, if *E. coli* colonies with different morphologies or AMR phenotypes (lactose fermenting, non-lactose fermenting and β-haemolytic) were identified, multiple colonies were collected. The number of isolates collected per patient ranged from one to four isolates,

except for one patient from which a total of 20 isolates were collected during a long-term hospitalisation (described in detail in Chapter 4).

All isolates included in this study were tested for their antimicrobial susceptibility phenotype using 13 antimicrobials, also described in detail in Chapter 3.

2.2.4. Metadata

Patient demographic data were extracted from electronic hospital records in both studies. This included information about patients including age and gender (see Supplementary Tables S1). We inferred patient comorbidities, such as their HIV status or chronic hepatitis/liver cirrhosis, based either on the ward into which the patient was admitted, or from initial diagnoses from doctors when blood culture requests were submitted. These samples were denoted "HIV/liver disease" samples. Length of hospital stay was calculated by subtracting the date of discharge from the date of admission. The clinical outcomes for all patients included in this study were available. However, data on recent hospital admission and antibiotic treatment were only available for isolates that were collected in the 06EI study.

In an effort to look for the origin of infection, or coinfection at different bodily sites, the patient identifiers for all patients for whom we had associated blood-borne isolates were used to screen the urinary and peritoneal fluid (ascites) positive patient database to determine whether there were also corresponding records of *E. coli* isolates cultured from the liver abscesses or urinary tract infections from that patient too. However, we were only able to collect information about antimicrobial susceptibility phenotype for these *E. coli* isolates. Unfortunately, we could not include them in this study for sequencing because the samples were not being kept.

We define community acquired (CA) infections as being patients who test positive for *E. coli* from blood cultures collected during the first 48 hours of the patients admission to HTD and define hospital acquired (HA) infections as isolates cultured from blood collected at least 48 hours after admission.

2.2.5. Whole genome sequencing

Genomic DNA was extracted in OUCRU using Wizard Genomic DNA Purification kit (Promega, USA) and sent to the Sanger Institute. Approximately 2 μg of DNA was subjected to sequencing using Illumina HiSeq 2500 platform to generate 125 bp paired end reads.

2.2.6. Accessions

Short read sequence data from this study are available in the European Read Archive (ERA) under project accession ERP017161. All samples and their accession numbers are detailed in Supplementary Table S1 and S2.

2.2.7. Genome assembly and annotation

Annotated assemblies were produced using the pipeline described in detail in Page *et al* [80]. For each

sample, sequence reads were used to create multiple assemblies using VelvetOptimiser v2.2.5 [81] and Velvet v1.2 [82]. An assembly improvement step was applied to the assembly with the best N_{50}, and SSPACE was used to scaffold the contigs [83]. Sequence gaps were filled using GapFiller [84]. Automated annotation was performed using PROKKA v1.11 [85] and *Escherichia* databases from RefSeq [86].

2.2.8. Quality control of sequencing data

In order to confirm the quality of the sequenced bacterial genomes, a range of different tools were used to determine the quality of the short-read Illumina sequence data. Short reads data quality score was accessed using FastQC v0.11.4 [87]. FastQC compiles sequence quality into one report which details basic statistics of all reads in each sample, such as GC content percentage (GC%), total number of reads, and sequence quality score. All reports were aggregated using MultiQC [88] for a complete output and was manually assessed. Samples showing a GC% different from that expected for *E. coli* (49-52%) were excluded from the analysis. Sequence quality was further checked by mapping the reads to a reference genome *E. coli* K-12 strain MG1655 (EMBL accession number U00096). Samples were excluded if number of contigs was higher than 300 or the total reference genome coverage was less than 80%.

Kraken was used to assign taxonomic labels to short DNA sequences to confirm the genus and species of each sample [89]. CheckM was also used to test for genome heterogeneity and completeness [90].

2.2.9. *In silico* MLST and phylogroup classification

Multilocus sequence typing (MLST) is a traditional typing method that uses the combined sequence of seven housekeeping genes in the genome to assign the bacteria to a sequence type (ST). STs were determined from WGS data using MLST check [91], which compares the assembled genome against the *E. coli* MLST database (http://mlst.warwick.ac.uk/mlst/dbs/Ecoli) to assign a specific allelic profile.

Originally *E. coli* was classified into phylogroups A, B1, B2 and D: phylogroups were defined as group of isolates representing distinct clades in a phylogeny built from the ECOR collection [92]. Different phylogroups were believed to have differences in gene content that reflected different niches and lifestyle [93]. *E. coli* phylogroup of each isolate was determined by blasting *de novo* assemblies against three marker genes, *chuA*, *yjaA* and *tspE4.C2* proposed by Clement *et al* [94]. Those from our collection that could not be classified into these four phylogroups were label as “other phylogroups”.

2.2.10. Construction of core genome phylogeny

The core and pan genome of *Escherichia coli* was determined using ROARY [95], in which core genes are defined as those which produce products that share a minimum BLASTp percentage identity of 95% and are present in at least 95% of all isolates. The sequences of all core genes were concatenated into a multiple sequence alignment file. We also included 25 reference genomes from annotated reference

strains from different phylogroups, plus *E. fergusonii* and *E. albertii* (see Supplementary Table S3).

Approximate maximum likelihood phylogenetic trees, based on core gene alignments, were constructed using FastTree v2.1 [96]. SNP alignments of core genes was used to subdivide the population of the sequenced *E. coli* isolates by hierarchical clustering using hierBAPS [97]. Phylogenetic trees and associated metadata were visualized in iTOL [98] or Phandango [99].

2.3. Results

2.3.1. Sequencing and QC

DNA was prepared for whole genome sequencing from a total of 690 isolates (see Methods). Following library preparation and quality control steps, which took into account the combined results from FastQC, Kraken, CheckM, genome size as well as number of contigs (see Methods), 665 were taken forward for further analysis (see Figure 2.1). This included 506 genomes from isolates cultured from blood and 159 from isolates cultured from rectal swabs (described below; Supplementary Table S1 and S2).

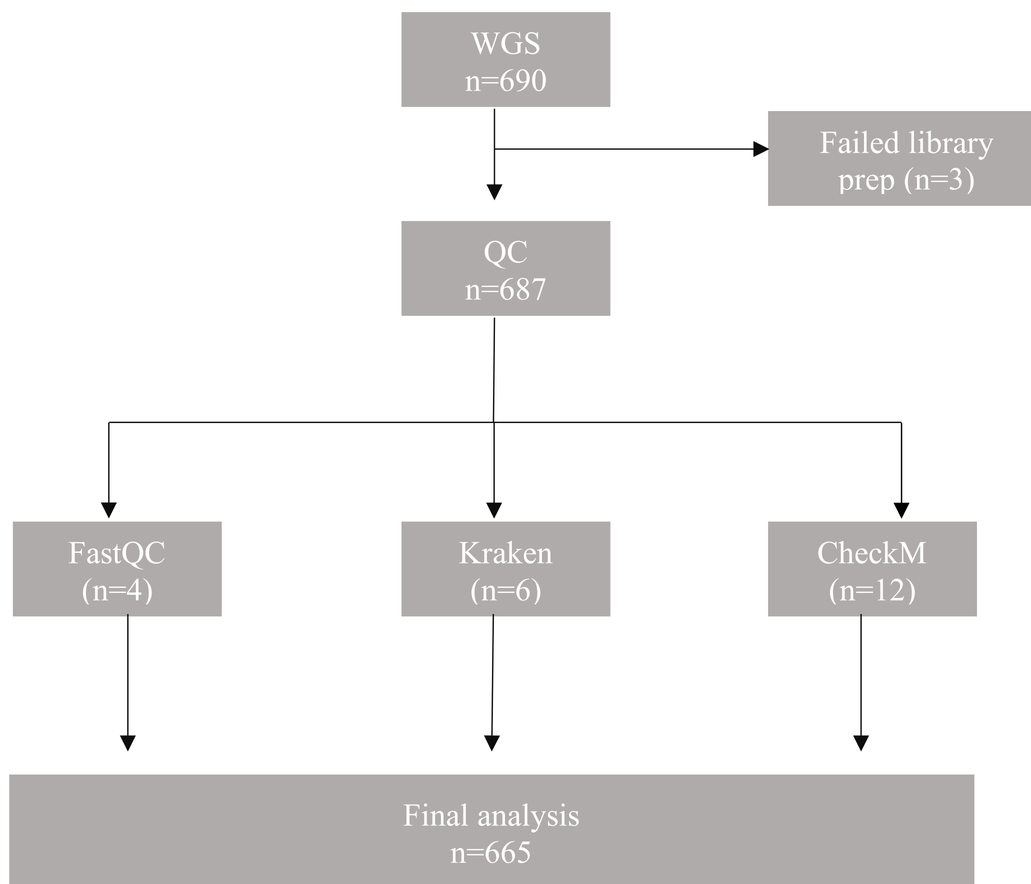


Figure 2.1. Sequencing and QC work flow for the samples included for whole genome sequencing (WGS) in this study. Indicated are the number of samples that passed through each stage of sequencing and QC. The numbers in brackets indicate the number of samples that failed at each sequencing stage or QC step.

2.3.2. Basic patient demographic and clinical outcome of hospitalised patients with BSIs.

The patient characteristics linked to the 506 *E. coli* genomes derived from BSIs are summarised in Table 2.1. Looking at the demographic data, 54% of all *E. coli* BSI patients were female and the median age of all patients with BSIs was 53 (interquartile range (IQR) 38 – 65 years old). Based on the ward into which the patient was admitted and/or the hospital records, immunocompromised patients infected with HIV accounted for 63/506 (12%) of all BSI patients while 109/506 (22%) BSI isolates were derived from patients with chronic hepatitis B/C or cirrhosis. If age was taken into account then for patients under 40 years of age, 37% of patients with BSI were also HIV positive, while in the BSI patients between 40 and 60 years of age, 7% of them were HIV positive and 32% had liver disease (Figure 2.2).

Table 2.1: Clinical characteristic of *E. coli* BSIs patients

Patient characteristic		**n=506**	**% - Median(IQR)**
Gender	female	270	53.36
	male	236	46.64
Age		53	38-65
Acquisition	CA (<48h)	445	87.94
	HA(>48h)	57	11.26
	unknown	4	0.79
Outcome	recover	368	72.8
	worsen*	56	11.1
	no change*	69	13.7
	death*	12	2.4
Hospital length		11	8-15
Comorbidity	Liver disease	109	21.54
	HIV	63	12.45

* "Death" means patients died at HTD. "worsen" means clinical symptom progress toward a severe infection manifesting as increased breathing rate, increased white cell counts, respiratory failure, coma. This is likely to equate to death since Vietnamese traditional practice is to be discharge patients home to die. "No change" means the patients clinical condition did not improved after treatment or they were transfered to another hospital and therefore lost to follow up.

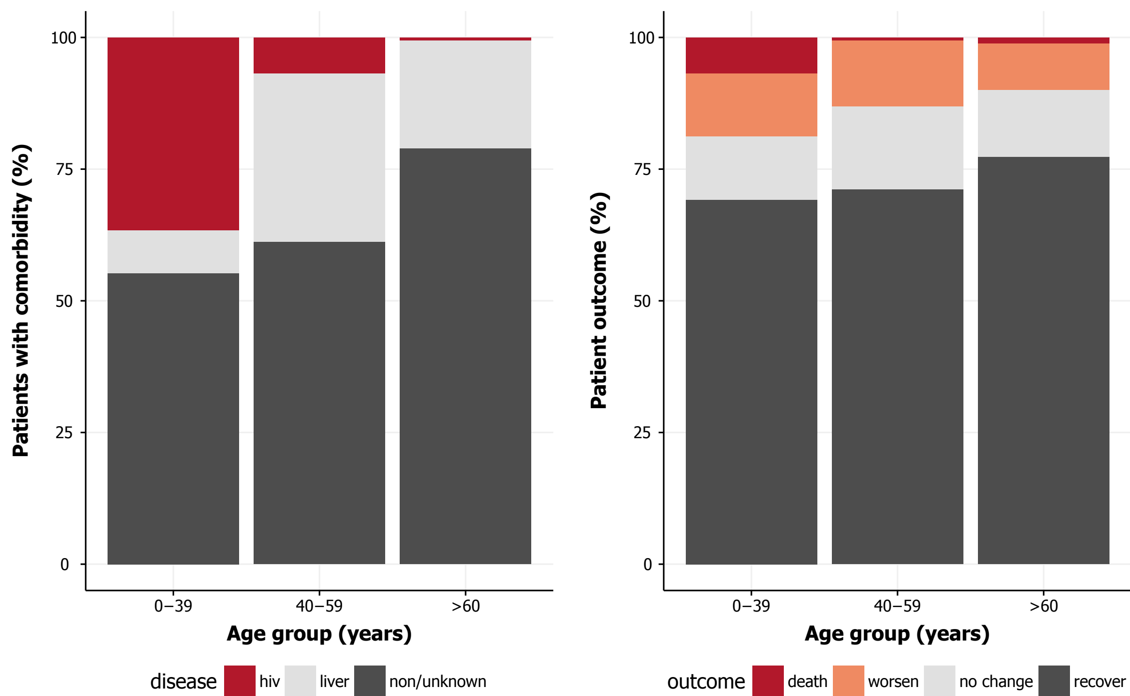


Figure 2.2: Comorbidity and outcome in patients with *E. coli* BSIs. Other comorbidity but not prevalent including diabetes, severe pneumonia, severe Dengue fever, cholecystitis and meningitis.

Overall, 368/506 (72.8%) patient conditions were improved or recovered when discharged. 56/506 (11%) was discharged to die at home, while 12/506 (2.4%) died at the hospital. Sixty-nine patients (69/506;13.6%) showed no improvement, either because they were transferred to another hospital at discharge or clinical symptoms had not improved after treatment and they had been discharged so there was no means to confirm clinical outcome. Mortality rates could range from 14% up to 27% (summarised in Figure 2.2). If gender is considered mortality rates were 58% and 42% in men and women, respectively. However, mortality was highest in HIV positive (8/12; 66.7%) patients of which the majority were men (6/8 patients; 75%), explaining the anomaly. The mean time to death at the hospital was 2 days (IQR 1-3 days) after admission, while patients, if they survived spent an average of 12-14 days recovering from BSIs.

2.3.3. The genomic diversity and population structure of BSI and carriage *E. coli* from Vietnamese patients attending HTD.

In order to construct an accurate phylogeny, we determined the core genome for the entire sequenced dataset and 22 reference *E. coli* genomes plus three *Escherichia* outgroups: one *E. fergusonii* and two *E. albertii* (Supplementary table S3). The core and pan genome of the *E. coli* was determined using ROARY (see Methods). The pan genome of all isolates analysed in this study, amounted to 42,856

genes (Figure 2.3). In total, 2,796 genes were conserved in more than 95% of all isolates. Since the average *E. coli* genome contains approximately 4,500 ~ 5,000 genes, this indicates that the core genes comprised approximately half of the average genome gene content. Interestingly, from the accessory genes, the number of genes present in less than 15% (103 isolates) of total genomes comprised 36,531 genes. This underlines the extreme level of diversity seen in these genomes and indicated that *E. coli* has an open genome (Figure 2.3), consistent with previous studies [100]. This also underlines the ability of *E. coli* to adapt to new environments and indicates that there is a high level of genetic diversity in *E. coli* present in patients attending HTD. Our pan genome prediction includes all mobile elements such as prophage and insertion sequences.

To determine the population structure of both BSIs and carriage *E. coli* included in this study, and including the reference sequences, we constructed a phylogenetic tree from a multiple sequence alignment of the core genes. We identified 443,135 single nucleotide polymorphisms (SNPs) over 2,700,577 nucleotides of core genes. Due to the high genetic diversity of the alignment, maximum likelihood phylogenetic trees were constructed using FastTree v2.1 [96]. The SNP alignments of core genes were used to subdivide the population of the sequenced *E. coli* isolates by hierarchical clustering using hierBAPS [97]. This analysis identified 15 primary BAPS clusters denoted as L1-L15 (Figure 2.4). The MLST ST and ECOR phypogroups was also predicted for all isolates in order to link the genomic data to the common terms of reference used in hospitals, like HTD, to describe *E. coli*.

Looking at MLST STs across our dataset, we identified a total of 117 STs amongst the 665 isolates analysed, of which 12/117 (10.2%) were new STs with a single locus variant (Supplementary table S1and S2). The ST distribution showed significant diversity but also highlighted the circulation of 12 dominant STs (ST131, ST95, ST1193, ST38, ST69, ST648, ST73, ST405, ST354, ST410, ST12, ST10), which account for nearly 70% (453/665) of all STs identified in this study (Figure 2.4). Among these, ST131 and ST95 were the most common in our collection, comprising 17.6% (117/665) and 8.7% (58/665) of all STs identified in this study, respectively. Apart from the 12 dominant STs, the remaining 105 STs, were each only defined in between 1 and 9 isolates. The latter are referred to as ‘Others’ in Figure 2.4. It is evident from Figure 2.4 that the 15 HierBAPS phylogenetic groupings are largely congruent with the MLST ST subdivisions apart from BAPS cluster 10 (L10), 11 (L11) and 14 (L14), which encompassed multiple different STs.

When comparing the WGS phylogeny with the ECOR phylogroup assignment for isolates collected in HTD, it was apparent that isolates were represented in all four historical phylogroups A, B1, B2 and D. We confirmed previous observations that phylogroup A and B1 are closely related lineages [101]. However, within each lineage A and B1, there is significant genetic diversity between isolates. For example, in phylogroup A, the median number of pairwise SNPs between isolates was 31,768 (range 0 – 72,133), whereas in phylogroup B1, the median was 28,791 SNPs (range 0 – 37,189). Genetic diversity is lower between isolates in phylogroup B2 (median 193, range 0 – 32,164) and phylogroup

D (median 3,417, range 0-51,318). It is also clear from Figure 2.4 that the ECOR phylogroups are broad subdivisions that include multiple STs and BAPS groupings. Although important to relate these findings to standard terms of reference, it is clear they do not offer sufficiently high resolution for fine scaled analysis.

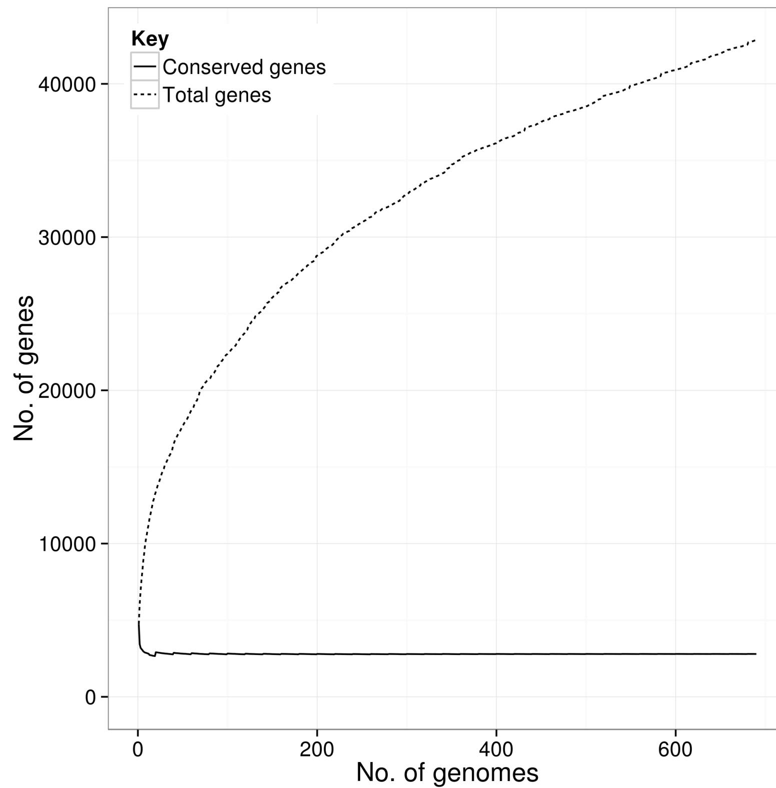


Figure 2.3: The pan genome of 687 *Escherichia coli*, two *Escherichia albertii* and one *Escherichia fergusonii* comprised 42,846 genes in total with 2,796 genes core genes present in more than 95% of all isolates. 33 genes were defined as soft core genes, which were genes present from 94 – 95% isolates. Finally, shell genes were 3,496 genes that appeared in more than 15% but less than 94% isolates.

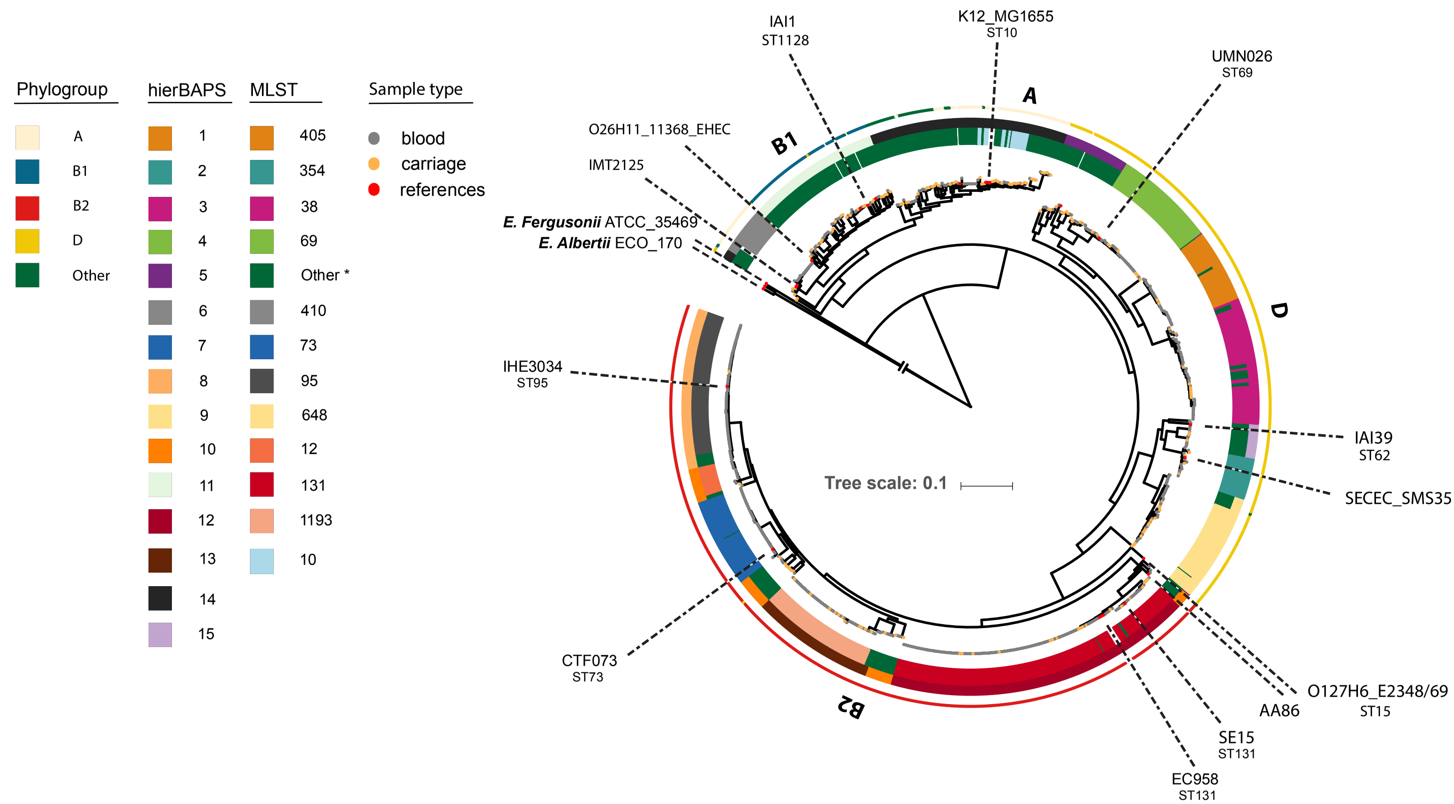


Figure 2.4: Population structure of 665 *E. coli* isolates analysed in this study. Rooted maximum-likelihood tree of *E. coli* from a core gene alignment consisting of 2,796 genes. Tree was rooted using *E. albertii* (ECO 170, KF1) and *E. fergusonii* (ATCC 35469) as outgroups. The tree includes data from 22 reference genomes representing four phylogroups, represent as red dots at the branch tips. The coloured ring from inside to outside includes MLST sequence type (ST) for the dominant 12 STs (See Figure 2.7 for representation of all STs), all hierBAPS clusters and the four ECOR phylogroups. Grey and yellow dots at the branch tips represent BSI isolates and rectal isolates, respectively. Other * denotes allelic variants of specific STs and other rare STs.

Relating whole genome phylogeny, BAPS groups, ECOR phylogroupings and MLST STs to patient clinical data.

The WGS phylogeny of the *E. coli* isolates was then used as a framework on to which MLST, ECOR phylogroup and the clinical metadata was mapped. When relating the phylogeny to clinical data including patient gender, comorbidities as well as linking patient BSIs samples to corresponding isolates reported from other sterile sites (urine/ peritoneal fluid), we identify several patterns. First, although isolates taken from female and male patients were distributed equally across the phylogeny (Figure 2.5), in certain BAPS clusters and STs there was a skew in the gender distribution (Figure 2.5). For example, in ST69(L4), 95(L8), 410(L6), 1193(L13) of phylogroup B2 and D (>60% patients were female) while in other STs of phylogroup B1 and A (L11/L14), majority of patients were male (>60%). These STs 69, 95 and 410 are known to be associated with urinary tract infections (UTIs) in women [102], and this was supported by having the corresponding isolates with the same antimicrobial resistant (AMR) profile from urine isolates listed in the HTD patient database. Secondly, looking at the patient demographic and clinical data for lineages included within ECOR phylogroups B1, A and D showed *E. coli* isolates more associated with carriage and male patients with liver disease as well as patients that also tested positive by culture for *E. coli* from peritoneal fluid (Figure 2.5). Interestingly, isolates taken from patients with HIV were more likely to be associated with phylogroup D and B2, but not exclusively.

Comparing ST and sample type, it was evident that STs within the B2 phylogroup were more likely to have been isolated from blood (52.17%) than stool (25.16%) (264/506 *versus* 40/159, p<0.001, Chi square test, Figure 2.4). In contrast, rectal swab isolates were enriched in phylogroups A and B1: 16.3% (26/159) phylogroup A in rectal swab *versus* 7.31% (37/506) in blood (p<0.001) and 12.58% (20/159) phylogroup B1 in rectal swab *versus* 5.53% (28/506) in blood (p=0.003). However, blood and rectal swab derived isolates were almost equally represented amongst isolates belonging to phylogroup D (31% in blood *versus* 36% in rectal swabs, p=0.275).

From looking at all 117 STs, there were a total of 42 STs only found in isolates taken from blood samples, 42 STs only found only in rectal swab derived isolates, and the remaining 33 STs were evenly distributed between blood and rectal swab *E. coli* (Figure 2.6). In our study, the most globally prevalent ST, ST131, was the most common ST isolated from both blood (19.7%, 100/506 isolates) and rectal swab samples (6.7%, 10/159 isolates) possibly suggesting that carriage in the gut is linked to pathogenic *E. coli* establishing BSIs in an otherwise healthy individual. Conversely, none of ST73 and only one of ST95 were found in rectal swabs (Figure 2.6). This suggests that if these two STs entered to blood after colonizing a different bodily site, then they either lack the ability to colonize the gut in sufficient number to be detected in our clinical screen, or that they may be more prone to colonizing bladder and kidney epithelial cells instead of gastrointestinal tract.

E. coli carriage isolates, taken from rectal swabs, showed many different STs and long branches consistent a highly diverse community of *E. coli* living together in our gastrointestinal tracts (Figure 2.7). However, none of our isolates belonged to the other *Escherichia* species, *E. fergusonii* or *E. albertii,* even when we sample a large collection of *E. coli* longitudinally over a 5-year period.

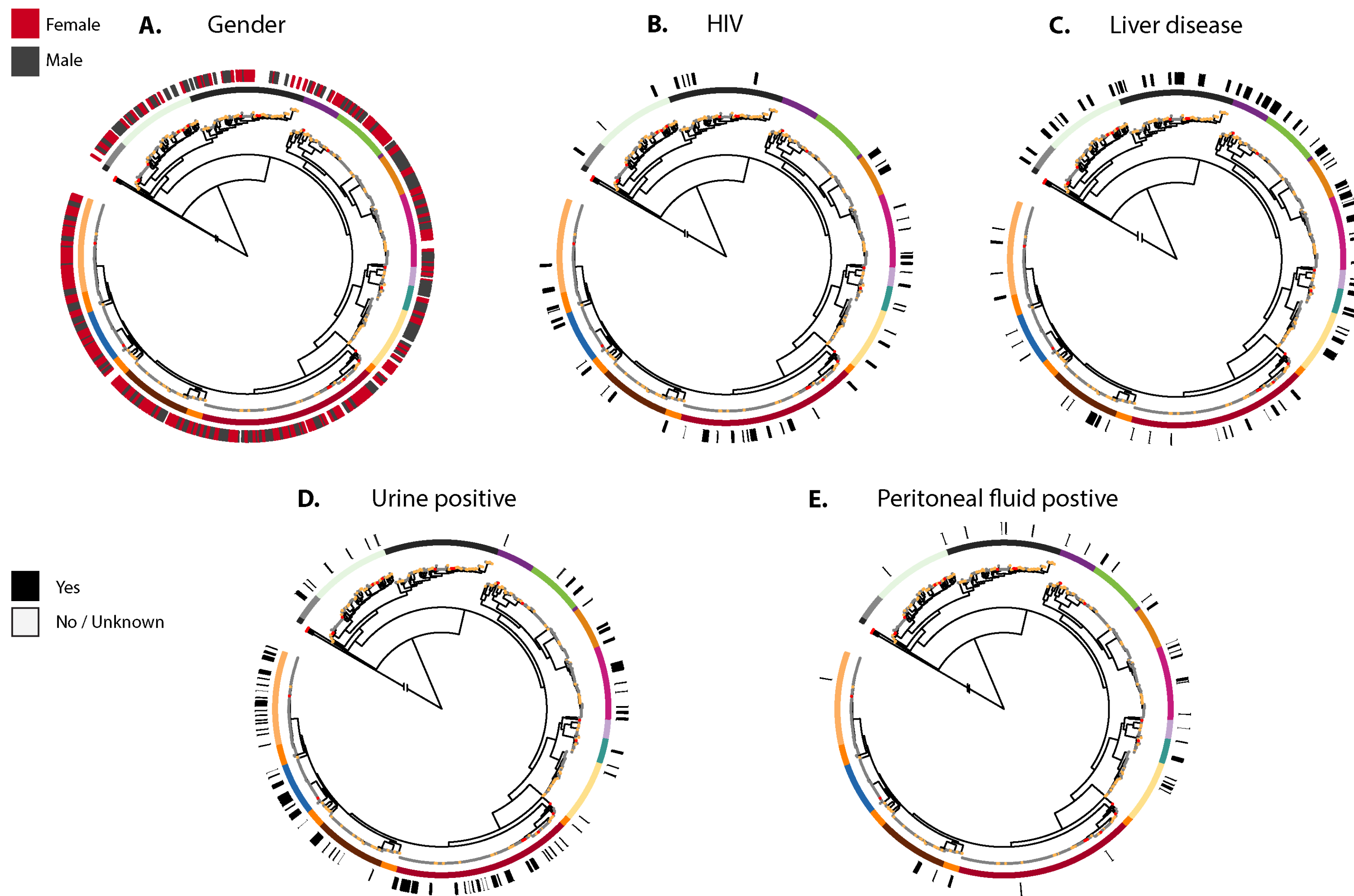


Figure 2.5: The phylogenies of *E. coli* taken from HTD, including reference strains, reported alongside clinical metadata for *E. coli* causing BSI. Inner rings denote BAPS clusters (see key of Figure 2.4), outer rings denote comorbidities or the presence of a corresponding isolate. See key for gender (A) or presence/absence (B-E).

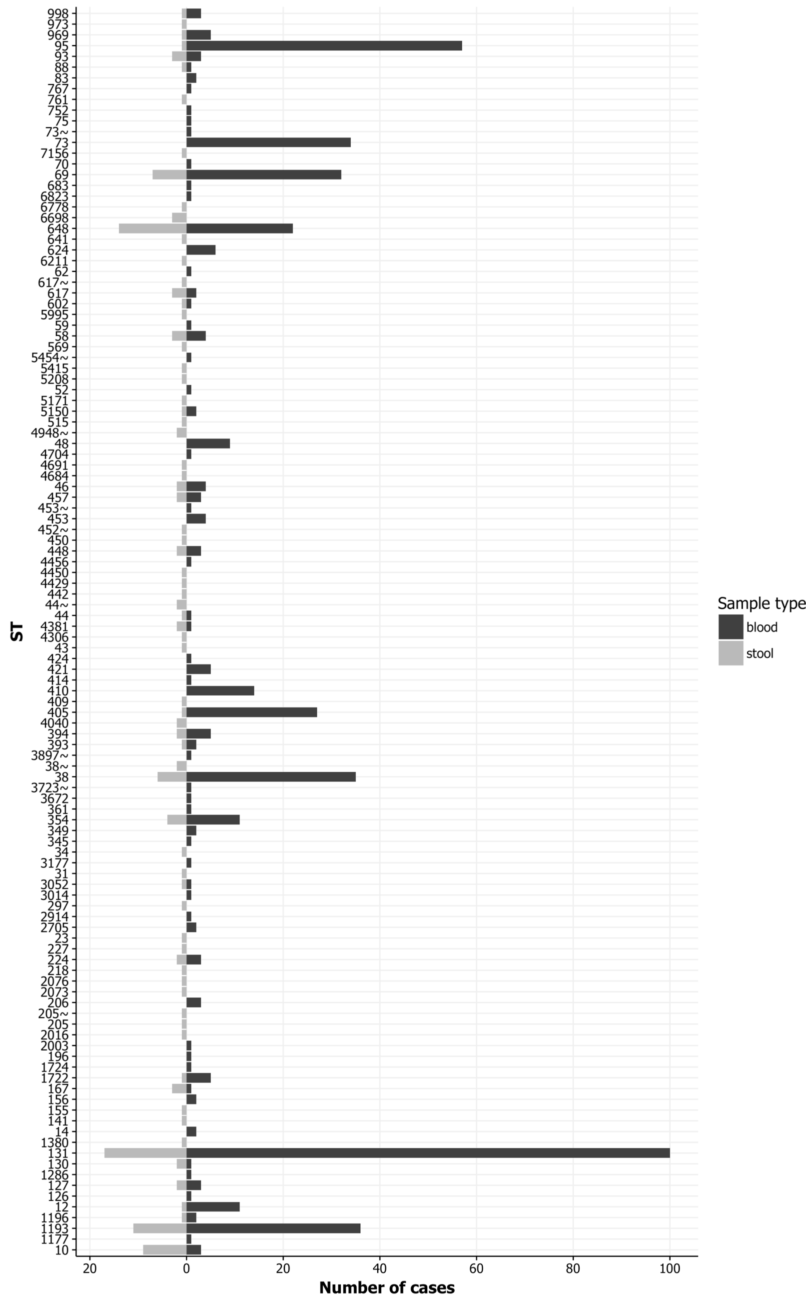
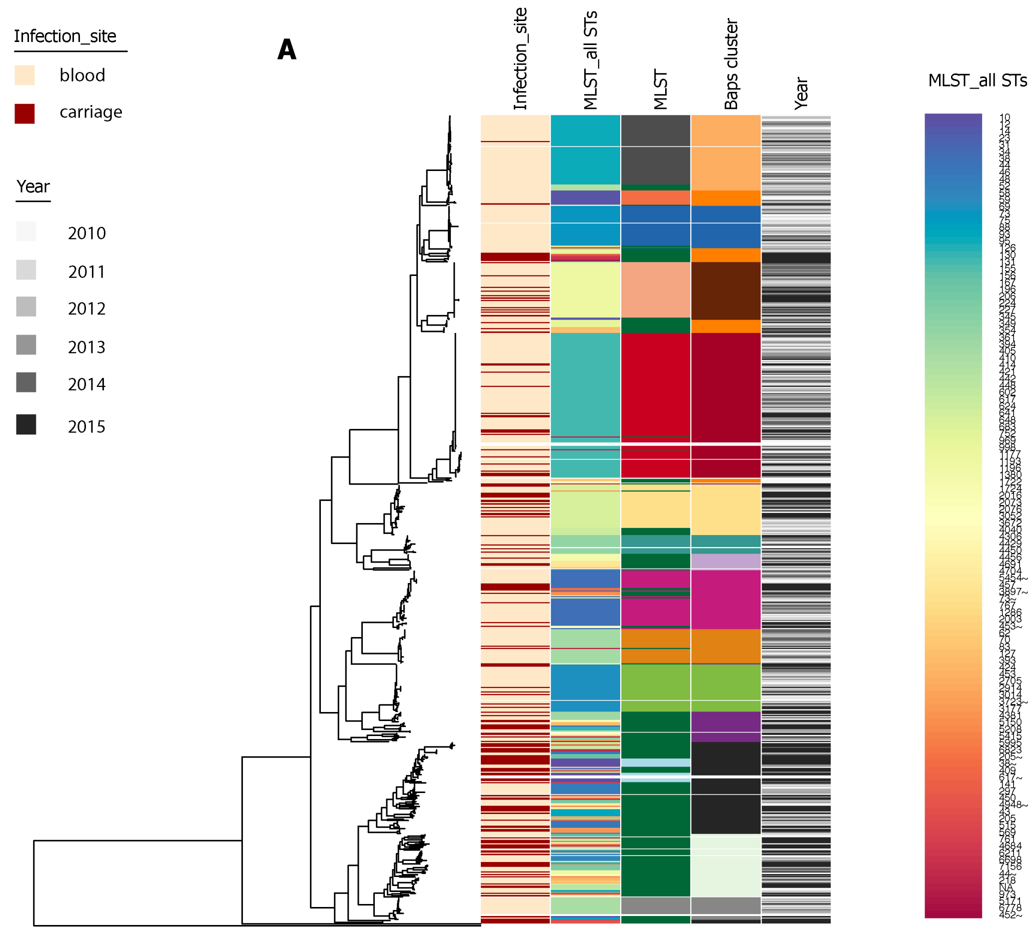


Figure 2.6: STs distribution of blood and rectal swab isolates. For colour see key.

In addition to differences in the relative abundance of STs across the whole collection we observed minor fluctuations in the relative proportion of different STs/BAPS clusters that comprised the total *E. coli* for different years (Figure 2.7), the exception being the proportion of ST131 which remained high throughout the whole 5-year period ranging from 18% to 22% of isolates sequenced/year. The most striking increase in numbers observed was for ST1193 which accounted for 3.7% *E. coli* BSIs in 2011, but then increased to 10% in 2014. We observed a reduction in the proportion of ST69, ST73 and other less common STs over the same time period.



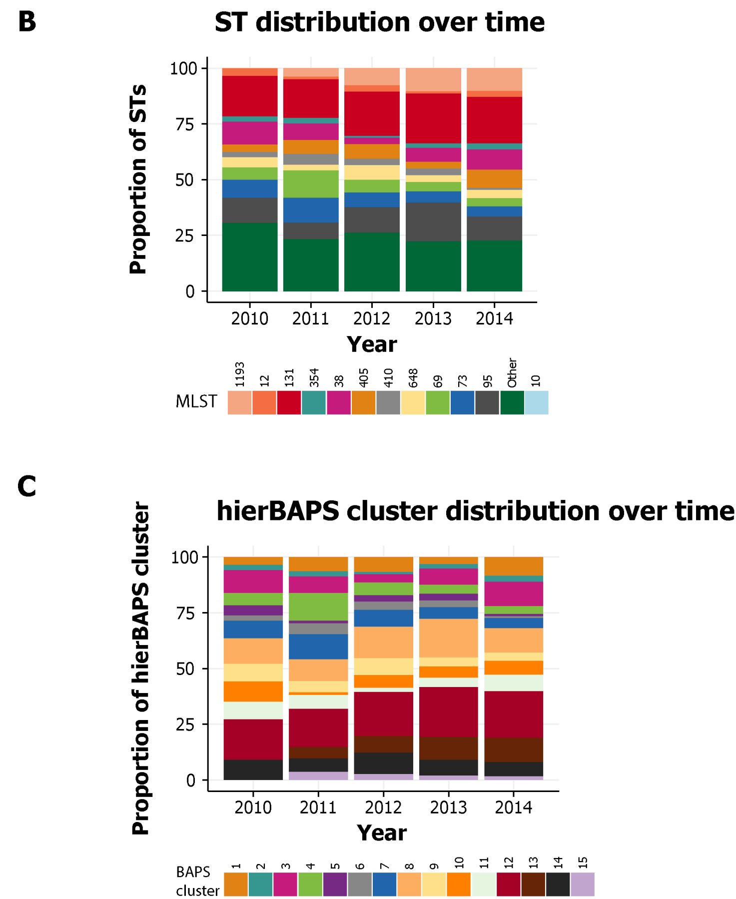


Figure 2.7: A core genome tree visualised in Phandango for blood and rectal swab isolates (A). The distribution of STs (B) and hierBAPS clusters (C) are reported for BSI *E. coli* over a five-year period from 2010 to 2014. Blood and rectal swab samples collected in 2015 were not included in panels B and C because they were not sampled randomly.

2.4. Discussion

Among 38, 000 hospital admissions to HTD every year, *E. coli* causes more BSIs than any other Gram-negative bacteria or fungi, accounting for around 20% of all BSI cases (HTD unpublished hospital records). Unlike seen in other nosocomial bacteria such as *Klebsiella pneumoniae* or *Staphylococcus aureus* [103, 104], the majority of BSI in our study (88%) were from patients with BSIs diagnosed on or within 2 days of admission and so deemed to have been community acquired. The remaining patients were considered to have HAI *E. coli* infections because they acquired the BSI >2 days after admission (this is discussed further in Chapter 4 for the longitudinally sampled patients). All of these patients had severe chronic diseases such as AIDS, cirrhosis, hepatitis B/C or tetanus which required long term hospitalization, and sometimes were seen to develop *E. coli* BSIs as secondary infections as described in detail in Chapter 4. Positively, in terms of patient care, of the 506 patients that had blood stream infections over the period of study 70% recovered and so likely responded to antimicrobial treatment. However, treatment took on average between 12-14 days representing a considerable burden on space and resource within HTD. It was also apparent from the basic patient meta data that was a slight gender bias in BSI towards females (53%) in the sample collection sequenced in this study. In contrast, BSI at HTD caused by all pathogens are generally biased towards occurring in men (65% of all BSI cases, unpublished data).

Aiming to understand the nature of *E. coli* causing BSIs in HTD, Vietnam as well to see if the trend of *E. coli* BSIs seen here mimics patterns of *E. coli* causing disease elsewhere in the world, we used WGS to investigate the population structure of isolates taken from patients with *E. coli* BSIs from HTD between Jan 2010 to Dec 2014. Also included were patients with matched blood stream and rectal swab isolates from 2015. The WGS data for all isolates was used to construct an accurate whole genome phylogeny. This was then used to predict BAPs clusters, ECOR phylogroups and MLST sequence types for all isolates. Not only is this the first study of its type at HTD, this is also the largest collection WGS of *E. coli* from one hospital to date.

We used these data to determine the composition of different STs with potential to carriage and causing BSIs in the HTD patients and showed that there was high diversity in isolates affecting patients at HTD including 15 primary BAPS clusters all the ECOR phypogroups and a total of 117 STs by in silico MLST. From these data, it was clear that the four dominant STs affecting patients attending HTD are the same in many studies internationally [102, 105]. Study by Adams-Sapper *et al* [105] found the same 4 STs dominant from 2007-2010 when investigating the nature of *E. coli* causing BSIs in United States. What was also apparent from these data was that these STs remained largely stable over the time of the study. The exception being ST 1193. We observed the introduction of ST1193 to patients at HTD in 2011, since then ST1193 has increased steadily to account for ~10% BSIs cases every year. Interestingly, there was little documentation of ST1193 in the literature until it was reported recently as

the second most common ST after ST131 in China in 2011 [106]. We hypothesised that the introduction and successful establishment in Vietnamese population was linked to antimicrobial resistant genes they carried. We will discuss in more detailed about ST1193 in final discussion in Chapter 5.

Here we also found strong associations between patient gender and comorbidity with the lineage of *E. coli* causing BSIs: lineages (STs and BAPS groups) within the ECOR phylogroup B1 and A were significantly associated with rectal swab samples, men and liver abscesses. Isolates in these phylogroup groups showed high pairwise differences in SNP content between isolates. Lineages within the ECOR phylogroup D were equally represented in blood and rectal swab samples. But perhaps most importantly our data showed that there was a significant association between members of phylogroup B2 and invasive disease at HTD. In addition, isolates belonging to B2 (as well phylogroup D) were also associated with female patients. For HIV patients (most of whom were male patients in our study), these data suggested that there was no significant association between the phylogenies of the isolates causing BSI and HIV occurrence. These data are consistent with previous studies which have shown that extraintestinal pathogenic *E. coli* (ExPEC; Chapter 1) commonly belong to phylogroup B2, and less frequently phylogroup D, whereas the more diverse commensal and intestinal pathogenic *E. coli* are often members of phylogroups A and B1 [13, 107, 108].

In summary, here we determined the full diversity, and flux in that diversity, of isolates from patients attending HTD. The following chapters will further investigate the possible link between clinical outcome and bacterial lineage by focussing on the presence or absence of known virulence genes and those associated with antimicrobial resistance.

Chapter 3: The accessory genome of *Escherichia coli* of patients attending HTD.

3.1. Introduction

In Chapter 2, we have shown that *E. coli* from rectal swabs and blood samples of patients attending HTD are extraordinarily diverse, with a core genome composed of only ~ 3,000 genes but a pan genome in excess of 40,000 genes, most of which are very specific to a subset of isolates (less than 15% of strains). Within this diversity there are likely to be many different non-pathogenic commensals as well as a range of different pathotypes (see Chapter 1). Since none of our rectal samples came from patients with enteric disease, we were not expecting to find intestinal pathogenic *E. coli* (InPEC). However, data described in Chapter 2 showed that there was a significant association between the phylogenies of the isolates sequenced and sample type, with members of phylogroup B2 significantly associated with blood stream infection at HTD. This is consistent with the literature which shows that extraintestinal pathogenic *E. coli* (ExPEC) dominate this phylogroup clinically [108]. Unlike InPEC which are obligate pathogens and have certain distinctive virulence factors that can be used to define them exclusively, ExPEC and commensal *E. coli* are not easily distinguished from one another [10], not least because ExPEC can also be found in healthy individuals. Nevertheless, there are a set of characteristic virulence genes that were extensively discovered by screening across ExPEC *E. coli* (adapted from [108, 109]; see Table 3.1). Of the ExPECs the most common cause of human disease comes from the uropathogenic *E. coli* (UPEC) subdivision. Whilst UPECs are the major cause of urinary tract infections they are often found in the gut of the infected individuals as well as causing more systemic disease [110].

Table 3.1: Common virulence factors identified in ExPEC

Functional category	Virulence factors	Gene detected
(1) Adhesion	Afimbrial adhesin	*afaABCDEF/draP*
	Type 1 fimbriae (*Fim*)	*fimABCDEFGHI*
	P fimbriae (*Pap*)	*papABCDEFGKX*
	S/F1C fimbriae (*Sfa*, *Foc*)	*sfaABHSY, focACDFGH*
	Common pilus	*yag_ecp*
	Adhesin	*fdeC*
	Dispersin (intestinal adhesin in EAEC)	*aap_aspU*
(2) Iron acquisition	Enterobactin	*entABCDEFS,fepABCD,fes*
	Salmochelin	*iroBCDEN*
	Aerobactin	*iucABCD,iutA*
	Yersiniabactin	*fyuA,ybtAEPQSTUX,irp12*
	Hemin uptake system	*chuASTUVWXY*
(3) Toxin	Enterotoxin	*astA,senB*
	Hemolysin	*hlyABCD*
	Cytotoxic necrotising factor 1 (CNF-1)	*cnf1*
	Serine protease autotransporters *Sat*, *Pic*	*sat, pic*
	Haemoglobin protease	*vat*
(4) Invasin	Invasion of brain endothelium	*ibeA*
(5) Protectins	K1 capsule	*kspDMT*
	Tir domain containing protein	*tcpC*

The genes in Table 3.1 were the main virulence determinants responsible for the mechanisms by which uropathogenic *E. coli* (UPEC) can cause disease, including cystitis and pyelonephritis [111]. These functions can be divided into at least five categories, with many of these factors indeed occuring with high redundancy [112]. In order to establish infection, UPEC first needs to attach to host cells through different types of fimbriae. Among them, Type 1 fimbriae, especially *fimH*, are essential to causing infection in UPEC [113]. Toxins such as *hly* and *vat* are used to lyse host cells, both for bacterial

nutrition and to disseminate to deeper tissues. Iron acquisition systems are used to chelate free iron from host, to enable bacterial proliferation [114]. Immune evasion is achieved by producing capsule (*kspDMT*), to protect bacteria from serum bactericidal activity of human plasma [115, 116].

Apart from virulence genes, antimicrobial resistance (AMR) genes are also an important component of the accessory genome since AMR genes confers selective advantage on members of bacterial populations when under antibiotic selection. MDR *E. coli* have also recently been listed by World Health Organization (WHO) as an urgent threat to humanity, and new antibiotics are urgently needed [64]. Currently, many genes mediating resistance to last resort antibiotics such as bla_{NDM} and *mcr-1* are reported to be located on plasmids [58, 117]. *E. coli* infections in Asia are reported as having the highest prevalence of ESBL-producing isolates, mainly by possessing the $bla_{\mathrm{CTX\text{-}M}}$ gene [118]. Tracking and understanding the diversity of AMR genes in *E. coli* causing infections in our collection is required to limit the spread of these bacteria in both community and hospital environments.

Plasmids have long been recognised as the vehicle by which AMR genes and virulence genes spread quickly between bacterial species. Acquisition of virulence and AMR plasmids might change the prevalence of a specific clone in disease-causing bacterial populations [119]. One classic example is that simply by acquiring the 230kb virulence plasmid pINV, which encodes T3SS, both *Shigella spp*. and EIEC can invade cells, live an intracellular lifestyle, and cause diarrheal disease [120]. Other distinct InPEC pathotypes are also mediated by virulence plasmids, such as plasmid pEAF which encodes an adherence factor in enteropathogenic *E. coli* (EPEC); plasmid pAA encoding aggregative adherent fimbria (AAF) in enteroaggrative *E. coli* (EAEC); and colonisation factors, ST and LT toxin can be located on a plasmid in enterotoxigeneic *E. coli* (ETEC) [121]. Plasmids in *Enterobacteriaceae* are diverse in terms of their size and the number of genes they carry. A *rep* gene PCR based method has been developed to classify plasmids as belonging to different Incompatibility (Inc) classes, called replicon typing scheme [122].

However, information about virulence, AMR and plasmid replicon genes are often missing from the core genome analysis because by definition these are not core genes. Thus, it is essential to screen for different components of accessory genomes and to relate them back to the core genome phylogenetic tree to look for genes that differently distributed between them.

Aim:

We hypothesised that specific virulence factors might explain the virulence potential of ExPEC relative to carriage *E. coli*, as well as explaining the successful expansion of several common lineages such as ST131, ST1193, ST73 and ST95. In order to test this hypothesis, we used a systematic approach to scan the *E. coli* genomes in our study for the presence of known and well-defined virulence genes, as well as for genes known to confer antimicrobial resistance. We then compared the distribution of these genes between the randomly sampled isolates cultured from blood and rectal swabs (see Chapter 2 Section

2.2). We used these data to look for any associations between genes and sample site, and to study differences between different lineages and STs.

Specific questions to be answered:

1) Are there any genes significantly associated with disease phenotype (blood/carriage)?

2) How are virulence genes distributed between different lineages?

3) Do AMR genes confer a selective advantage on population of *E. coli* causing invasive disease?

3.2. Methods

3.2.1. *In silico* virulence, AMR, and plasmid replicon type gene detection

ARIBA [123] was used to screen genomes for the presence of antimicrobial resistance (AMR) genes. Briefly, this algorithm mapped raw reads to reference gene sequences, used mapped reads to perform local assemblies, and to report gene presence or absence as well as whether the gene in question existed as a new variant. We also use ARIBA to screen for virulence genes and replicon types. AMR gene sequences were obtained from the ResFinder database [124], virulence gene sequences were taken from the VFDB database [125], and plasmid replicon types from the PlasmidFinder database [119]. PlasmidSPADES [126] was used to identify contigs likely to be plasmids from SPADES assembly graphs [127] and plasmid contigs were visualised using Bandage [128].

3.2.2. Antimicrobial phenotype testing

All isolates were tested for their antimicrobial susceptibility phenotype using 13 antimicrobials including amikacin (30 μg) (AMK), amoxicillin (20 μg) / clavulanic acid (10 μg) (AMC), ceftazidime (30 μg) (CAZ), ceftriaxone (30 μg) (CRO), cefepime (30 μg) (FEP), ciprofloxacin (5 μg) (CIP), ofloxacin (5 μg) (OFX), trimethoprim (1.25 μg) /sulfamethoxazole (23.75 μg) (SXT), imipenem (10 μg) (IPM), ertapenem (10 μg) (ETP), meropenem(10 μg) (MEM), tazocine or piperacillin (100 μg) / tazobactam (10 μg) (TZP), ticarcillin (75 μg) / clavulanic acid (10 μg) (TCC). Antimicrobial susceptibility was determined using the disk diffusion method following Clinical and Laboratory Standard Institute (CLSI) guidelines [129]. An EBSL-producing phenotype was determined by the double disk diffusion method [130], using a combination of third-generation cephalosporin discs, *i.e.* ceftazidime, ceftriaxone, cefepime and amoxicillin (20 μg) / clavulanic acid (10 μg). All phenotypic tests were performed by Microbiology staff at HTD as part of clinical routine laboratory work.

3.2.3. Statistical analysis

Logistic regression was used to calculate the association between a gene and an outcome of being invasive or carriage phenotype in Stata v.13 (StataCorp). The P value was adjusted for multiple comparisons using the Bonferroni correction method. Values of $P < 0.05$ after Bonferroni correction were considered to be significant. Differences in gene distribution in each group were calculated using

Chi squared or Fisher exact test (when observation in contingency table <5). Distribution of total AMR genes and replicon genes between STs/BAPS lineage or blood/carriage were computed using the Wilcoxon test. Gene presence/absence data were visualised in Phandango [99] and plots were produced by ggplot2 package [131] in R (R Foundation).

3.3. Results

3.3.1. The distribution of virulence factors between blood and rectal swab isolates

Using the virulence database 'VFDB' [125], containing 2,600 reference genes across Gram-negative and Gram-positive bacteria, we identified a total of 205 virulence genes across our strain collection, classified into five groups: adhesion, iron acquisition, toxins, invasins and protectins (Table 3.1). By comparing the distribution of these genes between blood and rectal swab isolates our analysis, we identified several genes that might serve as predictors for an invasive phenotype ($OR \geq 2$, $p \leq 0.05$). These include loci related to adhesion (*pap* cluster, *fdeC*), iron acquisition system (yersiniabactin, salmochelin, aerobactin), toxins (*hly*, *vat*) and protectin (*kspD*) (Figure 3.1). One of the most significant associations was between P fimbriae and invasive disease (*papBCDEFGIJK*; $OR \geq 3$, $p \leq 10^{-8}$). P fimbriae mediate the ability to adhere to kidney epithelial cells and are strongly associated with ExPEC and invasive disease [132]. Of the siderophores showing a strong odds ratio for blood stream infections, the highest score was seen for yersiniabactin (*ybt*, *irp*, *fyuA*) (OR 3.3, 95% CI 2.2-5, $p = 4.26 \times 10^{-8}$) followed by salmochelin (*iro*) (OR 2.7, 95% CI 1.64-4.5, $p = 3.5 \times 10^{-4}$) and aerobactin (*iuc*) (OR 2.2, 95% CI 1.6-3.3, $p = 4 \times 10^{-5}$). Enterobactin, an archetypal siderophore system, was identified across all *E. coli* isolates included in this study, and therefore is not considered to be characteristic of ExPEC, consistent with previous studies [133].

We also detect the association between member of serine protease autotransporters (SPATE) family including *sat*, *pic* and *vat*, which were also more likely to be present in invasive isolates than carriage. Each autotransporter has been shown to cause vacuolating of host tissues [112]. Other detected virulence factors have also been well-studied, including hemolysin (*hlyABCD*), *kspMD* which mediates K1 capsule, and *ibeA*, all of which increase the odds of causing invasive disease.

Apart from genes that are enriched in blood isolates, we also found genes/loci that are more likely to be associated with intestinal pathogenic *E. coli* (InPEC) isolates. These included *esp* genes which encode type three secretion system effector proteins and are encoded on the locus of enterocyte effacement (LEE), a pathogeneicity island, characteristic of enterohemorrhagic *E. coli* (EHEC) and EPEC pathotypes, both members of the InPEC (see section 3.1). Included within this locus we detected *espX* which is only known to be found on the LEE in EHEC isolates [134] (Figure 3.2). In addition, we also found many other virulence genes characteristic of O157:H7 such as *cesABDLT*, *etgA*, *nleBCEGH*, *sepDLQ*, *escCDEFGIJLNMPRSTUV* in one patient with a fatal outcome, suggesting that EHEC

O157:H7 infection (ST752; data not shown) is rare in Vietnam, but can subsequently gain access to the bloodstream.

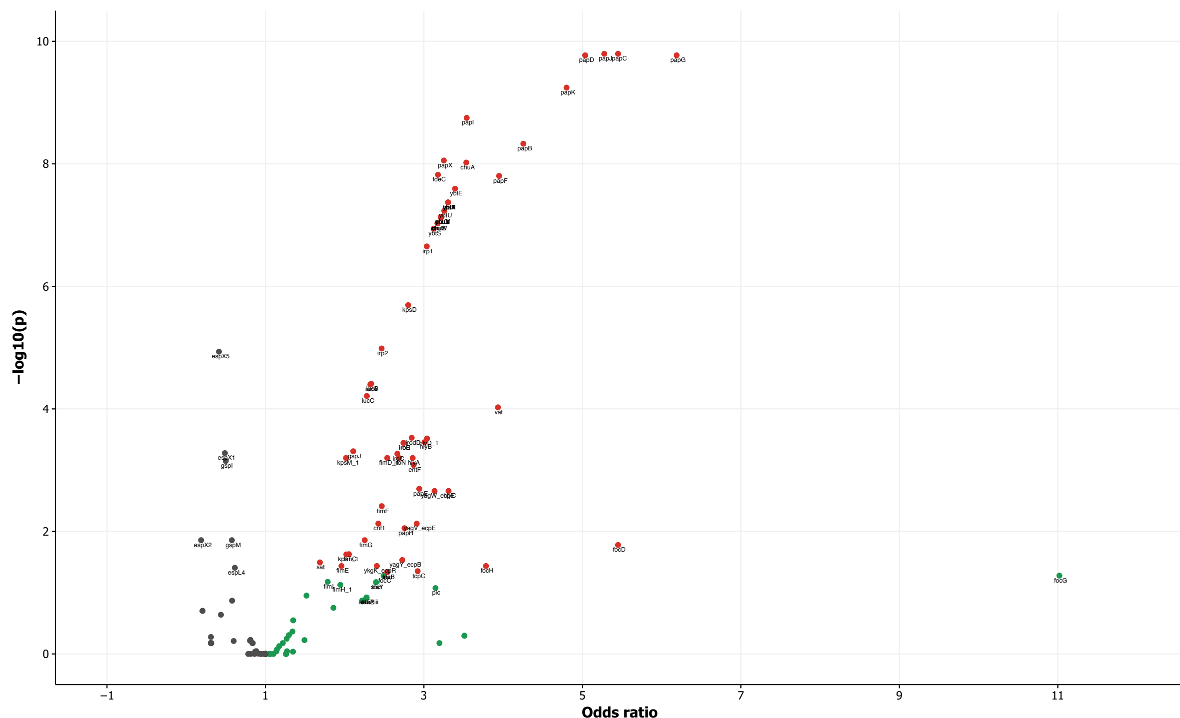


Figure 3.1: Volcano plot for odds ratio versus $-\log_{10}$ p-value of gene distribution between invasive and carriage isolates. The graph shows the odds ratio (x-axis) versus the $-\log_{10}$ of p-value (y-axis) from the results of the logistic regression model. Each dot represents a single gene. OR< 1 indicates that the gene is more likely to be associated with rectal swab isolates (coloured in grey) whereas genes with OR> 1 are more likely to be associated with blood isolates. Genes in red represent the association is statistically significant with $-\log_{10}$(p-value) ≥ 1.3 (*i.e.*, $p < 0.05$). Genes in green indicate that there is an association between tested gene and invasive disease but it is not statistically significant. All analysis was performed in Stata v.13 (StataCorp) and results were plotted using ggplot2 package in R (R computing).

3.3.2. The distribution of virulence factors between different BAPS lineages/ phylogroups/STs

When studying the distribution of virulence genes across lineages, we noticed that genes were clustered according to ECOR phylogroups (Figure 3.2). For examples the *chu* and *pap* genes are absent from phylogroup A (L14)/B1 (L11), whereas *esp* genes are mostly present in phylogroup D (L1,2,3,4,9) and phylogroup A/B1. A certain number of virulence genes are present in all isolates regardless of their lineages or potential pathogenic outcome. These include genes that mediate enterobactin, outer membrane protein A, type 1 fimbriae (*fim*), general secretion pathway (*gsp*) on T3SS, and common pilus (*yag_ecp*).

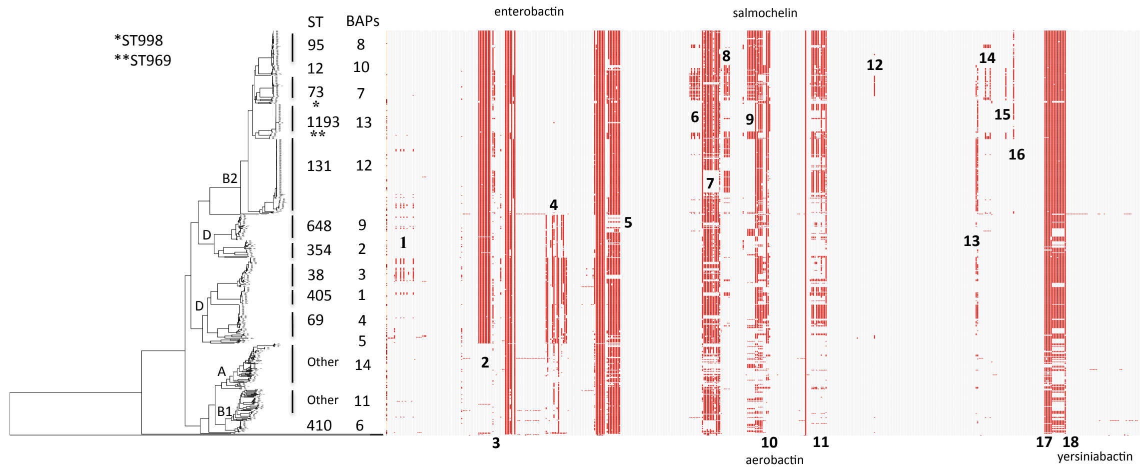


Figure 3.2: Phylogeny of *E. coli* lineages and virulence genes found in each isolate. Red dots represent gene presence; white indicates gene absence. The tree, rooted on *E. albertii*, includes STs and BAPS lineage information. Key gene differences within and between STs were highlighted (1 to 18). (1) Afimbrial adhesion (*afa*/*draP*), (2) Hemin uptake system (*chu*), (3) *papB*, (4) T3SS effector (*esp*), (5) adhesion (*fim*), (6) adhesion (*foc*), (7) T2SS general secretion pathway protein (*gsp*), (8) hemolysin (*hly*), (9) salmochelin (*iro*), (10) aerobactin (*iuc*, *iut*), (11) adhesion (*pap*), (12) autotransporter (*pic*), (13) toxin (*sat*, *senB*), (14) adhesion (*sfa*), (15) protectin (*tcp*), (16) toxin (*vat*), (17) adhesion (*yag*-Ecp), (18) yersiniabactin (*ybt*).

Breaking the phylogroups down into ST or BAPS lineages, we looked at the distribution of the virulence genes in more detail across the twelve most prevalent STs in our collection (Table 4.2). Within phylogroup B2, the four dominant STs are ST95 (BAPS L8), 73 (L7), 1193 (L13) and 131(L12). Although ST131 is a globally successful pathogenic clone of *E. coli,* ST73 carried eight more virulence genes than ST131, ST95 and ST1193 (Figure 3.2), and several of these genes were exclusive to ST73 in our collection. These included the *foc* gene cluster (*focA* to *focI*), *sfa* cluster (*sfaA*, *sfaF*), *hly* (*hlyA* to *hlyC*), *cnf1*, *pic*, *sat*, *vat* and *tcpC* (Table 4.2). ST73 in our study had more *Pap* genes than ST131, as well as having additional fimbriae such as S1 fimbriae (*sfa* operon) and F1C fimbriae (*foc* operon). Interestingly other rare STs in phylogroup B2, including ST12/ST969/ST998 (L10), appear to possess the same virulence profile as ST73 and so likely to be capable of causing invasive disease.

Even within a specific ST, we observed the gain and loss of several specific genes. For example, within ST131, not all isolates (62%) possess P fimbriae, and those that lack P fimbriae may harbour *afa/draP* (9.4%) and to a lesser extent *sfa/foc* (1.7%). In addition to having genes related to colonization, 50% to 80% of our ST131 isolates carry gene that encode toxins including *hly*, *sat*, *senB* and *cnf*-1 (Figure 4.2, Table 4.2). Of the other STs carrying *pap* genes 12% of the member of ST38 in our collection harbour

papG. In those that lack *papG*, 63.4% utilise afimbrial adhesin (*afa*) and Dr fimbrae (*draP*) which can bind to uroepithelial cells in the same way as P fimbriae. 12% of ST38 also use dispersin (*aap*/*aspU*), a protein mediated colonization factor characteristic of enteroaggrative *E. coli* pathotypes (EAEC) for binding.

Table 3.2: The percentage of isolates carrying specific virulence genes in the 12 most dominant sequence types of *E. coli* found in this study. Total represents the percentage of all isolates among the 665 isolates that possess the gene/locus.

Functional category	**Gene cluster**	**Genes**	**Phylogroup**													
			ST													
			10	**1193**	**12**	**131**	**354**	**38**	**405**	**410**	**648**	**69**	**73**	**95**	**other**	**Total (%)**
Adhesion	*afaABCDEF/draP*	*afaACDF*	0	0	0	9.4	0	63.4	14.3	0	13.9	2.6	0	0	3.8	8.3
		afaE1	0	0	0	2.6	0	2.4	14.3	0	8.3	0	0	0	0.5	1.8
		afaE2	0	0	0	0	0	12.2	0	0	0	0	0	0	0	0.8
		draP	0	0	0	9.4	0	63.4	14.3	0	13.9	2.6	0	0	3.8	8.3
	Type 1 fimbriae (Fim)	*fimA*	25	0	50	0.9	0	73.2	0	71.4	25	0	97.1	0	31.1	23.8
	fimABCDEFGHI	*fimB1*	100	95.7	83.3	18.8	73.3	82.9	100	100	30.6	100	100	82.8	79.2	71.6
		fimC	100	95.7	83.3	100	80	82.9	100	100	30.6	100	100	100	85.4	89.5
		fimD1	91.7	95.7	66.7	100	80	75.6	100	100	30.6	100	100	100	82.1	87.5
		fimE	100	95.7	83.3	99.1	80	82.9	100	100	30.6	100	100	100	86.3	89.6
		fimF	100	93.6	75	100	86.7	82.9	100	100	44.4	100	100	100	89.2	91.3
		fimG	100	95.7	83.3	100	86.7	82.9	100	100	44.4	100	85.3	100	89.2	90.8
		fimH	91.7	95.7	83.3	99.1	86.7	82.9	100	100	44.4	100	100	100	92	92.2
		fimI	100	95.7	83.3	100	80	82.9	100	92.9	30.6	100	100	100	81.6	88.1
	P fimbriae (Pap)	*papA*	0	0	0	0.9	0	7.3	21.4	0	0	0	2.9	0	0.5	1.8
	papABCDEFGKX	*papB*	0	23.4	91.7	45.3	0	17.1	71.4	21.4	8.3	79.5	85.3	94.8	21.7	40.5

		papC	0	27.7	100	60.7	0	12.2	50	7.1	2.8	79.5	97.1	94.8	17.5	41.1
		papD	0	27.7	100	62.4	0	12.2	53.6	7.1	2.8	79.5	97.1	94.8	17.9	41.7
		papE	0	4.3	25	1.7	0	19.5	0	0	19.4	5.1	44.1	93.1	11.8	17.7
		papF	0	27.7	75	62.4	0	22	53.6	7.1	16.7	82.1	94.1	94.8	9.9	40
		papG	0	27.7	83.3	62.4	0	12.2	53.6	7.1	0	76.9	97.1	93.1	7.5	37.6
		papH	0	4.3	66.7	4.3	0	7.3	25	0	2.8	74.4	79.4	3.4	6.1	14.6
		papI	16.7	93.6	91.7	91.5	0	48.8	57.1	7.1	22.2	76.9	85.3	94.8	23.1	55.9
		papJ	0	27.7	100	62.4	0	12.2	53.6	7.1	2.8	79.5	97.1	94.8	17.5	41.5
		papK	0	27.7	91.7	61.5	0	12.2	53.6	7.1	2.8	82.1	97.1	94.8	15.6	40.8
		papX	16.7	100	75	82.1	60	78	89.3	7.1	75	84.6	82.4	98.3	25	63
	S/F1C fimbriae (Sfa,Foc)	*sfaA*	0	0	0	0	0	0	0	0	2.8	0	0	8.6	1.9	1.5
	sfaABHSY	*sfaB*	0	0	91.7	1.7	0	0	0	0	2.8	0	85.3	8.6	9.4	10.2
		sfaH	0	0	0	0	0	0	0	0	2.8	0	14.7	8.6	4.7	3.2
		sfaS	0	0	0	0	0	0	0	0	2.8	0	0	8.6	4.2	2.3
		sfaY	0	0	91.7	1.7	0	0	0	0	2.8	0	79.4	8.6	9	9.8
	focACDFGH	*focA*	0	0	16.7	0	0	0	0	0	0	0	76.5	0	1.4	4.7
		focC	0	0	91.7	1.7	0	0	0	0	2.8	0	82.4	8.6	9.4	10.1
		focD	0	0	16.7	0.9	0	0	0	0	2.8	0	85.3	8.6	6.1	7.7
		focF	0	0	91.7	1.7	0	0	0	0	2.8	0	85.3	8.6	9.4	10.2
		focG	0	0	16.7	0	0	0	0	0	0	0	82.4	0	1.9	5.1

		focH	0	0	91.7	0	0	0	0	0	0	0	79.4	0	5.2	7.4
		focI	0	0	91.7	1.7	0	0	0	0	2.8	0	76.5	8.6	9.4	9.8
	Common pilus	*yagv_ecpE*	83.3	100	100	100	93.3	100	100	92.9	100	100	97.1	100	84.9	94.4
	yag_ecp	*yagw_ecpD*	66.7	100	100	100	93.3	100	100	92.9	100	100	100	100	84	94
		yagx_ecpC	83.3	100	100	100	93.3	100	100	92.9	100	100	100	100	83.5	94.1
		yagy_ecpB	83.3	100	100	100	93.3	100	100	92.9	100	100	100	100	88.2	95.6
		yagz_ecpA	83.3	100	100	100	93.3	100	100	92.9	100	100	100	100	85.4	94.7
	Adhesin	*fdeC*	50	97.9	100	97.4	0	100	100	64.3	0	0	100	100	41	65.4
	Dispersin (intestinal adhesin in EAEC)	*aap_aspU*	0	0	0	0	0	12.2	0	0	0	0	0	0	6.1	2.7

Functional category	**Gene cluster**	**Genes**	**ST**													
			10	**1193**	**12**	**131**	**354**	**38**	**405**	**410**	**648**	**69**	**73**	**95**	**other**	**Total (%)**
Iron acquisition	Enterobactin (*entABCDEFS*)	*entABCES*	100	100	100	100	100	100	100	100	100	100	100	100	100	100
		entD	100	0	100	0	100	100	100	100	100	97.4	100	100	74.1	66.9
		entF	0	0	0	82.1	0	0	0	0	0	0	100	0	4.2	20.9
	fepABCDG	*fepA*	100	100	100	100	100	100	100	100	100	100	100	55.2	97.6	95.3
		fepBCDG	100	100	100	100	100	100	100	100	100	100	100	100	99.5	99.8
	Salmochelin (*iroBCDEN*)	*iroBE*	16.7	0	91.7	0	0	2.4	0	35.7	33.3	12.8	82.4	96.6	23.6	25.6
		iroCD	16.7	0	91.7	0	0	2.4	0	35.7	30.6	12.8	82.4	94.8	23.1	25.1
		iroN	16.7	0	75	0	0	0	0	35.7	33.3	12.8	82.4	89.7	22.6	24.2
	Aerobactin (*iucABCD*, *iutA*)	*iucABC*	66.7	97.9	91.7	97.4	13.3	58.5	71.4	92.9	80.6	92.3	58.8	96.6	33.5	67.7
		iucD	8.3	0	83.3	0.9	0	19.5	25	0	11.1	69.2	2.9	0	11.8	12.6
		iutA	8.3	0	83.3	0.9	0	19.5	21.4	0	13.9	69.2	2.9	0	11.8	12.6
	Yersiniabactin (*fyuA*, *ybtAEPQSTUX*, *irp12*)	*fyuA*	50	100	100	100	40	87.8	100	71.4	58.3	94.9	100	100	51.4	78.3
		ybtAEPQSTUX	50	100	100	100	40	87.8	100	71.4	58.3	94.9	100	100	51.4	78.3
		irp1	50	100	91.7	100	40	87.8	100	71.4	58.3	94.9	100	96.6	50	77.4
		irp2	50	0	91.7	94	40	87.8	100	71.4	52.8	92.3	94.1	94.8	48.6	68
	Hemin uptake system (*chuASTUVWXY*)	*chuASTUVWXY*	0	100	100	100	100	100	100	0	100	100	100	100	44.8	78.5

Invasin	Invasion of brain endothelium	*ibeA*	0	0	0	0.9	73.3	0	0	0	0	0	0	12.1	7.5	5.3
Protectins	K1 capsule	*kspD*	16.7	100	91.7	98.3	100	100	64.3	0	91.7	82.1	97.1	100	51.9	77.4
		kspM	16.7	100	91.7	59.8	33.3	80.5	42.9	0	86.1	64.1	85.3	98.3	42.5	62
		kspT	16.7	100	0	0.9	0	0	0	0	0	0	2.9	100	11.8	20.2
	Tir domain containing protein	*tcpC*	0	0	83.3	0	0	0	0	0	0	0	97.1	0	7.1	8.7

Functional category	**Gene cluster**	**Genes**	**ST**													
			10	**1193**	**12**	**131**	**354**	**38**	**405**	**410**	**648**	**69**	**73**	**95**	**other**	**Total (%)**
Toxin	Enterotoxin	*astA*	16.7	0	8.3	2.6	13.3	14.6	0	7.1	5.6	0	5.9	0	20.3	9.3
		senB	25	85.1	58.3	80.3	0	34.1	32.1	21.4	0	66.7	44.1	0	13.2	35.9
	Hemolysin	*hlyAB*	0	4.3	91.7	55.6	0	7.3	21.4	0	0	5.1	91.2	0	12.3	22
		hlyC	0	4.3	75	54.7	0	7.3	21.4	0	0	2.6	44.1	0	0.9	15.3
	Cytotoxic necrotizing factor 1 (CNF-1)	*cnf1*	0	4.3	66.7	53.8	0	0	3.6	0	0	2.6	88.2	1.7	8.5	18.6
	Serine protease autotransporters Sat, Pic	*sat*	16.7	0	0	96.6	0	53.7	32.1	0	13.9	76.9	61.8	0	12.3	34.3
		pic	0	0	0	0	0	0	0	0	0	0	94.1	3.4	3.8	6.3
	Haemoglobin protease	*vat*	0	66	83.3	0	0	0	0	0	0	0	85.3	63.8	13.2	20.3

3.3.3. Antimicrobial resistance phenotype between blood/rectal swab isolates

We tested for phenotypic resistance to 13 antimicrobials belonging to five classes, including aminoglycosides, 3rd and 4th – generation cephalosporins, β-lactam plus inhibitor, carbapenem, trimethoprim and fluoroquinolones. Isolates exhibited the highest resistance to cotrimoxazole (STX) with 70% and 75% isolates in blood and carriage isolates respectively. Ninety-nine percent of isolates were sensitive to carbapenem (ertapenem, imipenem, meropenem) and aminoglycosides (amikacin). Resistance to 3rd – generation cephalosporins, which are used as empirical treatments against suspected BSIs in HTD, was ~ 50% (Figure 3.3).

Looking at the dominant STs, ST131 was the most prevalent ST in our blood collection (100/506, 19.76%). Its global spread has been partially attributed to its drug resistance profile that includes resistance to fluoroquinolones and third-generation β-lactams [135]. Of note the newly-emerging ST1193 clone seen in HTD, which accounted for 7.11% of cases (36/506), has the same resistance profile as ST131. Nevertheless, ST95 (57/506, 11.26%) and ST73 (34/506, 6.72%), the second and fifth prevalence STs, are generally susceptible to most antibiotics (See supplementary Tables S1 and S2 for complete list of resistance phenotypes).

Looking at all samples we observed no statistical significance between AMR phenotypes across blood and rectal swab isolates, except for combinations of β-lactam antimicrobials that include an inhibitor (amoxicillin-clavulanic acid, ticarcillin-clavulanic acid and piperacillin-tazobactam; highlighted in Figure 3.3). The most dramatic difference was seen in the case of piperacillin-tazobactam (TZP), a drug combination of extended spectrum penicillin (piperacillin) and tazobactam; tazobactam acts as an inhibitor of the β-lactamase enzyme. Fewer than 20% of BSI isolates showed reduced susceptibility to TZP and were classed as resistant, whereas >50% of all rectal swab isolates were classed as phenotypically resistant (Figure 3.3). This is important because TZP and carbapenem are used as an alternative treatment for ESBL-producing *E. coli* BSI at HTD.

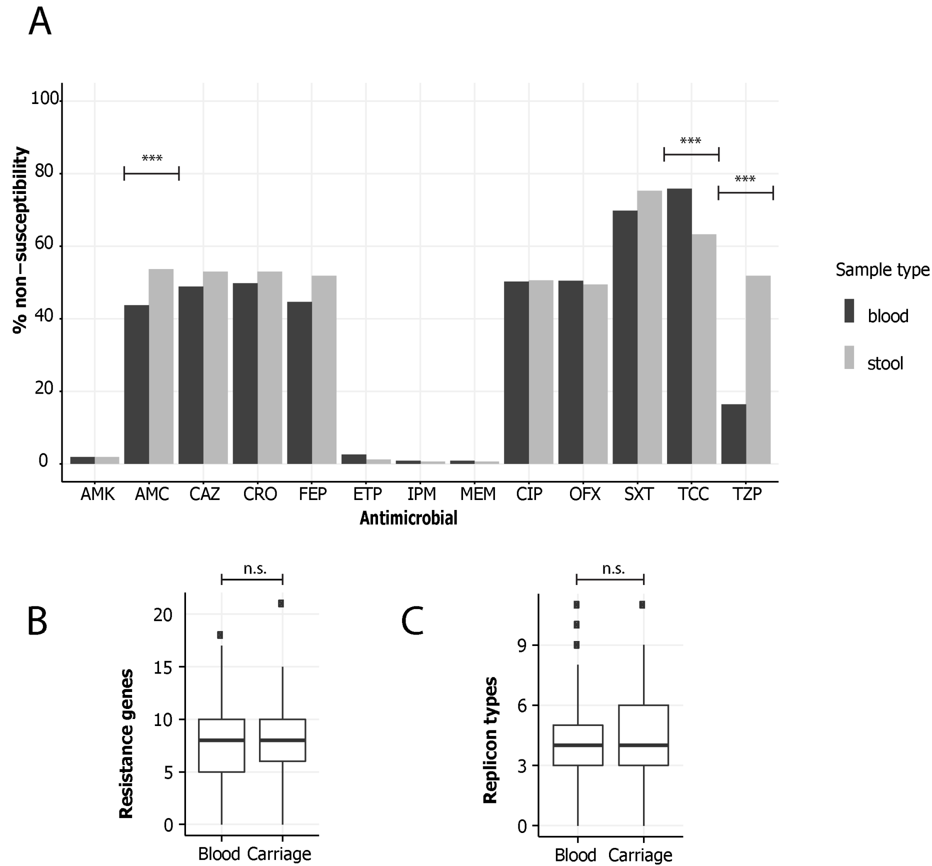


Figure 3.3: (A) The distribution of AMR resistance between blood and carriage isolate. Asterisks highlight the three antimicrobials showing differences in the levels of resistance between BSI isolates and rectal swab. (B) Total resistant genes and (C) total replicon genes found in two groups. n.s: not significant, *** $p < 0.001$; p-values were calculated using Wilcoxon test for continuous data and Chi square test for category data. Antimicrobial abbreviations: amikacin (AMK), amoxicillin-clavulanic acid (AMC), ceftazidime (CAZ), ceftriaxone (CRO), cefepime (FEP), ertapenem (ETP), imipenem (IPM), meropenem (MEM), ciprofloxacin (CIP), ofloxacin (OFX), trimethoprim-sulfamethoxazole (SXT), ticarcillin-clavulanic acid (TCC), piperacillin-tazobactam (TZP).

3.3.4. Antimicrobial resistance genes distribution in blood/ rectal swab

We screened *E. coli* genomes in this study for resistance genes and plasmid replicons that are associated with AMR to evaluate the diversity of AMR genes and plasmids in our *E. coli* collection. We found a total of 54 genes that have the potential to confer resistance to not only drugs against which these isolates exhibit a resistant phenotype (β -lactam, fluoroquinolone, *etc.*), but also to members of other classes including rifampicin, chloramphenicol, macrolide, fosfomycin, lincosamide, tetracyclin, streptomycin and colistin (Figure 3.5). The general lack of difference in phenotypic AMR profile between blood and carriage isolates was supported by comparing the total number of AMR genes and

plasmid replicon genes in BSI and rectal swab isolates: overall there was no statistical significant difference (median 8 vs 8, Wilcoxon test) (Figure 3.3).

Consistent with the phenotypic data showing that the majority of our isolates remained sensitive to carbapenem - we only found three isolates harbouring *bla*$_{NDM-4}$ or *bla*$_{NDM-5}$, and the presence of these genes corresponded to the resistance phenotype of the specific isolate (Supplementary Table S6). Although there was no difference in the profile of genes when comparing blood and rectal swab derived isolates overall, the total number of AMR genes and replicon types are substantially different between STs/BAPS lineage (Figures 3.4 & 3.5). ST131 and ST1193 contained more AMR genes than "susceptible" STs such as ST73 and ST95 (median 9 *versus* 6, p= 4.07 x 10^{-06}, Wilcoxon test). However, ST131 contain less replicon types than ST1193 (median 4 *versus* 6, p= 4.463 x 10^{-11}, Wilcoxon test).

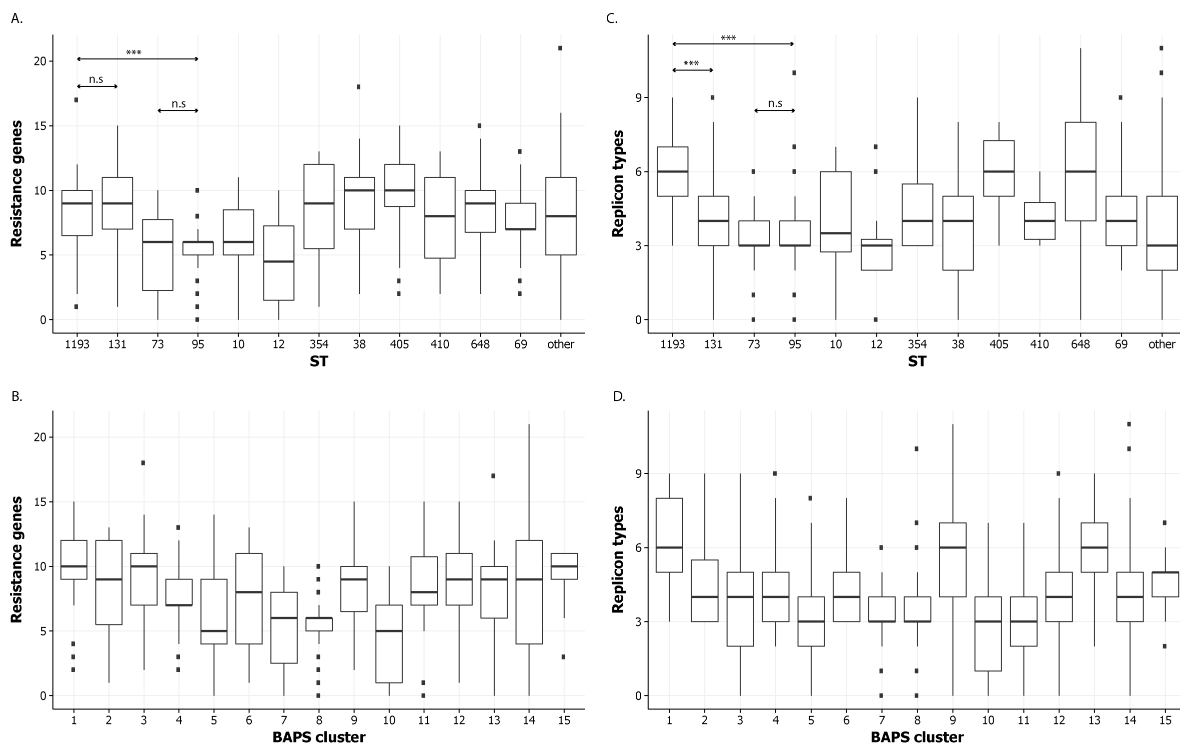


Figure 3.4: Total number of resistance genes and replicon types distributed across the most prevalent STs and among all 15 BAPS lineages in this dataset. Resistance genes found in different STs (A) and BAPs lineages (B). Total replicon types found in STs (C) and BAPs lineages (D). n.s: not significant, *** $p<0.001$ (Wilcoxon test)

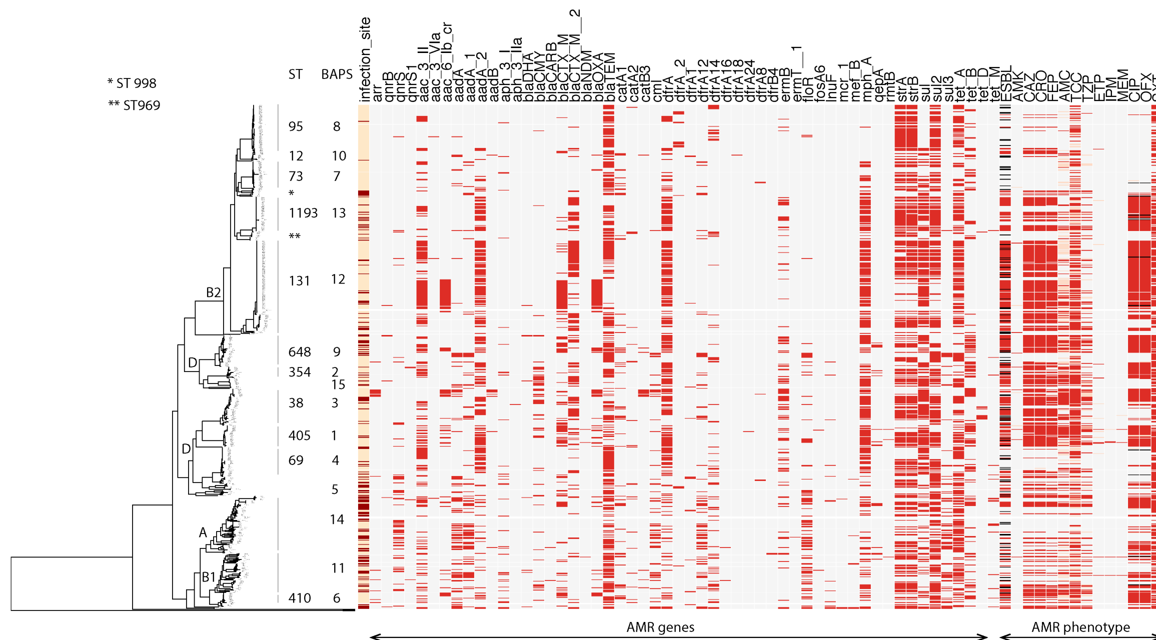


Figure 3.5: AMR genotypes and phenotypes across all *E. coli* isolates in this collection. Columns report infection site (blood/carriage, see key in Fig 2.7), red is present while white is absent, black were missing phenotype data (ESBL, CIP, OFX) since ESBL phenotypes were not assessed for isolates from 2010. Phylogroups are denoted on the branches of the phylogenetic tree.

Since extended spectrum β-lactams are one of the most important classes of antimicrobial at HTD we looked for ESBL-genes across all lineages (Table 3.3). We found that the majority of ESBL-producing isolates belonging to ST1193 carried *bla*$_{\mathrm{CTX-M-27}}$ (27/31, 87%), whereas ST131 isolates carried different *bla*$_{\mathrm{CTX-M}}$ gene classes such as *bla*$_{\mathrm{CTX-M-14}}$ (33/106, 31%), *bla*$_{\mathrm{CTX-M-15}}$ (32/106, 32%) or *bla*$_{\mathrm{CTX-M-27}}$ (22/106, 20.8%). However, a small number of ST131 isolates contained multiple *bla*$_{\mathrm{CTX-M}}$ genes and ST131 isolates carried the greatest range of ESBL genes, suggesting that this genetic background has the ability to acquire and maintain multiple plasmids containing multiple *bla*$_{\mathrm{CTX-M}}$ genes. This contrasted with ST1193 which almost exclusively carried *bla*$_{\mathrm{CTX-M-27}}$. The absence of multiple classes A (cefotaximase (CTX-M), temoneira (TEM) and sulfhydryl variable (SHV); Figure 3.5) ESBLs was also noticeable in other lineages such as the STs within BAPS lineages L10, L8 and L7. These lineages carried the *bla*$_{\mathrm{TEM}}$ ESBL gene but, apart from essentially sporadic isolates, lacked other class A ESBL genes (Figure 3.5 and Figure 2.7 for the BAPS lineages). Importantly it is noticeable that these lineages were also largely susceptible to TZP as well as many other front-line antimicrobials (Figure 3.5).

We also detected the newly discovered *mcr*-1 in three isolates from patients attending HTD - one from an invasive isolate (ST617) and two from consecutive rectal swabs of day 0 and day 5 in one patient in ICU (Supplementary Table S6). Looking at the assembled genome data it was clear that the *mcr*-1 gene in the rectal swab isolates were located on an IncI plasmid, with the same structure as plasmid pHNSHP45 by Liu *et al* [136] (data was confirmed by visualizing *mcr-1* genes on contigs in Bandage

[128]). However, in the blood isolate, this gene was flanked by *IS*Apl1transposon and the *tnpA* gene on a large chromosomal contig, suggesting that the gene might be chromosomally integrated rather than plasmid-borne (Figure 3.6). Chromosomally integrated *mcr*-1 has been reported recently in clinical *E. coli* isolates from Vietnam and *E. coli* isolated from retail chicken meat in Switzerland - these reports highlight the easy mobilisation of *mcr-1* across different genetic locations [137, 138].

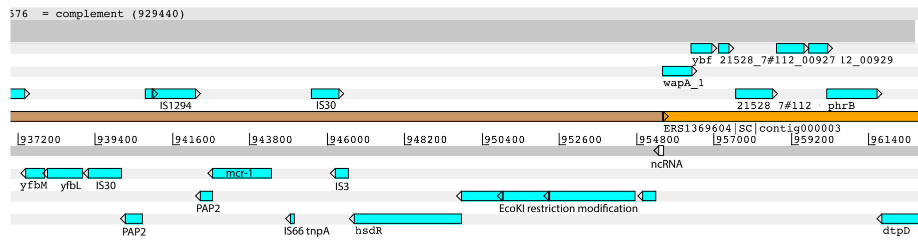


Figure 3.6: The genetic environment of *mcr-1* on a chromosomal contig from isolate 221214_63810, visualized using Artemis [139]

Table 3.3: *bla* CTX-M distribution among different STs (n = number of isolates)

	ST (n)												
CTX-M genes	**10**	**1193**	**12**	**131**	**354**	**38**	**405**	**410**	**648**	**69**	**95**	**other**	**Total**
*bla*CTX_M_15 + *bla*CTX_M_14	0	0	0	7	0	0	0	0	0	0	0	0	7
*bla*CTX_M_15 + *bla*CTX_M_24	0	0	0	0	0	0	1	0	0	0	0	0	1
*bla*CTX_M_15 + *bla*CTX_M_27	0	1	0	5	0	0	1	0	1	0	0	2	10
*bla*CTX_M_15 + *bla*CTX_M_45	0	0	0	2	0	0	0	0	0	0	0	0	2
*bla*CTX_M_15	1	1	2	34	1	4	15	5	8	1	1	24	97
*bla*CTX_M_3	0	0	0	0	0	0	0	0	1	0	0	0	1
*bla*CTX_M_55 + *bla*CTX_M_24	0	0	0	1	0	0	0	0	0	0	0	0	1
*bla*CTX_M_55	0	2	0	1	0	2	1	1	5	1	0	10	23
*bla*CTX_M_14	0	0	1	33	2	3	1	2	4	2	0	7	55
*bla*CTX_M_24	0	0	0	1	3	2	2	0	0	0	0	1	9
*bla*CTX_M_27	0	27	0	22	1	13	4	0	2	4	1	25	99
Total isolates	1	31	3	106	7	24	25	8	21	8	2	69	305

Table 3.4: Plasmid replicon type identified in *E. coli* collection. (n = number of isolates)

	ST (n)													
Replicon	**10**	**1193**	**12**	**131**	**354**	**38**	**405**	**410**	**648**	**69**	**73**	**95**	**other**	**Total**
col156	3	46	0	97	1	21	13	4	19	28	15	4	34	285
col8282	2	8	1	15	2	2	3	0	13	7	1	4	13	71
colE10	0	0	0	0	0	0	0	0	0	0	0	0	3	3
colRNAI	6	14	1	14	4	2	1	8	5	4	14	31	51	155
col_BS512	2	44	1	10	3	3	15	0	10	7	3	3	21	122
col_BS512_1	2	44	1	10	3	4	18	0	10	6	2	1	21	122
col_kphs6	0	0	0	0	0	0	0	0	0	0	0	0	4	4
col_mg828	0	9	1	10	2	1	12	1	23	7	6	9	12	93
col_mp18	1	1	0	1	1	0	0	0	0	0	1	0	2	7
colPVC	1	0	0	0	0	0	1	0	4	2	0	0	5	13
incA/C	0	0	0	0	0	0	0	0	0	0	0	0	2	2
incB/O/K/Z	2	2	2	13	2	0	2	0	5	4	0	3	6	41
incFI	8	3	9	69	2	31	24	14	21	26	18	55	127	407
incFIA	3	47	3	69	14	18	20	6	29	11	9	19	43	291
incFIA_hi1	0	0	0	0	0	0	1	0	1	0	0	0	17	19

incFIB	0	0	0	1	11	3	1	0	5	0	2	0	2	25
incFIB_ap001918	8	45	11	108	4	33	25	13	26	38	22	56	120	509
incFIB_K	0	0	0	0	0	0	0	0	0	0	1	1	19	21
incFIB_phcm2	0	0	0	0	0	0	0	0	0	0	0	0	2	2
incFIB_plf82	0	0	0	1	0	1	0	0	0	0	0	0	2	4
incFII_p	0	0	3	53	0	9	9	0	12	1	6	1	29	123
incHI1a	0	0	0	0	0	0	0	0	1	0	0	0	6	7
incHI1b	0	0	0	0	0	0	0	0	1	0	0	0	4	5
incHI2	0	0	0	0	0	1	0	0	1	1	0	0	4	7
incHI2a	0	0	0	0	0	1	0	0	1	1	0	0	5	8
incI1	2	8	2	7	7	10	8	6	4	0	0	1	42	97
incI2	0	0	0	1	0	1	0	0	1	0	0	0	5	8
incL_M	0	0	0	0	0	0	1	0	1	0	0	0	1	3
incL_M_pmu407	0	0	0	0	0	0	0	0	1	0	0	0	1	2
incN	0	0	0	2	0	1	2	1	0	1	0	0	6	13
incN1	0	0	0	0	0	0	0	0	0	0	0	0	1	1
incQ1	0	6	3	3	8	0	1	1	4	9	2	14	25	76
incR	0	1	0	0	0	0	0	0	0	1	0	0	12	14

incX1	3	2	0	0	0	0	0	0	5	3	0	2	30	45
incX1_1	3	1	0	0	0	0	0	0	2	0	0	1	27	34
incX2	0	0	0	0	0	0	0	0	0	0	0	0	1	1
incX3_pec14	1	0	0	0	0	0	0	0	0	0	0	0	0	1
incX4	0	0	0	2	0	3	0	0	0	0	0	0	3	8
incX4_1	0	0	0	0	0	0	0	0	0	0	0	0	2	2
incY	0	0	0	2	3	4	5	2	3	2	2	3	17	43
p0111	1	1	0	6	0	1	11	1	3	0	0	2	15	41
psl483	0	0	0	0	0	0	0	0	1	0	0	0	1	2
Total	12	47	12	117	15	41	28	14	36	39	34	58	212	665

3.4. Discussion

In this chapter, we compared the distribution of different virulence related and antimicrobial resistance gene contents for the randomly sampled BSI and rectal swab isolates sequenced in this study. The most significant association between its presence and the ability to cause invasive disease was for P fimbriae (OR > 3, $p \leq 10^{-8}$). P fimbriae is an important colonization factor that enhances early establishment of *E. coli* in the urinary tract. P fimbriae are often associated with UPEC causing pyelonephritis. The reciprocal regulation of P fimbriae, type 1 fimbriae and flagella help UPEC to ascend from bladder to kidneys [132]. However, the correlation between P fimbriae and virulence remains inconclusive in literature [30].

After attachment to host tissue and following colonisation, in order to grow and proliferate, *E. coli* requires a siderophore system to acquire iron from the host, where free iron is maintained in an incredibly small concentration (10^{-24} M) [133]. Four siderophore systems have been studied in *E. coli*: enterobactin (*fepABCDG*, *entDFCEBA*, *fes*, *entS*), salmochelin (*iroNEDCB*), yersiniabactin (*ybt*, *fyuA*, *irp*) and aerobactin (*iucABCD*, *iutA*). The human host has many strategies to prevent heme acquisition from pathogenic bacteria, such as Lipocalin-2 which can bind and inactivate bacterial enterobactin [142] and IL-22 which induces the production of the heme scavenger haemopexin, further limiting host iron availability [143]. Therefore, possessing more than one siderophore system simultaneously may give an *E. coli* isolate increased fitness and a competitive advantage relative to other *E. coli* strains.

Several virulence genes showed lineage restriction. Phylogroup B2 isolates which include ST131, ST73, ST95 and ST1193 carried the highest number of genes characteristic of ExPECs, as would be expected from previous studies. This is consistent with data shown in Chapter 2 which showed an association between Phylogroup B2 and gender as well as BSIs. Women are known to be highly susceptible to UTI infections by uropathogenic members UPEC members of the ExPEC. However, it is clear that different lineages employ different fimbrial systems for attaching and colonising the host, and may not necessarily use the same pili or retain more than one due to possible functional redundancy. ST73 in our study possessed more *Pap* genes than ST131, as well as having additional fimbriae such as S1 fimbriae (*sfa* operon) and F1C fimbriae (*foc* operon). This might enhance its ability to bind to human bladder and kidney epithelium cells [144] as well as to form biofilms [145]. The most striking observation was that *tcpC* can only found in ST73 (97%) and in some less prevalent STs such as ST12, ST969 and ST998. TcpC is a secreted protein that binds to myeloid differentiation factor 88 (MyD88) and impedes Toll-like receptor (TLR4) signalling pathways in innate immunity, therefore promoting intracellular bacterial survival of pathogen [146]. The ability to colonise different cell types, together with having a sophisticated molecular strategy such as *tcpC* to subvert host defence, and harbouring a toxin such as *cnf1* to cause cell apoptosis [147], might explain why ST73 was able to emerge as a UPEC clone, with potential to translocate into the blood.

The high prevalence of AMR genes in both BSI and carriage isolates highlights the urgent threat posed by multidrug resistant *E. coli* that are resistant to all major antibiotics, and the challenge that the clinical management of community acquired *E. coli* BSIs presents. Luckily genes conferring resistance to last line antimicrobials, bla_{NDM} and *mcr*-1, were rare and only sporadically distributed in our *E. coli* collection at the time of testing.

One other important observation made here was for TZP. TZP is an intravenous combination antibiotic that is not available in the community. TZP, like clavulanic acid, is a combination therapy including a β-lactam inhibitor (tazobactam), inhibiting the β -lactamase produced by ESBL-producing bacteria, so that the effectiveness of ESBL antibiotic combined with it, in this case piperacillin, is preserved. However, TZP is more efficient than clavulanic acid at inhibiting a wider range of ESBLs. Resistance mechanisms for TZP are not well defined but the efficiency of the inhibition is thought to be affected by the type of the ESBL produced, since the inhibitor binds to and blocks the enzyme active site [148]. Sugumar *et al*., (2014) have suggested that possession of bla_{OXA} in combination with other *bla* genes leads to high resistance to the Tazobactam treatment. Consistent with this from looking at the carriage isolates there was a significantly higher proportion of bla_{OXA} and bla_{CMY} genes present in TZP resistant isolates than TZP sensitive isolates (25.1% *versus* 8%; 20.4% *versus* 5.2%, $p < 0.001$, Chi square test) suggesting that Tazobactam is less efficient at inhibiting the products of bla_{OXA} and bla_{CMY} than other genes such as bla_{TEM} (48.5% *versus* 61.8%, p=0.002, Chi square test). Figure 3.5 also shows that phylogroup B2, BAPS lineages L10, L8 and L7 (ST95, ST73 and ST1193) are dominated by isolates carrying bla_{TEM} genes with only sporadic possession of other *bla* genes. Phenotypically these isolates were largely susceptible to TZP (Figure 3.5) and these susceptible isolates represent a significant proportion (28%) of all BSI isolates sampled in this study. This may explain why the proportion of blood isolates showing resistance to TZP was significantly lower than rectal swab isolates (Figure 3.3), which are more strongly associated with more diverse, less frequently sampled STs from lineages possessing multiple types of *bla* gene. It is not clear if this could be associated with the intravenously delivery with blood isolates exposed to high therapeutic doses compared to gut isolates due to differential drug penetration. Although as we have shown common BSI STs are present in rectal swab samples.

If the TZP-sensitive ST131 isolate numbers (87/100) from blood (the most prevalent ST causing BSI at HTD) are combined with ST95, ST73 and ST1193 then these 4 STs represent 48% of all BSIs sampled in this study. What's difficult to explain from the above is that a high proportion of ST131 isolates carry bla_{OXA} in combination with other *bla* genes. Within the ST131 TZP-sensitive isolates from blood 61 and 26 are bla_{OXA} negative or positive, respectively. Of the ST131 TZP-resistant isolates 6 and 7 are bla_{OXA} negative or positive, respectively. In addition, all of the bla_{OXA} positive isolates also possess either bla_{CTX} or bla_{TEM}. What is more, by performing a genome wide association study using SCOARY [149] which takes the gene presence/absence data from ROARY (see 2.2), then uses the phenotype to

work out gene distribution between 2 groups (TZP resistant or sensitive), in addition to *bla* genes several other genes showed a positive association, including *acrAB* (multidrug efflux pump), *mphA* (multidrug transporter), *ompC* (outermembrane protein linked to resistance) and the multiple antibiotic resistance protein MarB. Hence, TZP resistance appears to also be associated with general mechanisms of resistance and likely to be multifactorial like macrolide resistance. This may imply genetic background is as important as the possession of specific *bla* genes for resistance.

What is clear is that there is that the B2 lineages that were dominant in our BSI isolate collection remain significantly more susceptible to TZP than other phylogroups more associated with rectal swab. We have shown that once established in a community the relative proportion of different STs is stable over time. If we can use this data to identify the less well understood mechanisms of resistance, such as to TZP, we may be able to reach a more proactive state matching any changes in disease patterns in a hospital with flux in the bacterial population and then acquisition of different mechanisms of resistance and change interventions accordingly. Unfortunately, there was insufficient time to pursue this line of investigation further.

Chapter 4: Understanding the relationships between BSI and carriage *E. coli* isolates in patients attending HTD, Vietnam.

4.1. Introduction

Our results from Chapter 2 demonstrated that *E. coli* isolates from rectal swab were more likely to be members of ECOR phylogroup A or B1 than they were to belong to B2 or D. This is consistent with previous reports that have shown that phylogroups A and B1 are often identified as “commensal” or “harmless” bacteria and are less often the cause of BSIs. Although there are no single or defined sets of genes that absolutely distinguish invasive from commensal *E. coli* [10], invasive *E. coli* (ExPEC) often belonged to phylogroup B2 or D and harbour genes related to adhesion, immune evasion, iron acquisition and toxin production [109] as shown for the HTD isolates in Chapter 3.

In our 06EI study (see Section 2.2, Chapter 2), *E. coli* were cultured from 700/834 (84%) HTD ICU patients (data unpublished). Patients when staying long-term in the ICUs are highly vulnerable, and easily contract HAIs, due to their close contact with healthcare workers and other patients [150, 151]. HAI-associated outbreaks in ICU affect both adults and neonates and often involve strains that are highly resistant to therapeutic antibiotic usage in the hospital environments – these are frequently multidrug resistant (MDR) which is associated with a prolonged duration of hospital admission and requires more intense treatment and monitoring which significantly increases treatment costs as well as being associated with, and increasing the risk of a negative clinical outcome [79, 118, 152].

Chapters 2 and 3 of this study showed that there were clear phylogenetic and genetic differences between the randomly sampled rectal swab and BSI isolates from HTD. However, despite being representative of the isolates causing disease they were all derived from unlinked patients. Samples collected in study 06EI provided a unique opportunity to understand the relationship between rectal swab (carriage) and BSI (invasive) isolates taken from the same patient. The samples included in this study also included those taken from a limited number of patients who had been sampled longitudinally during the course of their hospital stay and so it was also possible to compare these *E. coli* isolates to those taken from the same patient at admission either from rectal swab and or those cultured from blood.

Aim: To understand the genetic diversity and relationships of carriage and invasive *E. coli* isolated from the same patient in HTD, Vietnam.

The specific questions to answer:

1) What is the relationship between the paired *E. coli* causing BSIs in a patient to those cultured from their rectal swabs?

2) How does *E. coli* diversity change over time during hospitalisation within a single patient? Are patients carrying the same strains for the duration of their stay or are they reinfected with new strains during their hospitalisation?

4.2. Methods

4.2.1 Study participants

Patient recruitment and sample collection was as detailed in Section 2.2. The 06EI study at the ICU in HTD was conducted from November 2014 to January 2016. Eligible patients (age ≥ 15 years old) admitted to ICU gave informed consent, and a rectal swab was taken from each patient within the first two days post-hospitalisation. This was dubbed "baseline" *E. coli* carriage. Follow-up rectal swabs were taken twice a week until the patient was discharged and were classified as "longitudinal" isolates.

4.2.2. Isolate collection

Isolates were collected and stored as described in Section 2.2. If growth on solid media revealed the presence of two or more colony morphologies (lactose/non-lactose fermented/ β-haemolytic) then multiple colonies were selected for sequencing. DNA was prepared from these isolates and sent to Sanger Institute for sequencing on Illumina HiSeq platform as described in Chapter 2.

4.2.3. Phylogenetic analysis

Phylogenetic analysis was essentially as described in Chapter 2 Section 2.2 based on whole genome SNPs built from core gene alignment. Core gene alignment was constructed by ROARY, except that we only included a subset of 94 isolates and did not include the reference genomes of *E. fergusonii* and *E. albertii*. With the smaller sample size the core genes present in at least 95% of these isolates amounted to 3,077 genes, as expected higher than the total core genes (2,796) defined for the entire sample set in Chapter 2. The pan genome of 94 isolates also comprised of fewer genes (18,886).

4.2.4. AMR genes, replicon types and virulence genes identification

These genomes of these isolates were also screened for antimicrobial resistance, plasmid replicons and virulence genes using ARIBA as describe in Chapter 3 (see Section 3.2). Phylogenetic tree and gene presence/absence as described in Section 3.2.

4.3. Results

4.3.1. Patient and sample recruitment

After QC, we successfully generated genomes of matched *E. coli* isolates taken from rectal swab (carriage) and blood collected in 2015 from a subset of twenty-eight patients (of the original thirty-one patients with matched samples. See Section 2.2) attending HTD. Of these patients twenty-five had BSI on admission or in less than 2 days after admission and so were considered to represent patients with community acquired blood stream infections. Two patients were transferred from a different ward after

8 days of treatment from the same hospital (03-0029 and 03-0514) while the third patient (03-0283) developed BSIs while being treated for tetanus; therefore, these three patients were recorded as HAIs in the 06EI study.

Of these 28 patients, 8 were male and 20 were female. The median age was 53 years of age (range 27-68 year old). Majority patients had an underlying disease that may have predisposed them to bacterial BSIs: 11 patients had liver disease as a consequence of chronic hepatitis B/C or alcoholism. In the remaining 17 patients, other comorbidities included pneumonia, AIDS, tetanus, cholecystitis, diabetes and UTIs, while three others had no known comorbidities. All 8 male patients were recorded as being moderate/heavy smokers and drinking alcoholic beverages, while none of female patients reported that they either smoked or drank alcohol.

In summary, there were 94 isolates remained after QC for this analysis. This included 28 bloodstreams and 46 rectal swab carriage isolates that were collected on day 0 or 1 of admission to HTD ICU, plus 20 longitudinal isolates taken after day 1 from 7 patients. The maximum sampling period was 20 days post admission for patient P3 (03-283).

4.3.2. Genomic diversity among rectal swab and blood derived *E. coli* isolates on admission to HTD ICU

The core genome phylogenies of all 94 isolates detailed above are shown in Figure 4.1. The phylogenetic tree shows that there are 13 BAPs phylogenetic groups (except L1 and L8) represented in the tree. BAPs groups were previously defined in Figure 2.4 Section 2.3 Chapter 2 and transposed onto this analysis. By MLST these isolates include a total of 37 ST's, of which the dominant STs were ST1193, ST131, and ST648 (Supplementary Table S4). If this is divided by sample type then the most common STs in the blood samples were ST1193 (5 samples), ST131 (4 samples) and ST69 (3 samples). For rectal isolates, the dominant STs were ST131 (9 samples), ST1193 (6 samples) and ST648 (8 samples) (Figure 4.1).

Of the 28 patients included for 16/28 (57 %) the same *E. coli* STs were cultured from the blood and rectal swab samples. For the remaining 12 patients, the *E. coli* cultured from blood and carriage samples were of different STs. Its worth noting that rectal swab isolates for 2/28 patients, P1 (03-0048) and P6 (03-0209) (Figure 4.1) with matched samples were collected after day 0. This was because the samples taken on admission failed to grow when the samples were collected for this study. Since the patients would have received treatment in this time after admission, the likelihood is that this influenced nature of the *E. coli* population in the gut. Hence, it is possible that the percentage of identical isolates on admission in the blood and rectal swab was higher 57%.

For the 28 paired samples, the genomes of the blood and carriage isolates from 16 patients were on average only differentiated by a median of 2 SNPs (IQR 1-3) (Supplementary Table S5). Clearly this located these isolates on the same phylogenetic branch of the tree. Paired rectal swab and blood isolates

from the same patient can be seen as red and orange nodes (spots) from blood and rectal swab, respectively on Figure 4.1. The arcs shown in Figure 4.1 link the respective paired sample(s) from blood and rectal swab. If multiple isolates were collected on each time point (due to multiples colony morphologies on the resultant culture plates; see methods) from the rectal swab samples then the blood samples are linked by multiple arcs to their respective paired rectal swab nodes (Figure 4.1).

To give an estimation of diversity we plotted the pairwise SNP distances between isolates from the same patient with the same MLST profile or the same ST taken from different patients (Figure 4.2). We used ST to control for the variable inter-lineage (ST) diversity. For the genomes of isolates of the same ST taken from different samples from the same patient the SNP difference between isolates was within 3 SNPs (median 1, IQR 0 – 2.5). However, if the same ST was taken from different patients, regardless of sample type then they are on average 3,670 SNPs apart at the whole genome level, (median 2,517; IQR 119 – 7,135) (Figure 4.2).

There were two exceptions, the paired *E. coli* genomes from patients P18 (03-0301) and P19 (03-0322) (Supplementary Table S5), were separated by only 40 SNPs respectively although both belonged to same ST1193 and ST131, as well as the same BAPS lineage (labelled as P18 and P19 of Figure 4.1). However, unlike the other linked samples in this study, because we know the estimated mutation rate of *E. coli* ST131 is 1 SNPs/genome/year [153] and because the same STs from 2 different patients in our study also have at least 24 SNPs difference, if 2 isolates are >40 SNPs apart it is likely that the patient has acquired the strains from a different source, including other body sites, unrelated to their own rectal carriage strains or there was unsampled diversity not captured on the rectal swab.

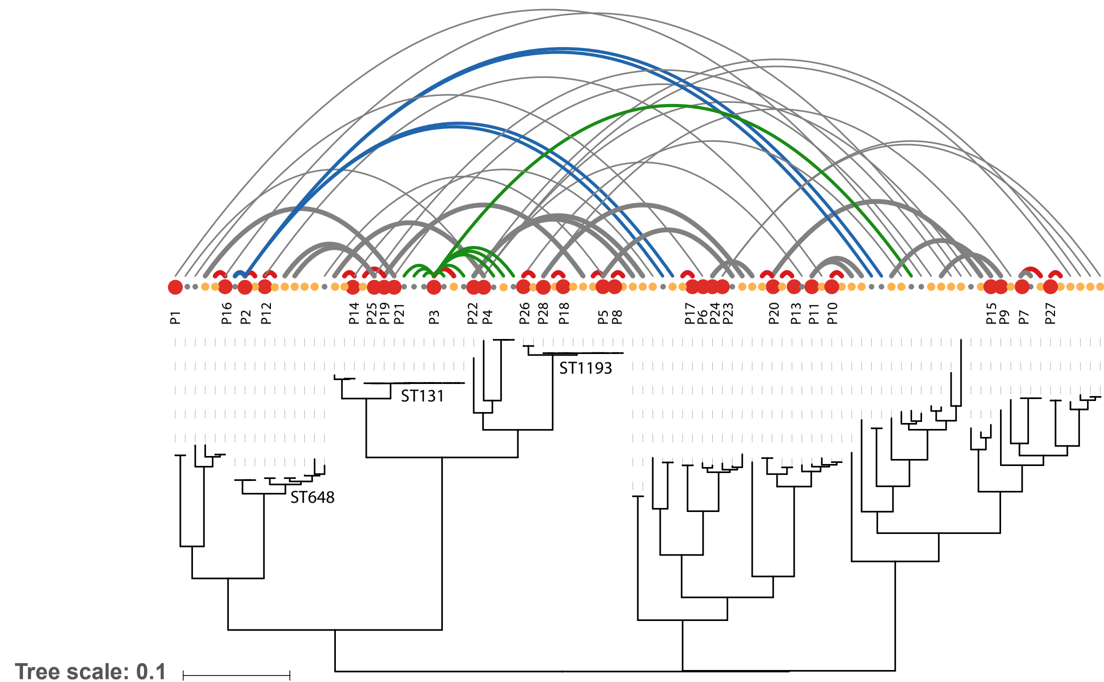


Figure 4.1: Phylogenetic diversity among rectal swab and blood derived *E. coli* isolates from HTD.

A phylogenetic tree constructed from the core genome of 94 isolates. Each node (coloured spot) represents one isolate and the arcs connecting the nodes link rectal swab isolates to their respective blood isolate within the same patient. Red nodes are isolates from blood (n=28), yellow nodes are baseline rectal swab isolates, while grey nodes are longitudinal rectal swab isolates. The red arc indicates the blood and rectal swab isolates were the same ST. The blue arc connects all the isolates taken from patient P2, the green arc connects isolates taken from patient P3 as detailed in Figure 4.3.

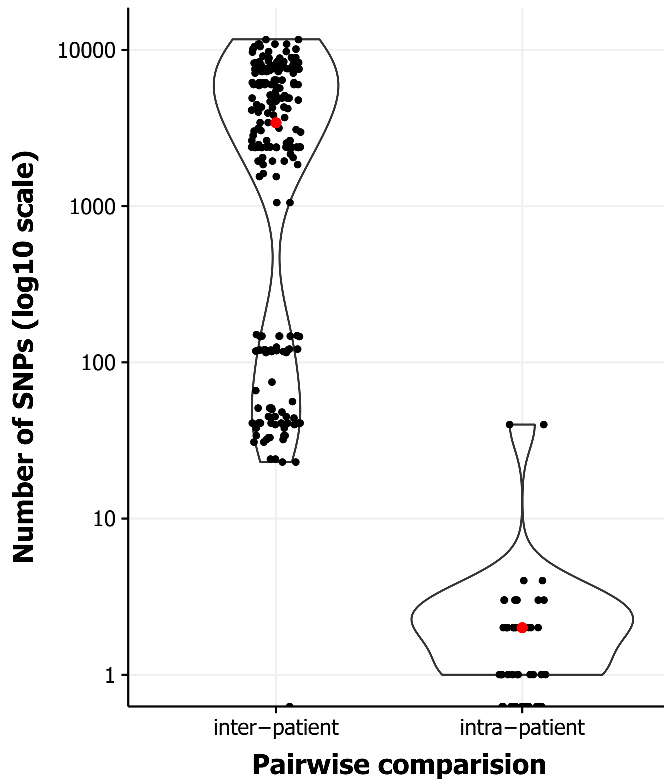


Figure 4.2: Pairwise SNP distances within and between patients based on the variant calls from the core gene alignment. A violin plot for SNP distances between isolates taken from the same patient (intra-patient) and in different patients (inter-patient) is reported. Only those paired samples separated by fewer than 12,000 SNPs are reported, since these capture genomic diversity within an MLST, while isolates between different MLST types had a median of 65,000 SNPs. Red dot represents the median SNP differences from each group. Fifteen pairs that are separated by 0 SNPs (*i.e.*, are identical) are reported in this figure as being separated by fewer than one SNP.

In an attempt to answer this question, we showed that of the remaining 12 patients where the blood and rectal swab isolates belonged to different STs, 6 patients were infected with UPEC clones such as ST73, ST69 and ST131. Looking through the hospital records these patients also had confirmed UTIs (while 2 had ascites, 1 with cholecystitis and 1 unknown), suggesting that the other route of infection possibly was through UTIs. Combined, this shows that for the majority of paired isolates the dominant isolate in the gut (and perhaps the bladder) belonged to the same ST and, from the core genome alignment, was almost indistinguishable from the isolates causing BSIs in that patient on admission to hospital.

4.3.2. Longitudinal diversity

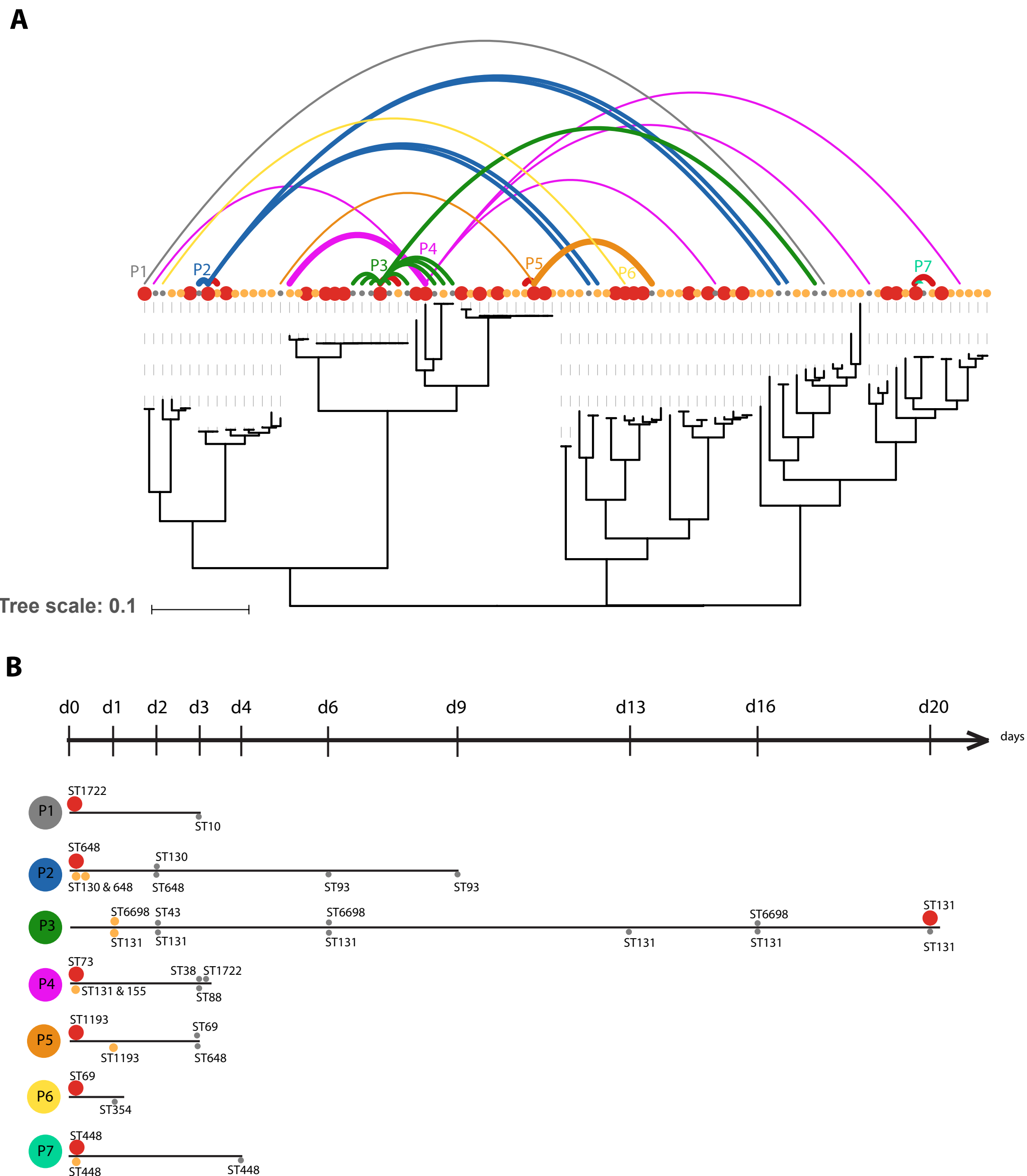


Figure 4.3: Twenty longitudinal isolates and from seven patients. The tree presented in this figure is the same as Figure 4.1 except that, for ease of viewing, only arcs from seven of the patients with multiple rectal swab samples are shown. Each arc with a different colour represents blood and rectal swab isolates taken on admission (red and orange nodes, respectively), or isolates from longitudinal samples (grey nodes) belonging to the same patient. The lower diagram shows the timeline of these longitudinal isolates and their STs.

How does the population structure of *E. coli* change over time during ICU hospitalisation and antibiotic treatment?

As mentioned above, within our sample collection there were samples that belonged to 7 patients who had been sampled multiple times after admission to HTD. The sections below attempt to bring all of the genomic and patient clinical meta data described in this and previous chapters together for these seven patients in order to describe how the population of *E. coli* in these patients has changed over time and highlight what factors may have influenced that. The phenotypic antibiotic resistance data discussed below is summarised in Supplementary Table S1 and S2 and was determined at the HTD Vietnam. The genes encoding the resistance were investigated in Chapter 3 (Supplementary Table S6). Figure 4.3 describes the STs of the *E. coli* taken from these patients plotted over the duration of their stay in the hospital with the sampling intervals labelled on it.

Patient P1 (03-0048) and P6 (03-0209)

Unfortunately, patient P1 (03-0048) and P6 (03-0209) did not have baseline isolates (strains could not be recovered from storage). However, looking at the longitudinal isolates collected from these patients after admission showed they were different to their blood isolates (ST10 and ST354) (Figure 4.3). These two patients both had liver disease (cirrhosis and ascites), and both suffered from community-acquired BSIs, infected with ESBL-producing *E. coli* belonging to ST1722 (P1) and ST69 (P6) respectively. P1 was treated with ceftriaxone and had a fatal outcome after three days. P6 was treated with imipenem for five days after admission, therefore this patient survived the BSIs although they were admitted for a total of 29 days in HTD. ST354 colonised P6 on day 1 and is an ESBL-producing *E. coli*, that carried both $bla_{\text{CTX-M-24}}$ and $bla_{\text{CMY-42}}$, and was phenotypically resistant to all tested antibiotics except carbapenem.

Patient P2 (03-0314)

Patient P2 (03-0314) was a patient suffering from cirrhosis. Initially, this patient's rectal swab sample was dominated by two STs, ST130 and ST648, both sensitive to 12 out of 13 tested antimicrobials (Supplementary Table S2). The blood isolate matched the rectal swab isolate with only one SNP difference. *E. coli* isolates were also cultured with the same AMR profile from peritoneal fluid (data not shown). It is very likely that this BSI emerged from translocation of *E. coli* from the gut. Patient P2 was treated with ceftriaxone. However, on day 6 and day 9, only MDR ESBL-producing *E. coli* of ST93 were seen in culture. These isolates harboured IncX and IncFII plasmid replicon types (result from Chapter 3). This might reflect either lineage replacement whereby only when sensitive STs such as ST130 and ST648 were killed off by exposure to ceftriaxone, we were then able to pick up drug resistant ST93. The other possibility was this patient acquired a hospital-associated MDR *E. coli*, although we have no means to identify transmission route unless we had sequenced all carriage isolates from other patients admitted to ICU around that time. The patient unfortunately did not survive.

Patient P4 and P5

Longitudinal rectal swab isolates taken from patients P4 (03-0559) and P5 (03-0710) were also found to be different STs from the original rectal swab isolate taken on admission. P4 was treated with ceftriaxone, since the admission BSI isolate, ST73, was found to be resistant to cotrimoxazole (SXT) and also had reduced susceptibility to ticarcillin/clavulanic acid (TCC) and amoxicillin/clavulanic acid (AMC). However, the rectal swab sample was dominated by ESBL-producing ST131 $bla_{CTX\text{-}M\text{-}14}$ and SXT resistant *E. coli* ST155 cultured from the rectal swab collected on admission. On day three post-admission, further rectal swab samples were taken and *E. coli* isolates of ST38, ST1722 and ST88 were grown. Interestingly, they were all phenotypically resistant to third- and fourth-generation cephalosporins and were ESBL-producing isolates (Supplementary Table S2). Screening for AMR genes using ARIBA and ResFinder revealed (See sections 3.3) that they all harboured $bla_{CTX\text{-}M}$ genes. ST38 isolates contained $bla_{CTX\text{-}M\text{-}27}$, ST1722 had both $bla_{CTX\text{-}M\text{-}27}$ and $bla_{DHA\text{-}1}$ while ST88 harboured $bla_{CTX\text{-}M\text{-}55}$.

Patient P5 had diabetes with confirmed UTIs, while carriage isolates were both the same ESBL-producing *E. coli* ST1193 harboured $bla_{CTX\text{-}M\text{-}27}$ so it was very likely that BSIs arose from the gut colonization then contaminated UTIs. After treatment with both ceftriaxone and amikacin, on day 3, two different STs including sensitive ST69 and MDR ST648 isolates with *aac(6')-Ib-cr*, $bla_{CTX\text{-}M\text{-}15}$, $bla_{OXA\text{-}1}$, $bla_{TEM\text{-}1B}$ correlating with the presence of a IncFII plasmid (data visualised in Bandage - not shown) were cultured from rectal swabs. This patient's condition did not improve after treatment, and they were transferred to another hospital, hence the opportunity to follow this up was lost.

Patient P7 (03-0514)

Patient P7 (03-0514) was one of three patients that were classified as HAIs in 06EI study since all three patients developed BSIs while staying in the hospital when treated for cirrhosis or tetanus. Since this patient was infected with the same ESBL-producing ST448 in the blood and rectal swab samples, they were treated with norfloxacin and ertapenem; nevertheless, ST448 was also shown to be present in rectal samples on day 4, with zero SNPs separating it from the original two isolates. This patient was discharged to die at home. ResFinder results revealed that these isolates carried $bla_{CMY\text{-}2}$.

Patient P3 (03-283)

Patient P3 (03-283) was admitted to the hospital suffering from tetanus. This patient did not have a bacteraemia on admission. This patient was swabbed six times during a 19-day stay in the ICU and *E. coli* isolates were cultured and retained from the rectal swabs on each occasion. The patient was screened for *E. coli* upon admission, and was treated with metrodinazole (as per standard guidelines for tetanus treatment). Interestingly, the cultured isolates showed multiple colony morphologies on each occasion and so two colonies were preserved from each of the rectal cultures, including those from day one. In addition, this patient developed a BSI on day 20 with the infection being recorded as a hospital-

acquired infection. It is clear (Figure 4.3) that patient P3 (03-283) rectal samples were dominated by two different *E. coli* MLST clones (ST131, ST6698), while ST43 seems likely to have colonised this patient transiently. It is also clear that the day 20 BSI isolate belonged to the same ST and phylogroup as the majority of the rectal isolates from this patient. This ST131 clone was picked up every sampling time because it is a β-haemolytic *E. coli*, hence it shows a phenotype that is distinct from other lactose fermenting clone (see Supplementary Table 3).

To see if there was evidence of variation between the longitudinally sampled ST131 isolates we re-drew the ST131 cluster in Phandango to include all seven isolates from this patient (Figure 4.4 carriage isolates denoted as LF (lactose-fermented) or NLF (non-lactose fermented) or Bhae(β-hemolytic) proceeding their sample ID, e.g. EI2793NLF, EI2793LF2/EI3892_Bhae; the day 20 BSI isolate is labelled as 010615-16405). An additional three BSIs isolates, and one carriage isolate from three different patients were added for context (Figure 4.4; Labelled as P19, P21, P25). Within seven longitudinal isolates taken from patient P3 03-0283, the blood isolate is most genetically closely related to isolate EI1605 β-haemolytic isolate from day 6, which is identical with 0 SNPs difference in the core genome alignment. The maximum SNP difference between these seven longitudinal isolates was 3 SNPs. In order to understand if there were any differences in the accessory genome between the blood and rectal swab isolates included here, we mapped sequencing reads against 205 genes known to be involved in invasive disease to look for gene presence or absence (Figure 4.4). On first analysis, this seemed to reveal something interesting. The blood isolates differed from the other five carriage isolates from the same patient by possessing two genes, *kspM*-1 and *papX* genes. PapX is important because it is a protein which represses the synthesis of P fimbriae, which we showed was highly correlated with invasive disease at HTD (described in Chapter 3). The impact of this gene would be to attenuate attachment and bacterial motility. Whilst *kspM*-1 mediated K1 capsules help *E. coli* escape host immunity response [154]. However, looking across all isolates included in Figure 4.4 it is clear that the presence or absence of these genes, although interesting, does not correlate with sample type: blood or rectal swab although it does highlight some within-host variation between isolates.

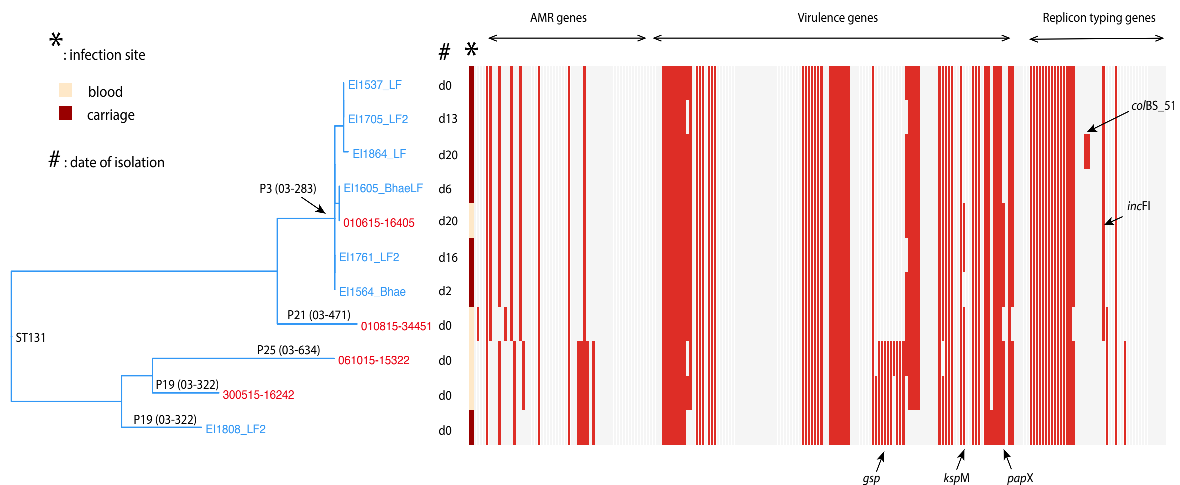


Figure 4.4: Graph shows phylogenetic relationship between isolates in patient P3 (03-283). The presence and absence of AMR and virulence genes was related to the core genome phylogeny of seven isolates from patient P3 and two other patients infected with ST131 BSIs.

4.4. Discussion

In our study, by using whole genome sequencing as a high-resolution method to dissect the complexity of *E. coli* populations, we found here that there is a substantial diversity of *E. coli* co-existing in each patient's gut; each patient carried at least two different STs from different phylogenetic lineages. Hence, our results are consistent with the study by Stoesser *et al* on multiple single *E. coli* isolates isolated from Cambodian children, in which it was found that one healthy individual can carry up to 10 different STs and a range of virulence and AMR determinants located on multiple plasmids [155].

It is also clear that 57 % of BSIs collected from our patients for whom we had matched rectal swab isolates were of the same ST and differed by 0 or only a small number of SNPs from their rectal swab isolates, which suggests that they come from their own microbiota. This has been observed in other ICU studies for *Klebsiella pneumoniae* [156, 157]. However, although we were not able to infer the route of infection for the other cases a significant number of these other patients carried isolates strongly associated with UTIs in their blood, were recorded as having UTI's and were largely female. This may suggest that this is the other major source of BSIs in these patients and is consistent with the findings of Chapter 2 and 3 showing a strong gender bias for the invasive isolates and a link to UTI's.

The diversity of multiple STs living in the same gut flora lead to an interesting point as to whether drug-sensitive *E. coli* could acquire AMR genes through plasmid transfer from AMR strains under antibiotic pressure. Although gut inflammation has been shown to boost the conjugation of a colicin-resistance plasmid between *Salmonella* Typhimurium and commensal *E. coli* in a mouse model [158], our finding in a small subset of patients rather show the lineage replacement of sensitive STs by those resistant to

the drug used for treatment (as shown in patients P2, P4 and P5). The repeated isolation of the same *E. coli* carried bla_{CMY-27} in patient P7 (03-514) (Figure 4.3) show the potential of treatment failure of ertapenem on bla_{CMY-27}-producing *E. coli*.

Our study has limitation. Due to the number of strains that needed to be identified and stored during the study, we assumed that all *E. coli* isolates that were of the same phenotype (colony morphology, colour on MacConkey agar), and of the same AMR pattern, were of the same genotype. This meant that only one or two representative isolates from each phenotype was kept and sequenced. This could lead to an underestimation of the real diversity of *E. coli* present in the gut and might explain why 43% of our cases harbour different STs between blood and rectal swab on admission.

Chapter 5 Final discussion

The majority of recent large-scale sequencing studies of *E. coli* have focused on the most prevalent and highly MDR clone, ST131, and have aimed to find factors responsible for its success as a pathogen [135, 153, 159]. Rather than focusing on a single ST or MDR clone, our study took a longitudinal approach, to capture the diversity of *E. coli* causing bacteraemia in a single hospital over a five-year period to provide basic understanding of E. coli population structure, both from invasive and carriage isolates.

By sequencing randomly more than 650 invasive and carriage isolates over 5 years, our study revealed that there was high genetic diversity in *E. coli*, with member of all classic ECOR phylogroups, comprised of 117 different STs or 15 BAPs clusters. Although exhibiting high diversity, at the same time, all the major STs causing BSIs in Vietnamese patients typically mirror the diversity of ExPEC all over world, including ST131, ST73, ST95, ST69 and ST1193 as dominant STs [101, 105]. These STs have been known for a long time as ExPEC associated with BSIs and UTIs [102, 160]. Interestingly, among these STs, ST131 and ST1193 is common in both invasive and carriage while ST73 and ST95 only found in BSI samples.

Stratifying by year, we have shown that the population structure of dominant *E. coli* STs is stable overtime, with in general the proportion of each STs remaining relatively stable every year. This in itself is an extremely important observation to frame the work done to reduce infection at HTD. The only exception was the introduction of ST1193 in 2011, but once established the population regained stability again. This phenomenon was also observed in the 10-year collection of BSI *E. coli* in United Kingdom [161]. Since majority of BSI were present on admission and so community acquired, these support the hypothesis that there was a high prevalence of carriage of these ExPEC in healthy individuals outside the hospital and in the community. This is particularly significant for immunocompromised patients and patients with other severe comorbidities or unhealthy lifestyles, as we have shown.

As mention above, our study highlighted the introduction of ST1193 in causing BSIs as well as being present in carriage samples in Vietnamese patients. ST1193 was a fluoroquinolone (FQ) resistant clone found as early as 2007 in both human and companion animals (dogs) in Australia [162]. Isolates between those two groups shared the same virulence genes as well as AMR genes, suggesting it was shared between human and animal. ST1193 has also been documented as non-lactose fermenting FQ resistant *E. coli* causing UTIs in Korea [163], being the second most prevalence ST. In the United States, ST1193 was rare, accounted for only 0.9% FQ resistant isolates. However, ST1193 was recently shown to be amongst the top STs associated with community onset infection in 30 county hospitals across China, from 2010 to 2011 [106]. Apart from exhibiting FQ resistant phenotype, ST1193 emerged as a clone that harboured different $bla_{\text{CTX-M}}$ including $bla_{\text{CTX-M-15}}$, $bla_{\text{CMY-2}}$ [164] and $bla_{\text{CTX-M-55}}$ [165].

ST1193 in our collection interestingly, although FQ resistant, harboured a different *bla*$_{CTX-M}$ (*bla*$_{CTX-M-27}$). Collectively, having the same virulence repertoire as ST131, being able to colonize human gut, plus FQ resistant and ESBL-producing might explain the successful clonal expansion of ST1193 in Vietnam. However, further studies combing genomics and epidemiological will be needed to confirm if this is also replicated across and outside of Asia.

Our data confirm that the global emergence of ST131 as the most prevalent ST causing BSIs is also replicated in our hospital and seems to be much more diverse than previously thought. Previous reports [135] have indicated that ST131 isolated from across the globe is FQ resistant and also harbours *bla*$_{CTX-M-15}$. However, in our collection, ST131 did not only harbour *bla*$_{CTX-M-15}$ alone, but also carried plenty other CTX-M variants such as *bla*$_{CTX-M-14}$, *bla*$_{CTX-M-27}$ and *bla*$_{CTX-M-55}$. This highlights the versatility of ST131 in acquiring and maintaining MDR plasmids, in comparison to the other 2 most common ExPECs: ST73 and ST95. The differences observed between the AMR profiles of ST131/ST1193 and ST73/ST95 lead to another interesting point. Since we were unable to find ST73 and ST95 in carriage isolates, we speculate that these two STs might colonise another habitat outside of human gut. The gut and urinary tract are very distinct environments in terms of the availability of nutrients, immunological control, and the population of resident microbes [166]. Since only a small proportion of bacteria that have siderophore systems can live in the urethra [167], perhaps these UPEC clones ST73 and ST69, do not have a chance to encounter other wider community of bacteria including MDR one as in the gastrointestinal tract. Therefore, there are not many chances to pick up mobile genetic elements like plasmids and transposon that carried resistant gene around as ST131.

Although our work in Chapter 4 included a relatively small sample size of 28 patients, this is the first report using WGS to confirm that the majority of BSIs patients (57%) are infected with *E. coli* that are indistinguishable or almost indistinguishable from the dominant strain(s) colonising their guts. This implies that patients were infected with carriage isolates already colonizing their body sites and not through hospital acquired infections as we had initially hypothesised. However, there were exceptions by which haft of the cases were different ST between blood and carriage. For a significant proportion of these isolates we showed these patients had UTI infections, sometimes with *E. coli* that had matching AMR patterns. Although this is perhaps strong circumstantial evidence that their BSIs might originate from prior UTIs we could not confirm this route of infection because the isolates were not stored. If that was true, it will inform future study design by showing that positive urine samples from other bodily sites are critical in order to fully understand colonisation and subsequently invasive infection of *E. coli*. If there were to be a future study, we would collect and sequence all paired *E. coli* isolated from blood, urine and rectal swab on admission from each patient and at a larger scale. By comparing the distribution of STs between different body sites, in the future we might be able to predict which STs can cause BSIs and which STs is just limited to cause bladder infection and also study the interaction of multiple STs coexist in the gut.

Now knowing that colonizing of ExPEC strains such as ST131 and ST1193 pose high risk to subsequently infection in immune compromised patients, either through bacterial translocation in the gut or by seeding bacteria from the gut to the urethra and subsequently bladder infection in the elderly, there are several approaches can be implemented to prevent BSIs infection. The first one is by decolonizing or limiting the presence of these ExPEC and MDR STs in high risk patients. Randomized clinical trial with *Lactobacillus plantarum* has showed that it can prevent 42% of sepsis cases in newborn between control and cases [168]. In term of preventing UTIs, asymptomatic bacteriuria *E. coli* strain 83972 has been used to colonized the urethra and compete with virulent UPEC [169, 170]. Other approach includes treating the gut with sugar molecules resemble sugar structure located on gut epithelial cells that *FimH* can bind, hence also reducing the colonization of *FimH* producing *E. coli* [171]. Since possessing P fimbriae and yersiniabactin has been shown, by us and others, to be the important virulence factors for *E. coli* invasive disease, developing vaccines against these 2 candidates could be a potential approach as well [172, 173].

Our study has several limitations. First of all, because this is a retrospective study we had little control on the sampling strategy either for the random samples or the matched samples from the same patient. We also only had access to rectal swab and blood samples and yet it is clear that other sample such as urines were collected for some patients with UTIs, but unfortunately they were not stored. Also, some of the patient meta data was predicted, for example liver disease was inferred due to patient care and was not through laboratory confirmed testing such as the AST liver function test. The same was true of the CD4 count for HIV status.

Another limitation of the study was the in ability to assemble plasmids from WGS data. In our dataset, we observed substantial diversity of AMR genes and plasmid replicon types, but due to the nature of short reads sequence produced by Illumina sequencing as well as the presence of repetitive elements such as transposases, we were unable to reconstruct the full plasmid structure and the detailed association between certain type of Inc plasmid and certain AMR genes as well as the genetic backbone surround the genes. One plasmid can harbour multiple replicon types apart from their original replicon, and multiple plasmids may co-exist in one cell. This makes it technically challenging to infer total numbers of plasmids based on total replicon types alone. We also used plasmidSPADES [126] to assemble plasmid (data not shown), however, plasmidSPADES operates by comparing read coverage difference between chromosome and plasmid DNA. Since big plasmids (> 90 kb) are not always present in high copy number, they are present as equivalent read depth to the chromosome and therefore are not recognized by plasmidSPADES. Hence these contigs often remain fragmented plasmicidic contigs. In the next study, we would like to employ Long read sequencing also used to elucidate the diversity of AMR genes each clone carried and build as a framework of to track plasmid movement between Enterobacteriace species, especially in our ICU collection.

In conclusion, the success of certain STs is a multifactorial phenomenon including possessing certain virulence genes (attachment, invasion, iron acquisition, toxin producing) and harbouring AMR genes as selective advantage against antimicrobial treatment. Strain that are capable of doing both things, such as ST131 and ST1193, will continue to account for a high proportion of BSIs cases.

References

1. Shulman, S.T., H.C. Friedmann, and R.H. Sims, *Theodor Escherich: the first pediatric infectious diseases physician?* Clin Infect Dis, 2007. **45**(8): p. 1025-9.
2. Blount, Z.D., *The unexhausted potential of* E. coli. eLife, 2015. **4**.
3. Tenaillon, O., et al., *The population genetics of commensal* Escherichia coli. Nat Rev Microbiol, 2010. **8**(3): p. 207-17.
4. Sender, R., S. Fuchs, and R. Milo, *Revised estimates for the number of human and bacteria cells in the body.* PLoS Biol, 2016. **14**(8): p. e1002533.
5. Stecher, B. and W.D. Hardt, *The role of microbiota in infectious disease.* Trends Microbiol, 2008. **16**(3): p. 107-14.
6. Berg, R.D., *The indigenous gastrointestinal microflora.* Trends Microbiol, 1996. **4**(11): p. 430-5.
7. Hudault, S., J. Guignot, and A.L. Servin, Escherichia coli *strains colonising the gastrointestinal tract protect germfree mice against* Salmonella *typhimurium infection.* Gut, 2001. **49**(1): p. 47-55.
8. Bergholz, P.W., J.D. Noar, and D.H. Buckley, *Environmental patterns are imposed on the population structure of* Escherichia coli *after fecal deposition.* Appl Environ Microbiol, 2011. **77**(1): p. 211-9.
9. Goto, D.K. and T. Yan, *Genotypic diversity of* Escherichia coli *in the water and soil of tropical watersheds in Hawaii.* Appl Environ Microbiol, 2011. **77**(12): p. 3988-97.
10. Leimbach, A., J. Hacker, and U. Dobrindt, E. coli *as an all-rounder: the thin line between commensalism and pathogenicity.* Curr Top Microbiol Immunol, 2013. **358**: p. 3-32.
11. Nowrouzian, F., et al., Escherichia coli *in infants' intestinal microflora: colonization rate, strain turnover, and virulence gene carriage.* Pediatr Res, 2003. **54**(1): p. 8-14.
12. Caugant, D.A., B.R. Levin, and R.K. Selander, *Genetic diversity and temporal variation in the* E. coli *population of a human host.* Genetics, 1981. **98**(3): p. 467-90.
13. Selander, R.K. and B.R. Levin, *Genetic diversity and structure in* Escherichia coli *populations.* Science, 1980. **210**(4469): p. 545-7.
14. Escobar-Paramo, P., et al., *Large-scale population structure of human commensal* Escherichia coli *isolates.* Appl Environ Microbiol, 2004. **70**(9): p. 5698-700.
15. Skurnik, D., et al., *Characteristics of human intestinal* Escherichia coli *with changing environments.* Environ Microbiol, 2008. **10**(8): p. 2132-7.
16. Massot, M., et al., *Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of* Escherichia coli *from community subjects in the Paris area in 2010 and evolution over 30 years.* Microbiology, 2016. **162**(4): p. 642-50.
17. Smati, M., et al., *Quantitative analysis of commensal* Escherichia coli *populations reveals host-specific enterotypes at the intra-species level.* MicrobiologyOpen, 2015. **4**(4): p. 604-15.

18. Touchon, M., et al., *Organised genome dynamics in the* Escherichia coli *species results in highly diverse adaptive paths.* PLoS Genet, 2009. **5**(1): p. e1000344.
19. Lukjancenko, O., T.M. Wassenaar, and D.W. Ussery, *Comparison of 61 sequenced* Escherichia coli *genomes.* Microbiol Ecol, 2010. **60**(4): p. 708-20.
20. Kaas, R.S., et al., *Estimating variation within the genes and inferring the phylogeny of 186 sequenced diverse* Escherichia coli *genomes.* BMC Genomics, 2012. **13**: p. 577.
21. McInerney, J.O., A. McNally, and M.J. O'Connell, *Why prokaryotes have pangenomes.* Nat Microbiol, 2017. **2**: p. 17040.
22. Land, M., et al., *Insights from 20 years of bacterial genome sequencing.* Funct Integr Genomics, 2015. **15**(2): p. 141-61.
23. Cook, H. and D.W. Ussery, *Sigma factors in a thousand* E. coli *genomes.* Environ Microbiol, 2013. **15**(12): p. 3121-9.
24. Russo, T., *Medical and economic impact of extraintestinal infections due to* Escherichia coli*: focus on an increasingly important endemic problem.* Microbes and Infection, 2003. **5**(5): p. 449-456.
25. Scharff, R.L., *Economic burden from health losses due to foodborne illness in the United States.* J Food Prot, 2012. **75**(1): p. 123-31.
26. Kotloff, K.L., et al., *Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study.* Lancet, 2013. **382**(9888): p. 209-22.
27. Lanata, C.F., et al., *Global causes of diarrheal disease mortality in children <5 years of age: a systematic review.* PLoS One, 2013. **8**(9): p. e72788.
28. Flores-Mireles, A.L., et al., *Urinary tract infections: epidemiology, mechanisms of infection and treatment options.* Nat Rev Microbiol, 2015. **13**(5): p. 269-84.
29. Kaper, J.B., J.P. Nataro, and H.L. Mobley, *Pathogenic* Escherichia coli. Nat Rev Microbiol, 2004. **2**(2): p. 123-40.
30. Croxen, M.A. and B.B. Finlay, *Molecular mechanisms of* Escherichia coli *pathogenicity.* Nat Rev Microbiol, 2010. **8**(1): p. 26-38.
31. Martin, G.S., et al., *The epidemiology of sepsis in the United States from 1979 through 2000.* N Engl J Med, 2003. **348**(16): p. 1546-54.
32. Cohen, J., *The immunopathogenesis of sepsis.* Nature, 2002. **420**(6917): p. 885-91.
33. Kalra, O.P. and A. Raizada, *Approach to a patient with urosepsis.* J Glob Infect Dis, 2009. **1**(1): p. 57-63.
34. Marchetti, G., C. Tincati, and G. Silvestri, *Microbial translocation in the pathogenesis of HIV infection and AIDS.* Clin Microbiol Rev, 2013. **26**(1): p. 2-18.
35. Wiest, R., M. Lawson, and M. Geuking, *Pathological bacterial translocation in liver cirrhosis.* J Hepatol, 2014. **60**(1): p. 197-209.

36. Samet, A., et al., *Leukemia and risk of recurrent* Escherichia coli *bacteremia: genotyping implicates* E. coli *translocation from the colon to the bloodstream.* Eur J Clin Microbiol Infect Dis, 2013. **32**(11): p. 1393-400.
37. Green, S.I., et al., *Murine model of chemotherapy-induced extraintestinal pathogenic* Escherichia coli *translocation.* Infect Immun, 2015. **83**(8): p. 3243-56.
38. Laupland, K.B., et al., *Incidence, risk factors and outcomes of* Escherichia coli *bloodstream infections in a large Canadian region.* Clin Microbiol Infect, 2008. **14**(11): p. 1041-7.
39. Bou-Antoun, S., et al., *Descriptive epidemiology of* Escherichia coli *bacteraemia in England, April 2012 to March 2014.* Euro Surveill, 2016. **21**(35).
40. Jackson, L.A., et al., *Burden of community-onset* Escherichia coli *bacteremia in seniors.* J Infect Dis, 2005. **191**(9): p. 1523-9.
41. van der Mee-Marquet, N.L., et al., *Marked increase in incidence for bloodstream infections due to* Escherichia coli*, a side effect of previous antibiotic therapy in the elderly.* Front Microbiol, 2015. **6**: p. 646.
42. Schlackow, I., et al., *Increasing incidence of* Escherichia coli *bacteraemia is driven by an increase in antibiotic-resistant isolates: electronic database study in Oxfordshire 1999-2011.* J Antimicrob Chemother, 2012. **67**(6): p. 1514-24.
43. Reddy, E.A., A.V. Shaw, and J.A. Crump, *Community-acquired bloodstream infections in Africa: a systematic review and meta-analysis.* Lancet Infect Dis, 2010. **10**(6): p. 417-32.
44. Deen, J., et al., *Community-acquired bacterial bloodstream infections in developing countries in south and southeast Asia: a systematic review.* Lancet Infect Dis, 2012. **12**(6): p. 480-487.
45. Southeast Asia Infectious Disease Clinical Research Network, *Causes and outcomes of sepsis in southeast Asia: a multinational multicentre cross-sectional study.* Lancet Glob Health, 2017. **5**(2): p. e157-e167.
46. Huson, M.A., et al., *Community-acquired bacterial bloodstream infections in HIV-infected patients: a systematic review.* Clin Infect Dis, 2014. **58**(1): p. 79-92.
47. Vlieghe, E.R., et al., *Bloodstream infection among adults in Phnom Penh, Cambodia: key pathogens and resistance patterns.* PLoS One, 2013. **8**(3): p. e59775.
48. Nga, T.V., et al., *The decline of typhoid and the rise of non-typhoid salmonellae and fungal infections in a changing HIV landscape: bloodstream infection trends over 15 years in southern Vietnam.* Trans R Soc Trop Med Hyg, 2012. **106**(1): p. 26-34.
49. Dat, V.Q., et al., *Bacterial bloodstream infections in a tertiary infectious diseases hospital in Northern Vietnam: aetiology, drug resistance, and treatment outcome.* BMC Infect Dis, 2017. **17**(1): p. 493.
50. Dale, A.P. and N. Woodford, *Extra-intestinal pathogenic* Escherichia coli *(ExPEC): Disease, carriage and clones.* J Infect, 2015. **71**(6): p. 615-26.

51. Martin, G.S., *Sepsis, severe sepsis and septic shock: changes in incidence, pathogens and outcomes.* Expert Rev Anti Infect Ther, 2012. **10**(6): p. 701-6.
52. Lukac, P.J., R.A. Bonomo, and L.K. Logan, *Extended-spectrum beta-lactamase-producing Enterobacteriaceae in children: old foe, emerging threat.* Clin Infect Dis, 2015. **60**(9): p. 1389-97.
53. Nhu, N.T., et al., *Emergence of carbapenem-resistant* Acinetobacter baumannii *as the major cause of ventilator-associated pneumonia in intensive care unit patients at an infectious disease hospital in southern Vietnam.* J Med Microbiol, 2014. **63**: p. 1386-94.
54. Davies, J. and D. Davies, *Origins and evolution of antibiotic resistance.* Microbiol Mol Biol Rev, 2010. **74**(3): p. 417-33.
55. Aminov, R.I., *A brief history of the antibiotic era: lessons learned and challenges for the future.* Front Microbiol, 2010. **1**: p. 134.
56. Blair, J.M., et al., *Molecular mechanisms of antibiotic resistance.* Nat Rev Microbiol, 2015. **13**(1): p. 42-51.
57. Gao, R., et al., *Dissemination and mechanism for the MCR-1 colistin resistance.* PLoS Pathog, 2016. **12**(11): p. e1005957.
58. Liu, Y.Y., et al., *Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study.* Lancet Infect Dis, 2016. **16**(2): p. 161-8.
59. Tumbarello, M., et al., *Bloodstream infections caused by extended-spectrum-beta-lactamase-producing* Escherichia coli*: risk factors for inadequate initial antimicrobial therapy.* Antimicrob Agents Chemother, 2008. **52**(9): p. 3244-52.
60. Rodriguez-Bano, J., et al., *Risk factors and prognosis of nosocomial bloodstream infections caused by extended-spectrum-beta-lactamase-producing* Escherichia coli*.* J Clin Microbiol, 2010. **48**(5): p. 1726-31.
61. Thaden, J.T., et al., *Increased costs associated with bloodstream infections caused by multidrug-resistant Gram-negative bacteria are due primarily to patients with hospital-acquired infections.* Antimicrob Agents Chemother, 2017. **61**(3).
62. Lim, C., et al., *Epidemiology and burden of multidrug-resistant bacterial infection in a developing country.* eLife, 2016. **5**.
63. Hsu, L.Y., et al., *Carbapenem-resistant* Acinetobacter baumannii *and Enterobacteriaceae in South and Southeast Asia.* Clin Microbiol Rev, 2017. **30**(1): p. 1-22.
64. WHO, *Global priority list of antibiotic-resistant bacteria to guide research, discovery, and development of new antibiotics.* http://www.who.int/medicines/publications/global-priority-list-antibiotic-resistant-bacteria/en, 2017.
65. Leibovici, L., et al., *The benefit of appropriate empirical antibiotic treatment in patients with bloodstream infection.* J Intern Med, 1998. **244**(5): p. 379-86.

66. Timsit, J.F., et al., *Treatment of bloodstream infections in ICUs.* BMC Infect Dis, 2014. **14**: p. 489.
67. Bevan, E.R., A.M. Jones, and P.M. Hawkey, *Global epidemiology of CTX-M beta-lactamases: temporal and geographical shifts in genotype.* J Antimicrob Chemother, 2017. **72**(8): p. 2145-2155.
68. Zhao, W.H. and Z.Q. Hu, *Epidemiology and genetics of CTX-M extended-spectrum beta-lactamases in Gram-negative bacteria.* Crit Rev Microbiol, 2013. **39**(1): p. 79-101.
69. Nordmann, P., et al., *The emerging NDM carbapenemases.* Trends Microbiol, 2011. **19**(12): p. 588-95.
70. Carattoli, A., et al., *Identification of plasmids by PCR-based replicon typing.* J Microbiol Methods, 2005. **63**(3): p. 219-28.
71. Thomas, C.M. and K.M. Nielsen, *Mechanisms of, and barriers to, horizontal gene transfer between bacteria.* Nat Rev Microbiol, 2005. **3**(9): p. 711-21.
72. Brauner, A., et al., *Characterization of* Escherichia coli *isolated in blood, urine and faeces from bacteraemic patients and possible spread of infection.* APMIS, 1991. **99**(4): p. 381-6.
73. Bergeron, C.R., et al., *Chicken as reservoir for extraintestinal pathogenic* Escherichia coli *in humans, Canada.* Emerg Infect Dis, 2012. **18**(3): p. 415-21.
74. Vincent, C., et al., *Food reservoir for* Escherichia coli *causing urinary tract infections.* Emerg Infect Dis, 2010. **16**(1): p. 88-95.
75. de Been, M., et al., *Dissemination of cephalosporin resistance genes between* Escherichia coli *strains from farm animals and humans by specific plasmid lineages.* PLoS Genet, 2014. **10**(12): p. e1004776.
76. Borjesson, S., et al., *Limited dissemination of extended-spectrum beta-lactamase- and plasmid-encoded AmpC-producing* Escherichia coli *from food and farm animals, Sweden.* Emerg Infect Dis, 2016. **22**(4): p. 634-40.
77. Lazarus, B., et al., *Do human extraintestinal* Escherichia coli *infections resistant to expanded-spectrum cephalosporins originate from food-producing animals? A systematic review.* Clin Infect Dis, 2015. **60**(3): p. 439-52.
78. Luna, G.M., et al., *Extraintestinal* Escherichia coli *carrying virulence genes in coastal marine sediments.* Appl Environ Microbiol, 2010. **76**(17): p. 5659-68.
79. Chung The, H., et al., *A high-resolution genomic analysis of multidrug-resistant hospital outbreaks of* Klebsiella pneumoniae. EMBO Mol Med, 2015. **7**(3): p. 227-39.
80. Page, A.J., et al., *Robust high-throughput prokaryote de novo assembly and improvement pipeline for Illumina data.* Microb Genom, 2016. **2**(8): p. e000083.
81. Gladman, S. and T. Seemann, *Velvet Optimiser* https://github.com/tseemann/VelvetOptimiser. 2008.

82. Zerbino, D.R. and E. Birney, *Velvet: algorithms for de novo short read assembly using de Bruijn graphs.* Genome Res, 2008. **18**(5): p. 821-9.
83. Boetzer, M., et al., *Scaffolding pre-assembled contigs using SSPACE.* Bioinformatics, 2011. **27**(4): p. 578-9.
84. Boetzer, M. and W. Pirovano, *Toward almost closed genomes with GapFiller.* Genome Biol, 2012. **13**(6): p. R56.
85. Seemann, T., *Prokka: rapid prokaryotic genome annotation.* Bioinformatics, 2014. **30**(14): p. 2068-9.
86. Pruitt, K.D., et al., *NCBI Reference Sequences (RefSeq): current status, new features and genome annotation policy.* Nucleic Acids Res, 2012. **40**(Database issue): p. D130-5.
87. Andrews, S., *FastQC* https://www.bioinformatics.babraham.ac.uk/projects/fastqc/. 2010.
88. Ewels, P., et al., *MultiQC: summarize analysis results for multiple tools and samples in a single report.* Bioinformatics, 2016. **32**(19): p. 3047-8.
89. Wood, D.E. and S.L. Salzberg, *Kraken: ultrafast metagenomic sequence classification using exact alignments.* Genome Biol, 2014. **15**(3): p. R46.
90. Parks, D.H., et al., *CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes.* Genome Res, 2015. **25**(7): p. 1043-55.
91. Page, A.J., et al., *Comparison of multi-locus sequence typing software for next generation sequencing data.* bioRxiv, 2017.
92. Ochman, H. and R.K. Selander, *Standard reference strains of* Escherichia coli *from natural populations.* J Bacteriol, 1984. **157**(2): p. 690-3.
93. van Elsas, J.D., et al., *Survival of* Escherichia coli *in the environment: fundamental and public health aspects.* ISME J, 2011. **5**(2): p. 173-83.
94. Clermont, O., S. Bonacorsi, and E. Bingen, *Rapid and simple determination of the* Escherichia coli *phylogenetic group.* Appl Environ Microbiol, 2000. **66**(10): p. 4555-8.
95. Page, A.J., et al., *Roary: rapid large-scale prokaryote pan genome analysis.* Bioinformatics, 2015. **31**(22): p. 3691-3.
96. Price, M.N., P.S. Dehal, and A.P. Arkin, *FastTree: computing large minimum evolution trees with profiles instead of a distance matrix.* Mol Biol Evol, 2009. **26**(7): p. 1641-50.
97. Cheng, L., et al., *Hierarchical and spatially explicit clustering of DNA sequences with BAPS software.* Mol Biol Evol, 2013. **30**(5): p. 1224-8.
98. Letunic, I. and P. Bork, *Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees.* Nucleic Acids Res, 2016. **44**(W1): p. W242-5.
99. Hadfield, J., et al., *Phandango: an interactive viewer for bacterial population genomics.* bioRxiv, 2017.
100. Rasko, D.A., et al., *The pangenome structure of* Escherichia coli*: comparative genomic analysis of* E. coli *commensal and pathogenic isolates.* J Bacteriol, 2008. **190**(20): p. 6881-93.

101. Salipante, S.J., et al., *Large-scale genomic sequencing of extraintestinal pathogenic* Escherichia coli *strains.* Genome Res, 2015. **25**(1): p. 119-28.
102. Lau, S.H., et al., *Major uropathogenic* Escherichia coli *strain isolated in the northwest of England identified by multilocus sequence typing.* J Clin Microbiol, 2008. **46**(3): p. 1076-80.
103. Holt, K.E., et al., *Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in* Klebsiella pneumoniae*, an urgent threat to public health.* Proc Natl Acad Sci U S A, 2015. **112**(27): p. E3574-81.
104. Harris, S.R., et al., *Whole-genome sequencing for analysis of an outbreak of meticillin-resistant* Staphylococcus aureus*: a descriptive study.* Lancet Infect Dis, 2013. **13**(2): p. 130-136.
105. Adams-Sapper, S., et al., *Clonal composition and community clustering of drug-susceptible and -resistant* Escherichia coli *isolates from bloodstream infections.* Antimicrob Agents Chemother, 2013. **57**(1): p. 490-7.
106. Zhao, L., et al., *Molecular epidemiology and genetic diversity of fluoroquinolone-resistant* Escherichia coli *isolates from patients with community-onset infections in 30 Chinese county hospitals.* J Clin Microbiol, 2015. **53**(3): p. 766-70.
107. Boyd, E.F. and D.L. Hartl, *Chromosomal regions specific to pathogenic isolates of* Escherichia coli *have a phylogenetically clustered distribution.* J Bacteriol, 1998. **180**(5): p. 1159-65.
108. Kohler, C.D. and U. Dobrindt, *What defines extraintestinal pathogenic* Escherichia coli*?* Int J Med Microbiol, 2011. **301**(8): p. 642-7.
109. Starcic Erjavec, M. and D. Zgur-Bertok, *Virulence potential for extraintestinal infections among commensal* Escherichia coli *isolated from healthy humans - the Trojan horse within our gut.* FEMS Microbiol Lett, 2015. **362**(5).
110. Foxman, B., *Epidemiology of urinary tract infections: incidence, morbidity, and economic costs.* Am J Med, 2002. **113 Suppl 1A**: p. 5S-13S.
111. Wiles, T.J., R.R. Kulesus, and M.A. Mulvey, *Origins and virulence mechanisms of uropathogenic* Escherichia coli*.* Exp Mol Pathol, 2008. **85**(1): p. 11-9.
112. Luthje, P. and A. Brauner, *Virulence factors of uropathogenic E. coli and their interaction with the host.* Adv Microb Physiol, 2014. **65**: p. 337-72.
113. Snyder, J.A., et al., *Role of phase variation of type 1 fimbriae in a uropathogenic* Escherichia coli *cystitis isolate during urinary tract infection.* Infect Immun, 2006. **74**(2): p. 1387-93.
114. Holden, V.I. and M.A. Bachman, *Diverging roles of bacterial siderophores during infection.* Metallomics, 2015. **7**(6): p. 986-95.
115. Cross, A.S., et al., *The importance of the K1 capsule in invasive infections caused by* Escherichia coli*.* J Infect Dis, 1984. **149**(2): p. 184-93.
116. Allen, P.M., et al., *Contribution of capsular polysaccharide and surface properties to virulence of* Escherichia coli *K1.* Infect Immun, 1987. **55**(11): p. 2662-8.

117. Poirel, L., et al., *Genetic features of* bla*NDM-1-positive Enterobacteriaceae.* Antimicrob Agents Chemother, 2011. **55**(11): p. 5403-7.
118. Livermore, D.M., *Current epidemiology and growing resistance of Gram-negative pathogens.* Korean J Intern Med, 2012. **27**(2): p. 128-42.
119. Carattoli, A., et al., *In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing.* Antimicrob Agents Chemother, 2014. **58**(7): p. 3895-903.
120. Lan, R., et al., *Molecular evolutionary relationships of enteroinvasive* Escherichia coli *and* Shigella *spp.* Infect Immun, 2004. **72**(9): p. 5080-8.
121. Croxen, M.A., et al., *Recent advances in understanding enteric pathogenic* Escherichia coli. Clin Microbiol Rev, 2013. **26**(4): p. 822-80.
122. Carattoli, A., *Resistance plasmid families in Enterobacteriaceae.* Antimicrob Agents Chemother, 2009. **53**(6): p. 2227-38.
123. Hunt, M., et al., *ARIBA: rapid antimicrobial resistance genotyping directly from sequencing reads.* bioRxiv, 2017.
124. Kleinheinz, K.A., K.G. Joensen, and M.V. Larsen, *Applying the ResFinder and VirulenceFinder web-services for easy identification of acquired antibiotic resistance and* E. coli *virulence genes in bacteriophage and prophage nucleotide sequences.* Bacteriophage, 2014. **4**(1): p. e27943.
125. Chen, L., et al., *VFDB 2016: hierarchical and refined dataset for big data analysis--10 years on.* Nucleic Acids Res, 2016. **44**(D1): p. D694-7.
126. Antipov, D., et al., *plasmidSPAdes: assembling plasmids from whole genome sequencing data.* Bioinformatics, 2016. **32**(22): p. 3380-3387.
127. Bankevich, A., et al., *SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing.* J Comput Biol, 2012. **19**(5): p. 455-77.
128. Wick, R.R., et al., *Bandage: interactive visualization of de novo genome assemblies.* Bioinformatics, 2015. **31**(20): p. 3350-2.
129. CLSI, *Standards for antimicrobial susceptibility testing; twenty-fourth informational supplement, CLSI document M100-S24.* Clinical and Laboratory Standards Institute, Wayne, PA, 2014.
130. Drieux, L., et al., *Phenotypic detection of extended-spectrum beta-lactamase production in Enterobacteriaceae: review and bench guide.* Clin Microbiol Infect, 2008. **14 Suppl 1**: p. 90-103.
131. Wickham, H., *ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.* 2009.
132. Simms, A.N. and H.L. Mobley, *PapX, a P fimbrial operon-encoded inhibitor of motility in uropathogenic* Escherichia coli*.* Infect Immun, 2008. **76**(11): p. 4833-41.
133. Raymond, K.N., E.A. Dertz, and S.S. Kim, *Enterobactin: an archetype for microbial iron transport.* Proc Natl Acad Sci U S A, 2003. **100**(7): p. 3584-8.

134. Bugarel, M., et al., *Virulence gene profiling of enterohemorrhagic (EHEC) and enteropathogenic (EPEC)* Escherichia coli *strains: a basis for molecular risk assessment of typical and atypical EPEC strains.* BMC Microbiol, 2011. **11**: p. 142.
135. Petty, N.K., et al., *Global dissemination of a multidrug resistant* Escherichia coli *clone.* Proc Natl Acad Sci U S A, 2014. **111**(15): p. 5694-9.
136. Liu, Y.-Y., et al., *Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study.* Lancet Infect Dis, 2016. **16**(2): p. 161-168.
137. Tada, T., et al., *Emergence of colistin-resistant* Escherichia coli *clinical isolates harboring* mcr-*1 in Vietnam.* Int J Infect Dis, 2017.
138. Dona, V., et al., *Heterogeneous genetic location of* mcr-*1 in colistin-resistant* Escherichia coli *isolated from humans and retail chicken meat in Switzerland: Emergence of* mcr-*1-carrying IncK2 plasmids.* Antimicrob Agents Chemother, 2017.
139. Rutherford, K., et al., *Artemis: sequence visualization and annotation.* Bioinformatics, 2000. **16**(10): p. 944-5.
140. Josenhans, C. and S. Suerbaum, *The role of motility as a virulence factor in bacteria.* Int J Med Microbiol, 2002. **291**(8): p. 605-14.
141. Kao, C.Y., et al., *The complex interplay among bacterial motility and virulence factors in different* Escherichia coli *infections.* Eur J Clin Microbiol Infect Dis, 2014. **33**(12): p. 2157-62.
142. Fischbach, M.A., et al., *How pathogenic bacteria evade mammalian sabotage in the battle for iron.* Nat Chem Biol, 2006. **2**(3): p. 132-8.
143. Sakamoto, K., et al., *IL-22 controls iron-dependent nutritional immunity against systemic bacterial infections.* Sci Immunol, 2017. **2**(8).
144. Antao, E.M., L.H. Wieler, and C. Ewers, *Adhesive threads of extraintestinal pathogenic* Escherichia coli*.* Gut Pathog, 2009. **1**(1): p. 22.
145. Lasaro, M.A., et al., *F1C fimbriae play an important role in biofilm formation and intestinal colonization by the* Escherichia coli *commensal strain Nissle 1917.* Appl Environ Microbiol, 2009. **75**(1): p. 246-51.
146. Cirl, C., et al., *Subversion of Toll-like receptor signaling by a unique family of bacterial Toll/interleukin-1 receptor domain-containing proteins.* Nat Med, 2008. **14**(4): p. 399-406.
147. Garcia, T.A., et al., *Cytotoxic necrotizing factor 1 and hemolysin from uropathogenic* Escherichia coli *elicit different host responses in the murine bladder.* Infect Immun, 2013. **81**(1): p. 99-109.
148. Sugumar, M., et al., *Detection of OXA-1 beta-lactamase gene of* Klebsiella pneumoniae *from blood stream infections (BSI) by conventional PCR and in-silico analysis to understand the mechanism of OXA mediated resistance.* PLoS One, 2014. **9**(3): p. e91800.

149. Brynildsrud, O., et al., *Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary.* Genome Biol, 2016. **17**(1): p. 238.

150. Snitkin, E.S., et al., *Tracking a hospital outbreak of carbapenem-resistant* Klebsiella pneumoniae *with whole-genome sequencing.* Sci Transl Med, 2012. **4**(148): p. 148ra116.

151. Koser, C.U., et al., *Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak.* N Engl J Med, 2012. **366**(24): p. 2267-75.

152. Stoesser, N., et al., *Genome sequencing of an extended series of NDM-producing* Klebsiella pneumoniae *isolates from neonatal infections in a Nepali hospital characterizes the extent of community- versus hospital-associated transmission in an endemic setting.* Antimicrob Agents Chemother, 2014. **58**(12): p. 7347-57.

153. Stoesser, N., et al., *Evolutionary history of the global emergence of the* Escherichia coli *epidemic clone ST131.* mBio, 2016. **7**(2): p. e02162.

154. Reiss, D.J. and H.L. Mobley, *Determination of target sequence bound by PapX, repressor of bacterial motility, in* flhD *promoter using systematic evolution of ligands by exponential enrichment (SELEX) and high throughput sequencing.* J Biol Chem, 2011. **286**(52): p. 44726-38.

155. Stoesser, N., et al., *Extensive within-host diversity in fecally carried extended-spectrum-beta-lactamase-producing* Escherichia coli *isolates: Implications for transmission analyses.* J Clin Microbiol, 2015. **53**(7): p. 2122-31.

156. Gorrie, C.L., et al., *Gastrointestinal carriage is a major reservoir of* K. pneumoniae *infection in intensive care patients.* Clin Infect Dis, 2017.

157. Martin, R.M., et al., *Molecular epidemiology of colonizing and infecting isolates of* Klebsiella pneumoniae. mSphere, 2016. **1**(5).

158. Stecher, B., et al., *Gut inflammation can boost horizontal gene transfer between pathogenic and commensal Enterobacteriaceae.* Proc Natl Acad Sci U S A, 2012. **109**(4): p. 1269-74.

159. Ben Zakour, N.L., et al., *Sequential acquisition of virulence and fluoroquinolone resistance has shaped the evolution of* Escherichia coli *ST131.* mBio, 2016. **7**(2): p. e00347-16.

160. Alhashash, F., et al., *Multidrug-resistant* Escherichia coli *bacteremia.* Emerg Infect Dis, 2013. **19**(10): p. 1699-701.

161. Kallonen, T., et al., *Systematic longitudinal survey of invasive* Escherichia coli *in England demonstrates a stable population structure only transiently disturbed by the emergence of ST131.* Genome Res, 2017.

162. Platell, J.L., et al., *Prominence of an O75 clonal group (clonal complex 14) among non-ST131 fluoroquinolone-resistant* Escherichia coli *causing extraintestinal infections in humans and dogs in Australia.* Antimicrob Agents Chemother, 2012. **56**(7): p. 3898-904.

163. Chang, J., et al., *Prevalence and characteristics of lactose non-fermenting* Escherichia coli *in urinary isolates.* J Infect Chemother, 2014. **20**(11): p. 738-40.

164. Xia, S., et al., *Dominance of CTX-M-type extended-spectrum beta-lactamase (ESBL)-producing* Escherichia coli *isolated from patients with community-onset and hospital-onset infection in China.* PLoS One, 2014. **9**(7): p. e100707.
165. Xia, L., et al., *Prevalence of ST1193 clone and IncI1/ST16 plasmid in* E. coli *isolates carrying* bla*CTX-M-55 gene from urinary tract infections patients in China.* Sci Rep, 2017. **7**: p. 44866.
166. Chen, S.L., et al., *Genomic diversity and fitness of* E. coli *strains recovered from the intestinal and urinary tracts of women with recurrent urinary tract infection.* Sci Transl Med, 2013. **5**(184): p. 184ra60.
167. Watts, R.E., et al., *Contribution of siderophore systems to growth and urinary tract colonization of asymptomatic bacteriuria* Escherichia coli. Infect Immun, 2012. **80**(1): p. 333-44.
168. Panigrahi, P., et al., *A randomized synbiotic trial to prevent sepsis among infants in rural India.* Nature, 2017. **548**(7668): p. 407-412.
169. Ferrieres, L., V. Hancock, and P. Klemm, *Biofilm exclusion of uropathogenic bacteria by selected asymptomatic bacteriuria* Escherichia coli *strains.* Microbiology, 2007. **153**(Pt 6): p. 1711-9.
170. Abraham, S.N. and Y. Miao, *The nature of immune responses to urinary tract infections.* Nat Rev Immunol, 2015. **15**(10): p. 655-63.
171. Spaulding, C.N., et al., *Selective depletion of uropathogenic* E. coli *from the gut by a FimH antagonist.* Nature, 2017. **546**(7659): p. 528-532.
172. Roberts, J.A., et al., *Antibody responses and protection from pyelonephritis following vaccination with purified* Escherichia coli *PapDG protein.* J Urol, 2004. **171**(4): p. 1682-5.
173. Brumbaugh, A.R. and H.L. Mobley, *Preventing urinary tract infection: progress toward an effective* Escherichia coli *vaccine.* Expert Rev Vaccines, 2012. **11**(6): p. 663-76.

Supplementary table S1. Details and epidemiological data for 506 invasive *E. coli* isolates from HTD.

Sample_accessio	lane	SID	infection_sit	acites_pos	urine_pos	liverdisea	hiv	age	gender	year	ST_getST	HierBAPS	outcome_	ESBL	AMK1	AMC1	CAZ1	CIP1	CRO1	ETP1	FEP1	IPM1	MEM1	OFX1	SXT1	TCC1	TZP1
ERS1369064	21374_6#1	010210_63044	blood	0	0	1	0	58	male	2010	73	7	worsen	not tested	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369065	21374_6#2	020312_31262	blood	0	1	0	0	58	female	2012	69	4	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369066	21374_6#3	030113_24807	blood	0	0	0	0	65	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369067	21374_6#4	040111_28577	blood	0	1	0	0	26	female	2011	59	15	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369068	21374_6#5	050112_59161	blood	0	1	0	0	18	female	2012	1193	13	nochange	ESBL	S	R	R	R	R	S	S	S	S	R	S	R	S
ERS1369069	21374_6#6	050913_24509	blood	0	1	0	0	60	female	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369070	21374_6#7	060710_38217	blood	0	0	0	1	36	male	2010	648	9	death	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369071	21374_6#8	070311_64462	blood	0	0	0	1	47	male	2011	457	15	nochange	ESBL	S	S	R	N	R	S	R	S	S	N	R	N	R
ERS1369072	21374_6#9	010310_16446	blood	0	0	1	0	83	female	2010	405	1	recover	not tested	S	S	R	I	R	S	R	S	S	S	R	S	S
ERS1369073	21374_6#10	020313_28975	blood	0	1	0	0	58	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369074	21374_6#11	030212_60095	blood	0	1	0	0	77	female	2012	131	12	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369075	21374_6#12	040111_32241	blood	0	0	0	0	45	male	2011	131	12	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	I	S
ERS1369076	21374_6#13	050210_38304	blood	0	0	0	1	33	male	2010	131	12	recover	not tested	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369077	21374_6#14	051110_22963	blood	0	1	0	0	68	female	2010	131	12	recover	not tested	S	S	R	R	R	S	R	S	S	R	R	S	S
ERS1369078	21374_6#15	060710_38667	blood	0	0	0	1	33	male	2010	131	12	death	not tested	S	I	S	R	I	I	S	S	S	R	R	I	S
ERS1369079	21374_6#16	070413_16469	blood	1	0	0	0	60	male	2013	648	9	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369080	21374_6#17	010312_63946	blood	0	0	1	0	50	male	2012	46	14	nochange	no	S	R	S	R	S	S	S	S	S	R	S	R	R
ERS1369081	21374_6#18	020511_15392	blood	0	1	0	0	25	male	2011	156	11	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	R
ERS1369082	21374_6#19	030611_17852	blood	0	1	0	0	48	female	2011	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369083	21374_6#20	040210_36908	blood	0	0	0	0	59	male	2010	448	11	recover	not tested	S	I	S	R	S	S	S	S	S	R	R	S	S
ERS1369084	21374_6#21	050313_17674	blood	0	0	1	0	34	male	2013	69	4	worsen	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369086	21374_6#22	060810_38989	blood	0	0	0	1	20	female	2010	38	3	death	not tested	S	I	R	S	R	S	R	S	S	S	R	I	R
ERS1369087	21374_6#23	070910_15708	blood	0	0	0	0	86	female	2010	617	14	recover	not tested	S	R	R	R	R	S	R	S	S	R	R	I	R
ERS1369088	21374_6#24	010410_39662	blood	0	0	0	1	44	male	2010	448	11	worsen	not tested	S	R	R	R	R	S	R	S	S	R	R	R	I
ERS1369089	21374_6#25	020512_14688	blood	0	0	1	0	42	male	2012	648	9	nochange	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369090	21374_6#26	030710_15824	blood	0	1	0	0	30	male	2010	95	8	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369091	21374_6#27	040411_63509	blood	0	0	0	0	71	female	2011	131	12	recover	no	S	S	S	N	S	S	S	S	S	N	S	N	S
ERS1369092	21374_6#28	050313_63842	blood	0	0	1	0	46	male	2013	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369093	21374_6#29	051212_23101	blood	0	1	0	0	89	female	2012	131	12	recover	ESBL	S	N	R	R	R	S	R	S	S	R	R	R	S
ERS1369094	21374_6#30	060811_64562	blood	0	0	1	0	42	male	2011	69	4	nochange	no	S	R	S	S	I	S	S	S	S	S	R	R	S
ERS1369095	21374_6#31	071012_16492	blood	0	0	0	0	36	female	2012	131	12	recover	ESBL	S	I	R	S	R	S	R	S	S	S	R	I	I
ERS1369096	21374_6#32	010511_15334	blood	0	1	0	0	87	female	2011	624	9	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	R
ERS1369097	21374_6#33	020512_26756	blood	0	0	0	0	54	male	2012	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369099	21374_6#35	040613_32862	blood	0	0	0	0	18	male	2013	38	3	recover	no	S	S	S	R	S	S	S	S	S	R	S	S	S
ERS1369100	21374_6#36	050410_25629	blood	0	1	0	0	58	female	2010	4456	10	worsen	not tested	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369102	21374_6#38	060813_32422	blood	0	0	0	0	28	female	2013	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369103	21374_6#39	071210_15119	blood	0	0	0	0	50	male	2010	46	14	recover	not tested	S	I	S	R	S	S	S	S	S	R	R	R	S
ERS1369104	21374_6#40	010512_63740	blood	0	0	0	0	42	male	2012	1193	13	nochange	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1369105	21374_6#41	020611_25952	blood	0	0	0	0	42	male	2011	345	11	nochange	no	S	S	S	R	S	S	S	S	S	R	S	R	S
ERS1369106	21374_6#42	030910_15342	blood	0	0	0	0	51	female	2010	95	8	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369107	21374_6#43	040712_23077	blood	0	1	0	0	64	female	2012	410	6	recover	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	R
ERS1369108	21374_6#44	050511_25028	blood	0	0	0	0	87	female	2011	95	8	recover	no	S	S	S	S	S	S	S	S	S	N	S	S	S
ERS1369109	21374_6#45	060112_16270	blood	0	1	1	0	75	female	2012	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369110	21374_6#46	060813_39004	blood	0	0	0	1	31	male	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369111	21374_6#47	071211_62111	blood	1	0	0	0	60	female	2011	93	14	worsen	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369112	21374_6#48	010611_63725	blood	0	0	0	0	63	male	2011	69	4	recover	ESBL	S	S	R	N	R	S	I	S	S	S	S	R	R
ERS1369113	21374_6#49	020711_16348	blood	0	0	0	0	80	female	2011	69	4	recover	no	S	I	S	S	S	S	S	S	S	S	S	R	S
ERS1369114	21374_6#50	030911_26912	blood	0	0	0	0	28	female	2011	69	4	recover	no	S	S	S	S	S	I	S	S	S	S	S	R	S
ERS1369115	21374_6#51	041010_14650	blood	0	1	0	0	85	female	2010	131	12	recover	not tested	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1369116	21374_6#52	050512_18639	blood	0	0	0	0	76	male	2012	95	8	worsen	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369117	21374_6#53	060112_16340	blood	0	0	0	0	69	male	2012	69	4	recover	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1369118	21374_6#54	060911_40229	blood	0	0	0	1	49	male	2011	405	1	recover	no	S	S	S	I	S	S	S	S	S	I	R	S	S
ERS1369120	21374_6#55	011013_25877	blood	0	0	0	0	63	female	2013	131	12	nochange	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	I
ERS1369121	21374_6#56	021013_15589	blood	0	0	0	0	92	male	2013	38	3	recover	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1369122	21374_6#57	030912_27734	blood	0	0	0	0	66	female	2012	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369123	21374_6#58	041010_31744	blood	0	0	0	0	70	female	2010	73	7	recover	not tested	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369124	21374_6#59	050613_23258	blood	0	0	0	0	88	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369125	21374_6#60	060113_29765	blood	0	0	0	0	49	female	2013	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	R	S
ERS1369126	21374_6#61	060913_62309	blood	0	0	0	0	59	female	2013	48	14	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369127	21374_6#62	080113_63528	blood	1	0	1	0	73	female	2013	38	3	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	I
ERS1369128	21374_6#63	011013_64251	blood	0	0	1	0	58	female	2013	414	5	recover	ESBL	S	S	R	S	R	S	R	S	S	S	S	R	S
ERS1369129	21374_6#64	021110_63806	blood	0	0	0	0	49	male	2010	354	2	recover	not tested	S	R	S	R	S	S	S	S	S	N	I	S	S
ERS1369131	21374_6#66	041012_23826	blood	0	1	0	0	83	female	2012	73	7	recover	no	S	I	S	S	S	S	S	S	S	S	R	R	I

ERS1369133	21374_6#68	060212_30215	blood	0	1	0	0	20	female	2012	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369134	21374_6#69	061110_28130	blood	0	0	0	0	28	male	2010	131	12	recover	not tested	S	I	R	R	R	S	R	S	S	R	S	I	S
ERS1369135	21374_6#70	080212_60406	blood	0	1	0	0	41	female	2012	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369136	21374_6#71	011211_17500	blood	0	0	0	0	62	male	2011	73	7	recover	no	S	I	S	S	S	S	S	S	S	S	S	R	S
ERS1369137	21374_6#72	021111_37774	blood	0	1	0	0	19	male	2011	73	7	worsen	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369138	21374_6#73	030913_60478	blood	0	1	1	0	45	male	2013	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369139	21374_6#74	041111_15364	blood	0	0	0	0	32	male	2011	12	10	recover	no	S	R	S	S	S	S	S	S	S	S	R	R	I
ERS1369140	21374_6#75	050810_64826	blood	0	0	1	0	54	male	2010	5454~	14	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369142	21374_6#77	070111_14682	blood	0	0	0	0	61	male	2011	69	4	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369143	21374_6#78	080312_64481	blood	0	0	1	0	59	male	2012	1193	13	recover	no	S	S	S	R	S	S	S	S	S	N	R	I	S
ERS1369144	21374_6#79	020112_16057	blood	0	0	0	0	55	female	2012	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	I
ERS1369145	21374_6#80	021112_40646	blood	0	0	0	1	24	male	2012	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369146	21374_6#81	031013_15697	blood	0	0	0	0	80	female	2013	3897~	14	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369147	21374_6#82	041213_28684	blood	0	0	0	0	31	female	2013	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	I	S
ERS1369148	21374_6#83	050811_18095	blood	0	0	0	0	49	female	2011	69	4	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369149	21374_6#84	060611_14093	blood	0	0	1	1	47	male	2011	93	14	recover	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369150	21374_6#85	070211_17208	blood	0	1	0	0	67	male	2011	131	12	recover	ESBL	S	R	R	S	R	S	R	S	S	S	R	N	S
ERS1369151	21374_6#86	080413_16573	blood	1	0	1	0	43	male	2013	349	5	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369152	21374_6#87	020112_38656	blood	0	0	0	1	25	female	2012	46	14	recover	no	S	S	S	R	S	S	S	S	S	R	R	R	S
ERS1369153	21374_6#88	021112_62218	blood	0	0	1	0	71	female	2012	73~	7	nochange	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369154	21374_6#89	031113_14465	blood	0	0	1	0	47	male	2013	405	1	recover	ESBL	S	S	R	R	R	S	R	R	S	R	R	R	S
ERS1369155	21374_6#90	041213_65501	blood	0	0	1	0	83	female	2013	48	14	recover	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369156	21374_6#91	050813_63772	blood	1	0	1	0	37	male	2013	10	14	nochange	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369157	21374_6#92	060611_22114	blood	0	0	0	0	18	female	2011	1193	13	recover	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369158	21374_6#93	070212_16155	blood	0	0	0	1	40	male	2012	131	12	worsen	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369159	21374_6#94	080413_30642	blood	0	0	0	0	75	female	2013	131	12	recover	no	S	I	S	S	S	S	S	S	S	S	R	R	S
ERS1369160	21374_6#95	090613_17399	blood	0	0	0	0	20	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369161	21374_6#96	100313_29475	blood	0	0	0	0	38	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369162	21374_6#97	110412_16784	blood	1	0	1	0	36	male	2012	457	15	recover	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	I
ERS1369163	21374_6#98	120410_15481	blood	0	0	1	0	65	female	2010	1286	14	worsen	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369164	21374_6#99	121212_37983	blood	0	1	0	0	71	male	2012	421	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369165	21374_6#100	130611_28767	blood	0	1	0	0	55	female	2011	38	3	recover	ESBL	S	R	N	R	N	S	N	S	S	R	R	R	S
ERS1369166	21374_6#101	140510_14138	blood	0	1	0	0	41	female	2010	410	6	recover	not tested	S	S	S	R	S	S	S	S	S	R	S	S	S
ERS1369168	21374_6#103	090910_63514	blood	1	0	1	0	58	female	2010	69	4	recover	not tested	S	R	S	S	S	S	S	S	S	S	R	R	R
ERS1369169	21374_6#104	100410_26662	blood	0	0	0	0	57	female	2010	73	7	recover	not tested	S	S	S	N	S	S	S	S	S	S	S	S	S
ERS1369170	21374_6#105	110413_30780	blood	0	0	0	0	17	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369171	21374_6#106	120410_33535	blood	0	0	0	0	56	female	2010	38	3	recover	not tested	S	R	S	S	S	S	S	S	S	S	R	R	S
ERS1369172	21374_6#107	121213_17744	blood	0	0	1	0	65	male	2013	1193	13	recover	no	S	S	S	R	S	S	S	S	S	R	R	R	S
ERS1369173	21374_6#108	130712_40976	blood	0	0	0	1	45	male	2012	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369174	21374_6#109	140512_33791	blood	0	1	0	0	25	female	2012	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369175	21374_6#110	080711_26353	blood	0	0	0	0	43	male	2011	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369176	21374_6#111	090912_65733	blood	0	0	1	0	49	male	2012	196	11	nochange	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1369177	21374_6#112	100512_15249	blood	1	0	1	0	47	male	2012	1177	3	worsen	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369178	21374_6#113	110511_38660	blood	0	0	0	1	25	male	2011	38	3	worsen	no	S	R	S	R	I	S	S	S	S	R	R	R	I
ERS1369179	21374_6#114	120411_17672	blood	0	0	0	0	65	female	2011	410	6	recover	ESBL	S	I	N	N	R	S	R	S	S	N	R	R	I
ERS1369180	21374_6#115	130111_39537	blood	1	0	0	1	29	male	2011	405	1	death	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	R
ERS1369181	21374_6#116	130812_15828	blood	0	0	0	0	50	male	2012	405	1	worsen	no	R	R	R	R	R	R	R	I	R	R	S	R	R
ERS1369182	21374_6#117	140513_34180	blood	0	0	0	0	29	male	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369183	21374_6#118	080810_17020	blood	0	0	0	0	82	female	2010	131	12	nochange	not tested	S	S	R	R	R	S	R	S	S	R	R	R	R
ERS1369184	21374_6#119	091010_17587	blood	0	0	0	0	86	female	2010	38	3	recover	not tested	S	S	R	S	R	S	R	S	S	S	R	S	S
ERS1369185	21374_6#120	100712_23467	blood	0	0	0	0	61	female	2012	969	10	recover	no	S	I	S	S	S	S	S	S	S	S	S	R	R
ERS1369186	21374_6#121	110610_17919	blood	0	1	0	0	35	male	2010	131	12	recover	not tested	S	I	R	S	R	S	R	S	S	S	R	S	S
ERS1369187	21374_6#122	120412_16847	blood	0	0	1	0	42	male	2012	1193	13	nochange	no	S	S	S	R	S	S	S	S	S	R	R	R	S
ERS1369188	21374_6#123	130112_16947	blood	0	0	0	0	40	female	2012	648	9	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369189	21374_6#124	130912_22825	blood	0	0	0	0	83	female	2012	131	12	nochange	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	S
ERS1369190	21374_6#125	140610_15930	blood	0	0	0	0	26	male	2010	1196	11	recover	not tested	S	I	S	R	S	S	S	S	S	R	R	S	S
ERS1369191	21374_6#126	080810_41701	blood	0	0	0	1	29	male	2010	131	12	recover	not tested	S	S	S	R	S	S	S	S	S	R	S	S	S
ERS1369192	21374_6#127	091012_16757	blood	0	0	0	0	17	male	2012	206	14	recover	ESBL	S	R	R	R	R	S	S	S	S	R	R	R	I
ERS1369193	21374_6#128	101010_45248	blood	0	0	0	0	0	male	2010	95	8	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	R
ERS1369194	21374_6#129	110613_63593	blood	0	0	1	0	67	male	2013	58	11	nochange	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369195	21374_6#130	120511_27232	blood	0	0	0	0	77	female	2011	405	1	recover	ESBL	S	I	R	N	R	S	R	S	S	R	R	R	R
ERS1369196	21374_6#131	130112_38372	blood	0	0	0	1	30	male	2012	131	12	worsen	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1369197	21374_6#132	130912_28617	blood	0	0	0	0	15	male	2012	410	6	recover	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369198	21374_6#133	140610_15956	blood	0	0	0	0	49	male	2010	624	9	recover	not tested	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1369199	21374_6#134	080811_39122	blood	0	0	0	1	32	male	2011	73	7	nochange	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369200	21374_6#135	091112_28827	blood	0	0	0	0	34	female	2012	648	9	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S

ERS1369201	21374_6#136	101210_15433	blood	0	0	1	0	69	male	2010	394	5	worsen	not tested	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369202	21374_6#137	110812-15553-a	blood	0	0	0	0	30	male	2012	224	11	recover	no	S	R	S	S	S	S	S	S	S	S	R	R	I
ERS1369203	21374_6#138	120610_17238	blood	0	0	1	0	59	male	2010	405	1	recover	not tested	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369204	21374_6#139	130311_15842	blood	0	0	0	0	77	male	2011	131	12	recover	no	S	S	S	N	S	S	S	S	S	N	S	N	S
ERS1369205	21374_6#140	131111_25296	blood	0	1	0	0	86	female	2011	405	1	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	S
ERS1369206	21374_6#141	140611_40707	blood	1	0	0	1	30	male	2011	93	14	nochange	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1369207	21374_6#142	081210_15188	blood	0	0	0	0	70	female	2010	410	6	recover	not tested	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1369208	21374_6#143	091210_65287	blood	0	0	1	0	45	female	2010	95	8	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369209	21374_6#144	101211_38912	blood	0	0	0	1	46	male	2011	131	12	worsen	no	S	I	S	S	S	S	S	S	S	S	R	R	I
ERS1369210	21374_6#145	110812-15553-b	blood	0	0	0	0	30	male	2012	130	5	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	I
ERS1369211	21374_6#146	120611_28682	blood	0	0	0	0	55	female	2011	421	8	recover	no	S	I	R	S	R	S	R	S	S	S	R	R	S
ERS1369213	21374_6#147	131210_34369	blood	0	0	0	0	53	male	2010	73	7	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369214	21374_6#148	140810_29363	blood	0	0	0	0	71	female	2010	95	8	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369215	21374_6#149	090110_18826	blood	0	0	0	0	49	female	2010	95	8	worsen	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369216	21374_6#150	091211_31762	blood	0	1	0	0	57	female	2011	69	4	recover	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1369217	21374_6#151	110113_63731	blood	0	0	1	0	63	male	2013	410	6	nochange	no	S	I	S	R	S	S	S	S	S	R	R	R	S
ERS1369218	21374_6#152	130712_60834	blood	0	1	0	0	70	female	2012	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369219	21374_6#153	120612_14122	blood	0	0	0	0	76	male	2012	648	9	worsen	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	I
ERS1369220	21374_6#154	130412_23376	blood	0	0	0	0	72	female	2012	95	8	recover	no	S	I	S	S	S	S	S	S	S	S	R	R	S
ERS1369221	21374_6#155	131212_14879	blood	0	1	0	0	71	female	2012	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369222	21374_6#156	140813_23372	blood	0	0	0	0	63	male	2013	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369223	21374_6#157	090111_28849	blood	0	0	0	0	64	female	2011	1722	15	recover	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	R
ERS1369224	21374_6#158	100111_28944	blood	0	0	1	0	55	male	2011	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	S	S
ERS1369225	21374_6#159	110212_25160	blood	0	0	0	0	55	female	2012	73	7	recover	no	S	S	S	N	S	S	S	S	S	S	S	I	S
ERS1369226	21374_6#160	131212_27800	blood	0	0	0	0	54	male	2012	206	14	recover	no	S	S	S	I	S	S	S	S	S	I	R	S	S
ERS1369227	21374_6#161	120810_17357	blood	0	0	1	0	59	female	2010	4704	14	nochange	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369228	21374_6#162	130413_17159	blood	0	0	0	1	26	female	2013	354	2	worsen	no	S	R	R	R	R	S	S	S	S	R	R	R	S
ERS1369229	21374_6#163	140112_17022	blood	0	0	1	0	52	male	2012	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369230	21374_6#164	140912_31941	blood	0	1	0	0	64	male	2012	73	7	recover	no	S	I	S	S	S	S	N	S	S	S	S	R	I
ERS1369231	21374_6#165	090310_24803	blood	0	0	0	0	85	female	2010	361	14	recover	not tested	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369232	21374_6#166	100112_28410	blood	0	1	0	0	51	female	2012	998	10	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369233	21374_6#167	110213_15643	blood	0	0	1	0	40	male	2013	752	14	worsen	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369234	21374_6#168	111210_15437	blood	0	0	0	0	46	female	2010	131	12	recover	not tested	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369235	21374_6#169	120810_17399	blood	0	1	0	0	41	female	2010	69	4	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369236	21374_6#170	130512_11718	blood	0	0	0	0	1	male	2012	12	10	recover	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	S
ERS1369237	21374_6#171	140112_38074	blood	1	0	0	1	29	male	2012	354	2	death	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369238	21374_6#172	141011_17723	blood	0	0	0	0	70	male	2011	58	11	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369239	21374_6#173	090610_17708	blood	1	0	1	0	56	female	2010	38	3	worsen	not tested	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369241	21374_6#175	110213_65970	blood	0	0	0	0	54	female	2013	12	10	recover	ESBL	S	S	R	S	R	S	R	S	S	S	S	R	S
ERS1369242	21374_6#176	111210_26698	blood	0	0	0	0	47	male	2010	12	10	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369243	21374_6#177	120810_59518	blood	0	0	0	0	54	male	2010	648	9	recover	not tested	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1369244	21374_6#178	130513_39080	blood	0	0	0	1	32	male	2013	1722	15	recover	no	S	S	S	S	S	S	S	S	S	N	R	R	S
ERS1369245	21374_6#179	140211_62933	blood	0	0	0	0	57	female	2011	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	N	S
ERS1369246	21374_6#180	141111_28450	blood	0	0	0	0	61	female	2011	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369247	21374_6#181	090611_28541	blood	0	1	0	0	23	female	2011	131	12	recover	ESBL	S	I	R	S	R	S	R	S	S	S	R	R	S
ERS1369248	21374_6#182	100311_59210	blood	0	0	0	0	55	male	2011	354	2	nochange	no	S	R	R	N	R	S	S	S	S	N	R	N	S
ERS1369249	21374_6#183	110411_17579	blood	1	0	1	0	61	male	2011	69	4	worsen	no	S	S	S	N	S	S	S	S	S	N	R	I	S
ERS1369250	21374_6#184	120313_33618	blood	0	1	0	0	22	female	2013	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369251	21374_6#185	121013_11280	blood	0	0	0	0	0	male	2013	969	10	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369252	21374_6#186	130611_14674	blood	0	1	0	0	47	male	2011	131	12	nochange	ESBL	S	I	N	R	N	S	N	S	S	R	R	I	S
ERS1369253	21374_6#187	140412_17063	blood	1	0	0	0	42	male	2012	648	9	worsen	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369254	21374_6#188	141113_24428	blood	0	1	0	0	31	female	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369255	21374_6#189	151213_22608	blood	0	1	0	0	77	female	2013	58	11	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369256	21374_7#1	161010_14940	blood	0	0	0	0	93	male	2010	12	10	worsen	not tested	S	I	R	S	R	S	R	S	S	S	R	S	S
ERS1369257	21374_7#2	170910_16859	blood	0	0	0	0	72	male	2010	12	10	recover	not tested	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369258	21374_7#3	181013_79856	blood	0	0	0	0	47	male	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369259	21374_7#4	200213_22545	blood	0	1	0	0	44	female	2013	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369260	21374_7#5	210413_25416	blood	0	1	0	0	62	female	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369261	21374_7#6	220411_30083	blood	0	0	0	0	61	female	2011	14	13	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369262	21374_7#7	150311_59427	blood	0	0	0	0	45	male	2011	131	12	recover	ESBL	S	I	R	N	R	S	R	S	S	N	R	N	S
ERS1369263	21374_7#8	160213_32598	blood	0	1	0	0	47	female	2013	131	12	recover	no	S	I	S	R	S	S	S	S	S	R	R	I	I
ERS1369264	21374_7#9	161110_23558	blood	0	1	0	0	61	female	2010	95	8	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369265	21374_7#10	171211_58263	blood	0	1	0	0	70	male	2011	73	7	recover	no	S	R	S	S	S	S	S	S	S	S	R	R	S
ERS1369266	21374_7#11	181110_23652	blood	0	0	0	0	88	male	2010	131	12	nochange	not tested	S	S	R	R	R	S	R	S	S	R	I	I	S
ERS1369267	21374_7#12	200410_33742	blood	0	0	0	0	19	female	2010	69	4	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369268	21374_7#13	210510_14732	blood	0	1	0	0	82	female	2010	73	7	recover	not tested	S	R	S	S	S	S	S	S	S	S	S	S	S

ERS1369269	21374_7#14	220412_17835	blood	0	0	0	0	84	male	2012	405	1	worsen	no	S	R	R	R	R	S	S	S	S	R	S	R	S
ERS1369271	21374_7#16	160312_34208	blood	0	0	0	0	83	male	2012	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	S	S
ERS1369272	21374_7#17	161210_64854	blood	0	1	0	0	36	female	2010	3177	9	recover	not tested	S	S	R	R	R	S	R	S	S	R	N	R	R
ERS1369273	21374_7#18	180411_26186	blood	0	1	0	0	18	male	2011	624	9	nochange	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369274	21374_7#19	190111_25198	blood	0	0	0	0	67	male	2011	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369275	21374_7#20	200412_34318	blood	0	0	0	0	46	female	2012	394	5	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369276	21374_7#21	210912_62870	blood	0	0	0	0	64	female	2012	12	10	recover	no	S	I	S	S	S	S	S	S	S	S	S	R	S
ERS1369277	21374_7#22	220912_38772	blood	0	0	0	1	34	male	2012	131	12	worsen	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369278	21374_7#23	150513_29362	blood	0	1	0	0	51	female	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	S	R	S
ERS1369280	21374_7#25	170112_64769	blood	0	0	1	0	56	male	2012	405	1	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1369281	21374_7#26	180413_20302	blood	0	0	0	0	58	female	2013	69	4	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369282	21374_7#27	190210_15740	blood	1	0	1	0	45	male	2010	48	14	worsen	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369283	21374_7#28	200413_27917	blood	0	0	0	0	56	male	2013	131	12	recover	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369284	21374_7#29	211012_17801	blood	0	0	0	0	35	male	2012	624	9	worsen	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369285	21374_7#30	221011_26541	blood	0	1	0	0	57	female	2011	131	12	nochange	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369286	21374_7#31	150610_15560	blood	0	0	0	0	33	female	2010	969	10	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369287	21374_7#32	160412_17305	blood	0	0	0	0	82	female	2012	73	7	recover	no	I	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369288	21374_7#33	170310_17633	blood	0	0	1	0	65	female	2010	394	5	recover	not tested	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369289	21374_7#34	180413_63230	blood	0	0	1	1	39	female	2013	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369290	21374_7#35	190212_17108	blood	0	0	0	0	49	female	2012	1193	13	worsen	no	S	S	S	R	S	S	N	S	S	R	R	S	S
ERS1369291	21374_7#36	200610_65090	blood	0	0	0	0	40	male	2010	69	4	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369292	21374_7#37	211112_16469	blood	0	0	0	0	53	male	2012	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1369293	21374_7#38	221112_36416	blood	0	1	0	0	70	male	2012	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	R	S
ERS1369294	21374_7#39	150613_63444	blood	1	0	1	0	49	male	2013	38	3	worsen	ESBL	S	S	R	S	R	S	R	S	S	S	S	R	S
ERS1369295	21374_7#40	160610_15579	blood	0	0	0	0	63	male	2010	83	10	nochange	not tested	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369296	21374_7#41	170310_41641	blood	1	0	0	1	31	male	2010	131	12	nochange	not tested	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369297	21374_7#42	180511_30829	blood	0	1	0	0	56	female	2011	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369298	21374_7#43	190312_40740	blood	0	0	0	1	24	male	2012	405	1	worsen	ESBL	R	R	R	R	R	S	R	S	S	R	R	R	S
ERS1369299	21374_7#44	200810_14271	blood	0	0	1	0	44	male	2010	70	5	worsen	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369300	21374_7#45	211210_16960	blood	0	0	0	0	75	female	2010	38	3	recover	not tested	S	R	R	S	R	S	R	S	S	S	R	I	S
ERS1369301	21374_7#46	221213_26388	blood	0	0	0	0	47	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369302	21374_7#47	150910_16545	blood	0	1	1	0	63	male	2010	131	12	nochange	not tested	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1369303	21374_7#48	160610_15656	blood	0	0	1	0	73	female	2010	48	14	recover	not tested	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369304	21374_7#49	170412_38720	blood	0	0	0	1	26	male	2012	1722	15	death	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	I
ERS1369305	21374_7#50	180612_63824	blood	0	0	1	0	55	male	2012	421	8	recover	ESBL	S	S	R	S	R	S	R	S	S	S	S	R	S
ERS1369306	21374_7#51	190513_15714	blood	0	0	0	0	61	female	2013	1193	13	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369307	21374_7#52	200912_15265	blood	0	0	0	0	56	female	2012	48	14	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369308	21374_7#53	220210_24375	blood	0	0	0	0	65	male	2010	131	12	recover	not tested	S	R	S	R	R	S	S	S	S	R	R	S	S
ERS1369309	21374_7#54	230310_25278	blood	0	0	0	0	56	male	2010	58	11	recover	not tested	S	I	S	S	S	S	S	S	S	S	R	I	S
ERS1369310	21374_7#55	150912_62204	blood	0	0	1	0	76	female	2012	1193	13	recover	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369311	21374_7#56	160712_17087	blood	0	1	0	0	73	female	2012	1193	13	nochange	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1369312	21374_7#57	170413_63438	blood	0	0	1	0	61	female	2013	648	9	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	S
ERS1369313	21374_7#58	180813_23546	blood	0	0	0	0	23	male	2013	998	10	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369315	21374_7#60	200913_14489	blood	0	0	0	0	26	male	2013	131	12	recover	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1369316	21374_7#61	220212_40931	blood	0	0	0	1	25	male	2012	131	12	nochange	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369317	21374_7#62	230412_10585	blood	0	0	0	0	14	female	2012	410	6	recover	no	S	S	S	R	S	S	S	S	S	R	S	S	S
ERS1369318	21374_7#63	150913_14155	blood	0	0	0	0	79	female	2013	1193	13	recover	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369319	21374_7#64	160910_16741	blood	0	1	1	0	41	female	2010	424	11	worsen	not tested	S	I	S	R	S	S	S	S	S	R	R	I	S
ERS1369320	21374_7#65	170510_14388	blood	1	0	1	0	55	male	2010	38	3	worsen	not tested	S	S	R	S	R	S	R	S	S	S	S	S	S
ERS1369321	21374_7#66	180912_28943	blood	0	0	0	0	71	female	2012	95	8	recover	ESBL	S	S	R	S	R	S	R	S	S	S	S	R	S
ERS1369322	21374_7#67	191010_60577	blood	0	0	0	0	40	female	2010	95	8	recover	not tested	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369323	21374_7#68	201211_15050	blood	0	0	1	0	55	female	2011	69	4	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369324	21374_7#69	220213_21570	blood	0	0	0	0	83	female	2013	95	8	worsen	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369325	21374_7#70	230612_59785	blood	0	0	0	0	35	male	2012	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369326	21374_7#71	151012_33532	blood	0	1	0	0	49	female	2012	95	8	recover	no	S	R	R	S	R	S	S	S	S	S	R	R	S
ERS1369327	21374_7#72	160912_38033	blood	0	0	0	1	38	male	2012	12	10	recover	no	S	I	S	S	S	S	S	S	S	S	R	R	S
ERS1369328	21374_7#73	170511_16632	blood	0	0	0	0	69	female	2011	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369329	21374_7#74	180913_14365	blood	1	1	1	0	49	male	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369330	21374_7#75	191112_29598	blood	1	0	0	0	39	male	2012	38	3	recover	ESBL	S	R	R	R	R	I	R	S	S	R	R	R	I
ERS1369331	21374_7#76	210211_13696	blood	0	0	0	0	12	male	2011	648	9	recover	no	S	I	S	S	S	S	S	S	S	S	R	N	S
ERS1369332	21374_7#77	220213_28430	blood	0	1	0	0	82	male	2013	38	3	worsen	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	S
ERS1369333	21374_7#78	230713_31504	blood	0	1	0	0	32	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369335	21374_7#80	160913_25086	blood	0	0	0	0	26	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369336	21374_7#81	170710_24408	blood	0	0	0	0	57	female	2010	73	7	nochange	not tested	S	I	S	S	S	S	S	S	S	S	S	S	S
ERS1369337	21374_7#82	181011_14080	blood	0	0	0	0	52	male	2011	648	9	recover	no	S	I	R	S	I	S	S	S	S	I	R	R	S
ERS1369338	21374_7#83	191112_63467	blood	0	0	0	1	73	male	2012	46	14	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S

ERS1369339	21374_7#84	210410_27216	blood	0	0	0	0	57	female	2010	131	12	recover	not tested	S	I	R	R	R	S	R	S	S	R	R	S	S
ERS1369340	21374_7#85	220310_17922	blood	0	0	0	0	83	female	2010	73	7	recover	not tested	S	I	S	S	S	S	S	S	S	S	S	I	S
ERS1369341	21374_7#86	230910_39316	blood	0	0	0	1	27	male	2010	648	9	recover	not tested	S	I	R	R	R	S	R	S	S	R	R	R	R
ERS1369342	21374_7#87	151111_25407	blood	0	0	0	0	81	male	2011	410	6	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1369343	21374_7#88	160913_26331	blood	0	1	0	0	70	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369344	21374_7#89	170713_31171	blood	0	0	0	0	40	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369345	21374_7#90	181011_24017	blood	0	1	0	0	51	female	2011	95	8	recover	ESBL	S	I	R	S	R	S	R	S	S	S	R	R	S
ERS1369346	21374_7#91	200110_37250	blood	0	0	0	0	45	female	2010	83	10	nochange	not tested	R	I	S	S	S	S	S	S	S	S	S	I	S
ERS1369347	21374_7#92	210412_18156	blood	0	0	0	0	65	male	2012	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369348	21374_7#93	220313_15086	blood	0	0	0	0	82	male	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369349	21374_7#94	230913_14749	blood	0	0	0	0	73	female	2013	354	2	nochange	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369350	21374_7#95	240411_14784	blood	0	1	0	0	81	female	2011	1193	13	recover	no	S	S	S	R	S	S	S	S	S	R	S	I	S
ERS1369352	21374_7#97	251111_38569	blood	0	0	0	1	28	female	2011	410	6	nochange	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	I
ERS1369353	21374_7#98	270113_38499	blood	1	0	0	1	36	male	2013	10	14	death	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369354	21374_7#99	271213_29936	blood	0	0	0	0	35	female	2013	62	15	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369355	21374_7#100	290510_16912	blood	0	0	0	0	65	female	2010	95	8	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369356	21374_7#101	300611_16124	blood	1	0	0	0	59	male	2011	38	3	worsen	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369357	21374_7#102	231012_65175	blood	0	0	1	0	76	male	2012	48	14	nochange	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369358	21374_7#103	240411_18751	blood	0	1	0	0	52	male	2011	73	7	recover	no	S	I	S	S	S	S	S	S	S	S	R	R	S
ERS1369359	21374_7#104	241113_20386	blood	0	1	0	1	29	female	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369360	21374_7#105	251212_62621	blood	0	0	1	0	60	female	2012	88	6	nochange	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369361	21374_7#106	270210_24519	blood	0	1	0	0	68	female	2010	354	2	recover	not tested	S	I	S	R	S	S	S	S	S	R	R	I	S
ERS1369362	21374_7#107	280411_15098	blood	0	0	0	0	70	female	2011	69	4	worsen	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1369363	21374_7#108	290512_30440	blood	0	0	0	0	40	female	2012	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369364	21374_7#109	300611_65700	blood	0	0	1	0	64	female	2011	48	14	worsen	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369365	21374_7#110	231013_17462	blood	0	0	0	0	25	male	2013	2003	3	worsen	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1369366	21374_7#111	240510_40032	blood	0	0	0	1	30	male	2010	405	1	recover	not tested	S	R	R	R	R	S	R	S	S	R	R	I	S
ERS1369367	21374_7#112	241210_38427	blood	0	0	0	1	38	male	2010	998	10	worsen	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369368	21374_7#113	251213_38390	blood	0	1	0	1	29	female	2013	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369371	21374_7#116	290613_15256	blood	0	0	0	0	31	male	2013	1196	11	recover	no	S	I	S	R	S	S	S	S	S	R	R	R	S
ERS1369372	21374_7#117	300613_30142	blood	0	0	0	0	58	female	2013	69	4	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369373	21374_7#118	231110_40261	blood	0	0	0	1	33	male	2010	156	11	nochange	not tested	S	S	S	R	S	S	S	S	S	R	R	R	S
ERS1369374	21374_7#119	240511_31066	blood	0	0	0	0	74	female	2011	38	3	recover	no	S	R	S	N	S	S	S	S	S	S	R	R	I
ERS1369375	21374_7#120	241213_63038	blood	0	0	1	0	41	female	2013	767	11	nochange	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369376	21374_7#121	260113_14596	blood	0	0	0	0	78	female	2013	410	6	nochange	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	I
ERS1369377	21374_7#122	270213_33073	blood	0	0	0	0	55	female	2013	69	4	recover	no	S	R	S	S	S	S	S	S	S	S	R	R	S
ERS1369378	21374_7#123	280512_28240	blood	0	1	0	0	87	female	2012	69	4	recover	no	S	R	S	S	S	S	S	S	S	S	S	R	R
ERS1369379	21374_7#124	290813_24150	blood	0	0	0	0	56	female	2013	73	7	recover	no	S	I	S	S	S	S	S	S	S	S	S	R	S
ERS1369380	21374_7#125	300713_17762	blood	0	0	0	0	57	male	2013	1193	13	recover	no	S	S	S	R	S	S	S	S	S	R	R	R	S
ERS1369381	21374_7#126	231112_16756	blood	0	0	0	0	39	male	2012	1193	13	recover	no	S	S	S	R	S	S	S	S	S	R	R	R	S
ERS1369382	21374_7#127	240512_11939	blood	0	0	0	0	1	male	2012	1722	15	recover	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	S
ERS1369383	21374_7#128	250110_63710	blood	0	0	0	0	55	male	2010	624	9	recover	not tested	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1369384	21374_7#129	260313_44343	blood	0	0	0	0	1	male	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369385	21374_7#130	270312_28938	blood	0	1	0	0	24	female	2012	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369386	21374_7#131	280610_38514	blood	0	0	0	1	55	female	2010	131	12	recover	not tested	S	S	R	R	R	S	R	S	S	R	R	S	S
ERS1369387	21374_7#132	290913_15382	blood	0	0	0	0	54	male	2013	405	1	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369388	21374_7#133	300813_16771	blood	0	0	1	0	65	female	2013	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369389	21374_7#134	231112_63877	blood	1	0	1	0	60	male	2012	648	9	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369390	21374_7#135	240613_41098	blood	0	0	0	1	31	male	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369391	21374_7#136	250111_29994	blood	0	0	0	0	34	male	2011	75	11	death	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369392	21374_7#137	260511_17272	blood	0	0	0	0	37	female	2011	38	3	recover	no	S	R	R	N	R	S	I	S	S	R	R	R	I
ERS1369393	21374_7#138	270312_34295	blood	0	0	0	0	69	female	2012	131	12	recover	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	S
ERS1369394	21374_7#139	280611_15875	blood	0	0	0	0	45	male	2011	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369395	21374_7#140	291011_14882	blood	0	0	0	0	66	male	2011	52	14	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369396	21374_7#141	300913_31934	blood	0	1	0	0	23	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369397	21374_7#142	231210_27691	blood	0	1	0	0	52	female	2010	38	3	recover	not tested	S	R	R	R	R	S	R	S	S	R	S	R	I
ERS1369398	21374_7#143	240613_64466	blood	0	0	1	0	65	male	2013	648	9	recover	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	S
ERS1369399	21374_7#144	250313_26282	blood	0	1	0	0	52	female	2013	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369400	21374_7#145	260611_15759	blood	0	0	0	0	53	male	2011	453~	11	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369401	21374_7#146	270813_16280	blood	0	1	1	0	53	female	2013	38	3	nochange	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1369403	21374_7#147	291012_14405	blood	0	0	1	0	82	male	2012	405	1	nochange	no	S	R	R	R	R	S	R	S	S	R	R	R	I
ERS1369404	21374_7#148	301012_65908	blood	0	1	1	0	79	female	2012	405	1	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369406	21374_7#150	240812_39349	blood	0	0	0	1	39	male	2012	421	8	nochange	ESBL	S	S	R	S	R	S	R	S	S	S	R	S	S
ERS1369407	21374_7#151	250411_14897	blood	1	0	1	0	57	male	2011	1193	13	worsen	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369408	21374_7#152	260810_26401	blood	0	0	0	0	23	female	2010	648	9	recover	not tested	S	S	S	S	S	S	S	S	S	N	S	S	S
ERS1369409	21374_7#153	270813_39888	blood	0	0	0	1	31	male	2013	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S

ERS1369410	21374_7#154	281212_24521	blood	0	1	0	0	68	male	2012	131	12	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	S
ERS1369411	21374_7#155	291111_65777	blood	0	1	0	0	53	female	2011	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369412	21374_7#156	301111_29656	blood	0	1	0	0	66	female	2011	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369413	21374_7#157	240210_40634	blood	0	0	0	1	25	female	2010	38	3	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369414	21374_7#158	240812_64614	blood	0	0	1	0	51	male	2012	3014	14	recover	ESBL	S	S	I	S	R	S	S	S	S	R	S	R	S
ERS1369415	21374_7#159	250413_14080	blood	0	0	0	0	76	female	2013	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	I	S
ERS1369416	21374_7#160	260813_65252	blood	0	0	0	0	75	female	2013	3723~	10	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369417	21374_7#161	270910_34911	blood	0	0	0	0	43	male	2010	95	8	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369418	21374_7#162	281213_37649	blood	0	0	0	0	46	male	2013	73	7	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369419	21374_7#163	291211_63663	blood	1	0	1	0	61	male	2011	393	5	recover	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369420	21374_7#164	301213_23408	blood	0	1	0	0	32	female	2013	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	I	S
ERS1369421	21374_7#165	240310_38133	blood	0	0	0	1	32	male	2010	38	3	recover	not tested	S	R	R	R	R	S	S	S	S	R	S	R	I
ERS1369422	21374_7#166	240910_17728	blood	0	1	0	0	85	female	2010	131	12	recover	not tested	S	I	R	R	R	S	R	S	S	R	S	R	R
ERS1369423	21374_7#167	250512_30240	blood	0	0	0	0	77	female	2012	393	5	recover	no	S	R	R	R	R	S	S	S	S	R	R	R	I
ERS1369424	21374_7#168	260913_25623	blood	0	1	0	0	57	female	2013	349	5	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1369425	21374_7#169	271112_17165	blood	0	0	0	0	20	female	2012	648	9	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369426	21374_7#170	290110_19482	blood	0	0	0	0	78	female	2010	69	4	recover	not tested	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369427	21374_7#171	300111_16602	blood	0	0	0	0	41	female	2011	410	6	recover	ESBL	S	R	R	R	R	S	R	S	S	R	S	I	S
ERS1369428	21374_7#172	301213_26606	blood	0	0	0	0	63	female	2013	969	10	recover	no	S	I	S	S	S	S	S	S	S	S	R	R	I
ERS1369429	21374_7#173	240312_65905	blood	0	0	1	0	41	male	2012	69	4	recover	no	S	I	S	S	S	S	S	S	S	S	R	I	S
ERS1369430	21374_7#174	241012_65285	blood	0	1	1	0	60	female	2012	38	3	recover	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1369431	21374_7#175	250612_59884	blood	0	0	0	0	27	female	2012	127	10	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369432	21374_7#176	261012_30160	blood	0	1	0	0	54	female	2012	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369433	21374_7#177	271112_26332	blood	0	0	0	0	80	female	2012	69	4	recover	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369434	21374_7#178	290312_29062	blood	0	0	0	0	53	male	2012	131	12	recover	no	S	I	S	R	S	S	S	S	S	R	R	R	I
ERS1369435	21374_7#179	300413_14409	blood	0	0	0	0	79	male	2013	206	14	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369436	21374_7#180	310111_40848	blood	0	0	0	1	36	male	2011	38	3	nochange	no	S	R	R	R	R	S	S	S	S	R	R	R	R
ERS1369437	21374_7#181	240410_16433	blood	0	0	0	0	73	female	2010	453	11	recover	not tested	S	I	S	R	S	S	S	S	S	R	R	I	S
ERS1369438	21374_7#182	241111_65039	blood	1	0	1	0	56	male	2011	354	2	worsen	ESBL	S	R	R	R	R	R	R	S	S	R	S	R	S
ERS1369439	21374_7#183	250713_26493	blood	0	0	0	0	57	female	2013	624	9	recover	ESBL	S	R	R	R	R	R	R	S	S	R	R	R	S
ERS1369440	21374_7#184	261211_23317	blood	0	1	0	0	51	female	2011	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	S	S
ERS1369441	21374_7#185	271213_23259	blood	0	1	0	0	65	female	2013	405	1	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1369442	21374_7#186	290413_58223	blood	0	0	0	0	33	female	2013	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369443	21374_7#187	300511_25855	blood	0	1	0	0	63	female	2011	131	12	recover	ESBL	S	I	R	N	R	S	R	S	S	R	R	R	R
ERS1369444	21374_7#188	310310_63150	blood	0	0	1	0	64	male	2010	2914	5	worsen	not tested	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369445	21374_7#189	080114_27098	blood	0	0	0	0	72	female	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369446	21528_7#1	220214_14599	blood	1	0	0	0	50	male	2014	69	4	worsen	ESBL	S	S	R	S	R	S	R	S	S	S	R	I	S
ERS1369447	21528_7#2	270314_40994	blood	0	0	0	1	44	male	2014	1193	13	recover	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369448	21528_7#3	060514_38380	blood	0	0	0	1	46	male	2014	1193	13	recover	ESBL	S	S	R	R	R	I	R	S	S	R	R	I	S
ERS1369449	21528_7#4	190614_65107	blood	0	1	0	0	87	female	2014	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369450	21528_7#5	200714_26593	blood	0	0	0	0	67	female	2014	5150	3	recover	no	S	R	R	I	R	I	S	S	S	I	S	I	S
ERS1369451	21528_7#6	140914_62966	blood	1	0	0	0	54	female	2014	48	14	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369452	21528_7#7	310312_16067	blood	0	0	0	0	60	male	2012	48	14	worsen	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369453	21528_7#8	090114_31772	blood	0	1	0	0	32	female	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369454	21528_7#9	240214_26053	blood	0	0	0	0	64	female	2014	69	4	nochange	no	S	I	S	S	S	S	S	S	S	S	R	R	S
ERS1369455	21528_7#10	280314_62187	blood	0	1	0	0	59	male	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369456	21528_7#11	070514_65680	blood	1	0	1	0	61	male	2014	4381	14	nochange	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369457	21528_7#12	240614_34059	blood	0	0	0	0	50	male	2014	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369458	21528_7#13	280714_23051	blood	0	1	0	0	59	female	2014	1193	13	recover	ESBL	N	S	R	R	R	S	R	S	S	R	S	R	S
ERS1369459	21528_7#14	260914_28414	blood	0	0	0	0	64	female	2014	3672	10	nochange	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369460	21528_7#15	310313_15973	blood	1	0	1	0	49	female	2013	1193	13	recover	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1369461	21528_7#16	090114_27233	blood	0	0	0	0	77	female	2014	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369462	21528_7#17	240214_38904	blood	0	0	0	1	39	male	2014	405	1	recover	no	S	R	R	R	R	R	R	R	R	R	R	R	R
ERS1369463	21528_7#18	010414_62490	blood	0	1	1	0	60	female	2014	1724	11	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369464	21528_7#19	090514_25798	blood	0	0	0	0	71	female	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369465	21528_7#20	240614_65620	blood	0	0	1	0	63	female	2014	394	5	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369466	21528_7#21	300714_27372	blood	0	0	0	0	82	female	2014	648	9	nochange	ESBL	S	I	R	R	R	S	R	S	S	R	S	I	S
ERS1369467	21528_7#22	300914_64365	blood	0	0	1	0	50	male	2014	354	2	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	S
ERS1369468	21528_7#23	310511_40047	blood	0	0	0	1	27	male	2011	421	8	worsen	ESBL	S	S	R	N	R	R	R	S	S	S	R	S	S
ERS1369469	21528_7#24	150114_16293	blood	0	1	0	0	70	female	2014	73	7	recover	no	S	I	S	S	S	S	S	S	S	S	R	R	S
ERS1369470	21528_7#25	250214_30035	blood	0	0	0	0	24	female	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369471	21528_7#26	070414_37228	blood	0	0	0	0	54	male	2014	38	3	recover	no	I	R	S	R	S	S	S	S	S	R	S	R	I
ERS1369472	21528_7#27	090514_15608	blood	0	0	0	0	24	female	2014	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	I	S
ERS1369474	21528_7#29	010814_58240	blood	1	0	1	0	50	male	2014	14	13	worsen	no	S	S	S	N	S	S	S	S	S	S	R	I	S
ERS1369475	21528_7#30	300914_28670	blood	0	1	0	0	28	female	2014	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	S	S
ERS1369476	21528_7#31	310511_63560	blood	0	0	0	0	72	female	2011	405	1	recover	ESBL	S	I	R	N	R	S	R	S	S	R	R	R	R

ERS1369477	21528_7#32	310114_24944	blood	0	0	0	0	57	female	2014	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369478	21528_7#33	280214_63888	blood	0	1	1	0	60	male	2014	6823	11	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	I
ERS1369479	21528_7#34	070414_24073	blood	0	1	1	0	44	male	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369480	21528_7#35	100514_15663	blood	0	0	0	0	50	male	2014	167	14	nochange	ESBL	S	R	R	R	R	S	R	S	S	R	R	I	S
ERS1369481	21528_7#36	280614_32800	blood	0	0	0	0	24	female	2014	127	10	recover	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1369482	21528_7#37	030814_23341	blood	0	1	0	0	66	female	2014	73	7	recover	no	S	I	S	S	S	S	S	S	S	S	S	R	S
ERS1369483	21528_7#38	011014_64531	blood	0	0	1	0	59	male	2014	648	9	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369484	21528_7#39	310812_27582	blood	0	0	0	0	40	female	2012	44	14	nochange	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369485	21528_7#40	040214_38507	blood	0	0	0	1	34	male	2014	38	3	nochange	ESBL	S	I	R	S	R	S	R	S	S	S	R	R	I
ERS1369486	21528_7#41	050314_15390	blood	0	0	0	0	62	female	2014	224	11	recover	ESBL	S	I	R	R	R	R	R	R	R	N	S	R	R
ERS1369487	21528_7#42	080414_29092	blood	0	0	0	0	54	female	2014	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369488	21528_7#43	110514_25928	blood	0	1	0	0	76	male	2014	131	12	worsen	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	S
ERS1369489	21528_7#44	300614_24541	blood	0	0	0	0	82	female	2014	131	12	nochange	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369490	21528_7#45	030814_31413	blood	0	0	0	0	41	female	2014	73	7	recover	no	S	I	S	S	S	S	S	S	S	S	S	R	I
ERS1369491	21528_7#46	101014_32454	blood	0	1	0	0	20	female	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1369492	21528_7#47	311012_38900	blood	0	0	0	1	39	male	2012	95	8	recover	no	S	S	S	S	S	S	S	S	S	N	R	R	S
ERS1369493	21528_7#48	070214_17751	blood	0	0	0	0	70	female	2014	12	10	recover	ESBL	S	S	R	S	R	S	R	S	S	S	S	I	S
ERS1369494	21528_7#49	060314_26839	blood	0	1	0	0	59	female	2014	405	1	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	S
ERS1369495	21528_7#50	080414_62581	blood	0	0	1	0	49	male	2014	131	12	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	I
ERS1369496	21528_7#51	150514_22323	blood	0	0	0	0	71	female	2014	131	12	recover	no	S	I	S	R	S	S	S	S	S	R	R	R	S
ERS1369497	21528_7#52	010714_62215	blood	0	0	1	0	59	male	2014	648	9	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369498	21528_7#53	040814_31526	blood	0	0	0	0	54	female	2014	12	10	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369499	21528_7#54	161014_26100	blood	0	0	0	0	40	female	2014	69	4	recover	no	S	I	S	S	S	S	S	S	S	S	R	R	S
ERS1369500	21528_7#55	311013_62583	blood	0	1	0	0	70	female	2013	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	S	I	S
ERS1369501	21528_7#56	110214_35530	blood	0	0	0	0	19	male	2014	405	1	recover	ESBL	S	R	R	R	R	I	R	S	S	R	R	R	R
ERS1369502	21528_7#57	120314_19631	blood	0	0	0	0	29	male	2014	38	3	recover	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1369503	21528_7#58	210414_35060	blood	0	0	0	0	30	male	2014	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369504	21528_7#59	300514_63660	blood	1	0	0	0	61	male	2014	127	10	recover	no	S	R	S	S	S	S	S	S	S	S	R	R	R
ERS1369505	21528_7#60	040714_14504	blood	0	1	0	0	84	female	2014	131	12	death	no	S	I	S	R	S	S	S	S	S	R	R	R	S
ERS1369506	21528_7#61	160814_16268	blood	0	0	0	0	23	male	2014	2705	14	recover	ESBL	S	I	R	S	R	N	R	S	S	N	R	R	S
ERS1369507	21528_7#62	171014_65915	blood	0	0	0	0	64	male	2014	38	3	recover	no	S	R	R	R	R	S	S	S	S	R	R	R	S
ERS1369508	21528_7#63	311213_26706	blood	0	1	0	0	43	male	2013	410	6	nochange	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369509	21528_7#64	140214_35573	blood	0	0	0	0	80	male	2014	405	1	recover	ESBL	S	I	R	R	R	S	R	S	S	R	S	R	I
ERS1369510	21528_7#65	150314_36751	blood	0	0	0	0	54	male	2014	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369511	21528_7#66	220414_38635	blood	0	0	0	1	43	male	2014	73	7	nochange	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369512	21528_7#67	080614_37350	blood	0	0	0	0	46	female	2014	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369513	21528_7#68	090714_62800	blood	0	1	1	0	49	male	2014	38	3	recover	ESBL	S	I	R	R	R	N	R	S	S	R	R	R	R
ERS1369514	21528_7#69	230814_64725	blood	1	0	1	1	32	male	2014	405	1	worsen	ESBL	S	I	R	N	R	R	R	S	S	R	R	R	R
ERS1369515	21528_7#70	211014_26474	blood	0	1	0	0	52	male	2014	405	1	recover	ESBL	S	R	R	R	R	S	R	S	S	N	S	R	R
ERS1369516	21528_7#71	020114_31318	blood	0	0	0	0	51	female	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369517	21528_7#72	180214_41365	blood	0	0	0	1	43	female	2014	405	1	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	S
ERS1369518	21528_7#73	240314_16752	blood	0	0	0	0	38	female	2014	648	9	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369519	21528_7#74	270414_14826	blood	0	0	0	0	61	female	2014	1193	13	recover	ESBL	S	I	R	S	R	S	R	S	S	N	R	I	I
ERS1369520	21528_7#75	160614_64690	blood	1	0	1	0	64	male	2014	354	2	recover	ESBL	S	S	R	R	R	S	R	S	S	R	R	I	S
ERS1369521	21528_7#76	110714_63141	blood	0	0	1	0	63	male	2014	38	3	recover	no	S	R	R	S	R	S	R	S	I	S	S	R	S
ERS1369522	21528_7#77	070914_62168	blood	1	0	1	0	50	male	2014	131	12	recover	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	S
ERS1369523	21528_7#78	261014_62619	blood	0	0	1	0	50	male	2014	38	3	nochange	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	S
ERS1369524	21528_7#79	060114_23846	blood	0	0	0	0	67	male	2014	73	7	recover	no	S	S	S	I	S	S	S	S	S	S	R	I	S
ERS1369525	21528_7#80	200214_35781	blood	0	0	0	0	80	male	2014	405	1	recover	no	I	R	R	R	R	S	R	S	S	R	R	R	R
ERS1369526	21528_7#81	240314_23329	blood	0	1	0	0	73	female	2014	38	3	recover	ESBL	R	R	R	R	R	S	R	S	S	R	R	R	R
ERS1369527	21528_7#82	250414_10271	blood	0	0	0	0	2	female	2014	410	6	.	ESBL	S	R	R	R	R	S	S	S	S	R	N	R	R
ERS1369528	21528_7#83	180614_65040	blood	0	0	1	0	70	female	2014	131	12	recover	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1369529	21528_7#84	120714_63181	blood	0	0	1	0	38	male	2014	131	12	recover	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	I
ERS1369530	21528_7#85	110914_62662	blood	0	0	1	0	61	female	2014	10	14	recover	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1369531	21528_7#86	291014_38556	blood	0	0	0	1	29	male	2014	131	12	worsen	ESBL	S	I	R	R	R	S	R	S	S	R	S	R	S
ERS1369532	21528_7#87	070114_64030	blood	0	0	1	0	42	male	2014	131	12	nochange	ESBL	R	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369533	21528_7#88	200214_38821	blood	0	0	0	1	25	female	2014	131	12	nochange	ESBL	S	S	R	R	R	S	R	S	S	R	R	R	S
ERS1369534	21528_7#89	260314_23426	blood	0	1	0	0	74	female	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369535	21528_7#90	050514_65198	blood	0	1	1	0	49	male	2014	405	1	recover	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	I
ERS1369536	21528_7#91	190614_65151	blood	1	0	1	0	49	male	2014	453	11	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369537	21528_7#92	130714_30097	blood	0	1	0	0	64	female	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369538	21528_7#93	110914_62793	blood	0	0	1	0	41	male	2014	453	11	recover	ESBL	S	R	R	R	R	S	R	S	S	R	R	I	S
ERS1369539	21528_7#94	051114_27523	blood	0	1	0	0	63	female	2014	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	R	S
ERS1369540	21528_7#95	011214_29828	blood	0	0	0	0	53	male	2014	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1369544	21528_7#96	020412_29138	blood	0	0	0	0	24	female	2012	405	1	recover	no	S	R	S	R	R	S	R	S	S	R	R	R	S
ERS1369547	21528_7#97	051114_63612	blood	0	0	1	0	42	male	2014	354	2	recover	no	S	R	R	R	R	S	S	S	S	R	S	R	S

ERS1369548	21528_7#98	021214_22711	blood	0	0	0	0	63	female	2014	69	4	recover	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1369555	21528_7#99	081114_63951	blood	0	0	1	0	38	male	2014	131	12	nochange	ESBL	S	I	R	R	R	S	R	S	S	N	R	R	S
ERS1369556	21528_7#100	091214_62682	blood	0	0	1	0	35	female	2014	453	11	recover	no	S	S	S	S	S	S	S	S	S	S	S	R	S
ERS1369563	21528_7#101	091114_38685	blood	0	0	0	1	34	male	2014	2705	14	worsen	no	S	R	R	R	R	S	S	S	S	N	R	R	S
ERS1369564	21528_7#102	131214_26918	blood	0	1	0	0	80	male	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1369571	21528_7#103	091114_27828	blood	0	0	0	0	86	female	2014	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	N	R	R	S
ERS1369572	21528_7#104	141214_18149	blood	0	0	0	0	60	female	2014	131	12	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	R	S
ERS1369579	21528_7#105	101114_27962	blood	0	0	0	0	48	female	2014	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	R	S	R	R
ERS1369580	21528_7#106	141214_16448	blood	0	0	1	0	63	male	2014	1722	15	worsen	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	S
ERS1369587	21528_7#107	101114_64070	blood	0	0	1	0	35	female	2014	38	3	recover	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1369588	21528_7#108	191214_23791	blood	0	1	0	0	43	female	2014	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369595	21528_7#109	131114_64463	blood	0	1	0	0	49	female	2014	1193	13	recover	ESBL	S	S	R	R	R	S	R	S	S	N	R	I	S
ERS1369596	21528_7#110	221214_31521	blood	0	0	0	0	22	female	2014	457	15	recover	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	I
ERS1369603	21528_7#111	141114_14352	blood	0	0	0	0	28	female	2014	224	11	recover	no	S	S	S	R	S	S	S	S	S	R	S	I	S
ERS1369604	21528_7#112	221214_63810	blood	0	0	1	0	35	female	2014	617	14	recover	no	S	S	S	R	S	S	S	S	S	R	R	R	S
ERS1369611	21528_7#113	171114_28604	blood	0	0	0	0	79	female	2014	683	11	recover	no	S	R	R	S	R	S	S	S	S	S	R	R	I
ERS1369612	21528_7#114	101214_16015	blood	0	1	0	0	48	female	2014	38	3	worsen	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	S
ERS1369613	21528_7#115	241212_38008	blood	0	0	0	1	29	male	2012	38	3	nochange	no	S	R	S	R	S	S	S	S	S	R	R	R	I
ERS1369619	21528_7#116	171114_41043	blood	0	0	0	1	29	male	2014	12	10	worsen	no	S	S	S	S	S	S	S	S	S	N	S	S	S
ERS1369620	21528_7#117	271214_38789	blood	0	0	0	1	36	male	2014	969	10	worsen	no	S	I	S	S	S	S	S	S	S	S	R	R	S
ERS1369621	21528_7#118	030612_30623	blood	0	0	0	0	68	female	2012	69	4	recover	no	S	R	R	S	R	S	S	S	S	S	R	R	S
ERS1369627	21528_7#119	281114_38331	blood	0	0	0	1	56	female	2014	131	12	nochange	ESBL	S	I	R	R	R	S	S	S	S	N	S	I	S
ERS1369628	21528_7#120	281214_28015	blood	0	0	0	0	62	female	2014	95	8	recover	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1451217	21836_4#1	100115_14032	blood	0	0	0	1	27	female	2015	648	9	death	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1451218	21836_4#2	250315_63665	blood	0	0	1	0	53	female	2015	69	4	recover	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S
ERS1451219	21836_4#3	080215_16203	blood	0	0	0	0	67	female	2015	410	6	nochange	no	S	I	S	R	S	S	S	S	S	R	S	R	S
ERS1451225	21836_4#9	160315_63225	blood	0	0	1	0	31	male	2015	394	5	recover	ESBL	S	S	R	S	R	S	R	S	S	S	R	S	S
ERS1451226	21836_4#10	210815_15032	blood	0	0	1	0	55	female	2015	448	11	nochange	ESBL	S	R	R	N	R	S	S	S	S	N	R	R	I
ERS1451227	21836_4#11	311015_17508	blood	0	0	0	0	68	female	2015	1193	13	death	ESBL	S	S	R	N	R	S	R	S	S	N	R	S	S
ERS1451233	21836_4#17	010815_34451	blood	0	0	0	0	37	male	2015	NA	12	nochange	ESBL	I	R	R	N	R	S	R	S	S	N	R	R	S
ERS1451234	21836_4#18	061115_28103	blood	0	1	0	0	55	female	2015	1193	13	nochange	ESBL	S	S	R	N	R	S	R	S	S	N	S	I	S
ERS1451235	21836_4#19	201114_14705	blood	0	0	1	0	56	female	2014	5150	3	recover	ESBL	S	R	R	I	R	S	R	S	S	N	S	R	S
ERS1451241	21836_4#25	050915_16464	blood	0	1	0	0	40	female	2015	73	7	recover	no	S	I	S	N	S	S	S	S	S	N	R	I	S
ERS1451242	21836_4#26	310115_63575	blood	1	0	1	0	57	female	2015	3052	3	death	ESBL	S	S	R	S	R	S	R	S	S	S	S	R	S
ERS1451249	21836_4#33	160815_14563	blood	0	0	0	0	41	female	2015	126	10	nochange	no	S	S	S	N	S	S	S	S	S	N	S	S	S
ERS1451250	21836_4#34	061015_15322	blood	0	0	0	0	58	female	2015	131	12	nochange	ESBL	S	S	R	N	R	S	R	S	S	N	R	S	S
ERS1451258	21836_4#42	240815_15273	blood	0	0	1	0	59	male	2015	69	4	nochange	no	S	I	S	N	S	S	S	S	S	N	R	I	S
ERS1451265	21836_4#49	210715_16270	blood	0	0	0	0	44	female	2015	38	3	nochange	no	S	S	S	S	S	S	S	S	S	S	R	R	S
ERS1451273	21836_4#57	260515_15968	blood	1	0	1	0	46	male	2015	648	9	worsen	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451274	21836_4#58	301115_16266	blood	0	0	0	0	60	female	2015	1193	13	recover	no	S	I	S	N	S	S	S	S	S	N	S	I	S
ERS1451281	21836_4#65	300515_16242	blood	0	1	0	0	47	female	2015	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	I	S
ERS1451282	21836_4#66	210515_15556	blood	1	0	1	0	61	male	2015	1193	13	nochange	ESBL	S	S	R	R	R	S	R	S	S	R	R	S	S
ERS1451289	21836_4#73	100315_17879	blood	0	0	1	0	50	female	2015	354	2	recover	ESBL	S	R	R	R	R	S	S	S	S	R	S	R	S
ERS1451297	21836_4#81	290915_14739	blood	1	0	1	0	53	male	2015	69	4	nochange	ESBL	S	R	R	N	R	S	R	S	S	N	R	I	S
ERS1451298	21836_4#82	010615_16405	blood	0	0	0	0	64	female	2015	131	12	recover	ESBL	S	I	R	R	R	S	R	S	S	R	R	R	I
ERS1451305	21836_4#89	071115_14091	blood	0	0	0	0	53	male	2015	602	11	recover	ESBL	S	I	R	N	R	S	R	S	S	N	R	R	S
ERS1451306	21836_4#90	010215_15680	blood	0	0	0	0	67	female	2015	131	12	nochange	ESBL	S	S	R	S	R	S	R	S	S	S	R	R	S

Supplementary table S2. Details and epidemiological data for 159 carriage *E. coli* isolates from HTD.

Sample_accessio	lane	SID	infection_site	ST_getST	HierBAPS1	ESBL	AMK1	AMC1	CAZ1	CIP1	CRO1	ETP1	FEP1	IPM1	MEM1	OFX1	SXT1	TCC1	TZP1
ERS1451243	21836_4#27	EI298_NLF	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451251	21836_4#35	EI298_LF	rectal swab	38	3	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451259	21836_4#43	EI815_LF	rectal swab	394	5	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451267	21836_4#51	EI2793_NLF	rectal swab	354	2	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451275	21836_4#59	EI2793_LF2	rectal swab	167	14	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451283	21836_4#67	EI3399_NLF	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451291	21836_4#75	EI3399_LF	rectal swab	155	11	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451299	21836_4#83	EI3097_LF4	rectal swab	4040	5	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451307	21836_4#91	EI3097_LF1	rectal swab	58	11	ESBL	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1451220	21836_4#4	EI3097_LF2	rectal swab	4040	5	ESBL	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1451228	21836_4#12	EI777_LF	rectal swab	648	9	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1451236	21836_4#20	EI2619_LF4	rectal swab	23	6	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1451244	21836_4#28	EI2619_LF2	rectal swab	38	3	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451252	21836_4#36	EI1749_NLF	rectal swab	130	5	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451260	21836_4#44	EI1749_LF2	rectal swab	648	9	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451268	21836_4#52	EI1808_NLF	rectal swab	1193	13	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451276	21836_4#60	EI1808_LF2	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451284	21836_4#68	EI772_LF2	rectal swab	354	2	ESBL	R	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451292	21836_4#76	EI746_LF1	rectal swab	617	14	no	S	S	S	R	S	S	S	S	S	R	S	I	S
ERS1451300	21836_4#84	EI3807_LF	rectal swab	4306	5	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451308	21836_4#92	EI4355_LF_m	rectal swab	602	11	ESBL	R	R	R	R	R	R	R	R	R	R	R	R	R
ERS1451221	21836_4#5	EI3127_LF	rectal swab	448	11	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451229	21836_4#13	EI4341_BhaeNLF	rectal swab	1193	13	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	R
ERS1451237	21836_4#21	EI534_LF	rectal swab	3052	3	ESBL	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1451245	21836_4#29	EI3892_LF	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451253	21836_4#37	EI3892_Bhae	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451261	21836_4#45	EI3892_LF1	rectal swab	44	14	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451269	21836_4#53	EI3217_LF2	rectal swab	58	11	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451277	21836_4#61	EI3217_LF3	rectal swab	58	11	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451285	21836_4#69	EI2721_NLF	rectal swab	1193	13	no	S	S	S	R	S	S	S	S	S	R	S	S	S
ERS1451293	21836_4#77	EI4625_NLF	rectal swab	69	4	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451301	21836_4#85	EI4625_LF	rectal swab	5415	11	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451309	21836_4#93	EI1676_NLF	rectal swab	1193	13	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451222	21836_4#6	EI3211_NLF	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451230	21836_4#14	EI1537_BhaeNLF	rectal swab	6698	10	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451238	21836_4#22	EI1537_LF	rectal swab	131	12	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451246	21836_4#30	EI544_NLF	rectal swab	131	12	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451254	21836_4#38	EI544_LF	rectal swab	31	5	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1451262	21836_4#46	EI644_LF	rectal swab	5208	14	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1451270	21836_4#54	EI4263_LF	rectal swab	34	14	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451278	21836_4#62	EI4263_NLF	rectal swab	1193	13	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451286	21836_4#70	EI32_NLF	rectal swab	1193	13	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451302	21836_4#86	EI36_LF	rectal swab	167	14	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451310	21836_4#94	EI103_NLF	rectal swab	5150	3	ESBL	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1451223	21836_4#7	EI138_NLF	rectal swab	4381	14	no	S	I	S	S	S	S	S	S	S	S	S	S	S
ERS1451231	21836_4#15	EI138_LF	rectal swab	515	14	no	S	R	R	S	R	S	S	S	S	S	R	I	S

ERS1451239	21836_4#23	EI1888_LF2	rectal swab	93	14	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451247	21836_4#31	EI1787_NLF	rectal swab	130	5	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451255	21836_4#39	EI1787_LF2	rectal swab	648	9	no	R	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451263	21836_4#47	EI1830_LF2	rectal swab	93	14	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451271	21836_4#55	EI1605_NLF	rectal swab	6698	10	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451279	21836_4#63	EI1605_BhaeLF	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451287	21836_4#71	EI1761_NLF	rectal swab	6698	10	ESBL	S	R	R	R	R	S	R	S	S	R	R	I	S
ERS1451295	21836_4#79	EI1864_LF	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451303	21836_4#87	EI1564_Bhae	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451311	21836_4#95	EI1705_LF2	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451224	21836_4#8	EI1761_LF2	rectal swab	131	12	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451232	21836_4#16	EI1564_LF	rectal swab	43	14	ESBL	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1451240	21836_4#24	EI191_LF	rectal swab	10	14	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1451248	21836_4#32	EI950_LF	rectal swab	354	2	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451272	21836_4#56	EI3449_NLF	rectal swab	38	3	ESBL	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1451288	21836_4#72	EI3449_LF2	rectal swab	1722	15	ESBL	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1451280	21836_4#64	EI3449_LF1	rectal swab	88	6	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451296	21631_1#80	EI3185_Ecoli	rectal swab	448	11	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451304	21836_4#88	EI4381_LF	rectal swab	69	4	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1451312	21631_1#96	EI4381_NLF	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451320	21836_4#104	EI1755_LF	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451360	21836_4#143	EI1880_Ecoli	rectal swab	569	10	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451368	21836_4#151	EI1882_Ecoli	rectal swab	6211	11	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451313	21836_4#97	EI1994_Ecoli	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451328	21836_4#112	EI1765_LF1	rectal swab	450	14	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1451336	21836_4#120	EI1818_Ecoli	rectal swab	7156	5	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451400	21836_4#183	EI1951_LF	rectal swab	2073	11	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451376	21631_1#159	EI1896_Ecoli	rectal swab	46	14	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1451344	21836_4#128	EI1870_Ecoli	rectal swab	10	14	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451392	21631_1#175	EI1941_Ecoli	rectal swab	1380	5	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451384	21836_4#167	EI1912_Ecoli	rectal swab	5995	14	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451322	21631_1#106	EI2358_Bhae	rectal swab	405	1	ESBL	S	R	R	I	R	S	R	S	S	I	R	R	R
ERS1451337	21836_4#121	EI2302_LF1	rectal swab	44~	14	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1451355	21836_4#138	EI2129_LF3	rectal swab	44~	14	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1451379	21836_4#162	EI2129_NLF	rectal swab	95	8	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451345	21836_4#129	EI2314_NLF	rectal swab	NA	12	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451353	21836_4#136	EI2316_LF2	rectal swab	973	5	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451361	21836_4#144	EI2316_NLF1	rectal swab	69	4	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451329	21836_4#113	EI2296_LF	rectal swab	69	4	no	S	S	S	I	S	S	S	S	S	S	R	I	S
ERS1451377	21836_4#160	EI2339_BhaeLF2	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451385	21836_4#168	EI2347_LF	rectal swab	69	4	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1451314	21836_4#98	EI2353_NLF	rectal swab	227	14	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451294	21836_4#78	EI2453_BhaeLF2	rectal swab	12	10	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451256	21836_4#40	EI2455_LF2	rectal swab	4450	11	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451264	21631_1#48	EI2465_NLF	rectal swab	1193	13	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1451386	21836_4#169	EI3000_LF	rectal swab	4429	14	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451323	21836_4#107	EI3063_LF	rectal swab	38~	3	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451354	21836_4#137	EI2872_LF2	rectal swab	38~	3	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451394	21836_4#177	EI3004_LF	rectal swab	442	11	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R

ERS1451402	21836_4#185	EI3006_LF2	rectal swab	2016	11	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451338	21836_4#122	EI2828_LF2	rectal swab	224	11	no	S	R	R	R	R	S	S	S	S	R	R	I	S
ERS1451315	21836_4#99	EI3014_LF2	rectal swab	38	3	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451362	21836_4#145	EI2888_LF	rectal swab	617	14	no	S	S	S	R	S	S	S	S	S	R	R	I	R
ERS1451330	21836_4#114	EI2822_LF2	rectal swab	297	11	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451370	21836_4#153	EI2949_LF	rectal swab	69	4	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451378	21836_4#161	EI2987_Bhae	rectal swab	127	10	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451316	21836_4#100	EI3644_LF	rectal swab	457	15	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451339	21836_4#123	EI3425_LF1	rectal swab	10	14	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1451396	21836_4#179	EI3648_Bhae	rectal swab	969	10	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451371	21836_4#154	EI3503_LF	rectal swab	131	12	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451347	21631_1#131	EI3455_NLF	rectal swab	1193	13	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451325	21836_4#109	EI3964_LF1	rectal swab	617~	14	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451356	21836_4#139	EI3861_LF1	rectal swab	617	14	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R
ERS1451363	21836_4#146	EI3467_Bhae	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451369	21836_4#152	EI3966_LF	rectal swab	10	14	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451332	21836_4#116	EI3821_LF	rectal swab	457	15	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451364	21836_4#147	EI3865_LF1	rectal swab	4381	14	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451395	21836_4#178	EI3602_LF	rectal swab	641	11	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451324	21836_4#108	EI3637_NLF	rectal swab	409	14	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451372	21836_4#155	EI3867_LF2	rectal swab	761	14	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451380	21836_4#163	EI3873_LF1	rectal swab	NA	1	ESBL	S	R	R	R	R	I	R	S	S	R	R	R	R
ERS1451333	21836_4#117	EI3974_NLF	rectal swab	10	14	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451340	21836_4#124	EI3976_LF2	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451388	21836_4#171	EI3875_LF1	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451404	21631_1#187	EI3896_LF	rectal swab	2076	5	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451317	21836_4#101	EI3947_Bhae	rectal swab	69	4	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451387	21836_4#170	EI4108_Bhae	rectal swab	127	10	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451349	21836_4#132	EI4005_NLF	rectal swab	1193	13	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451403	21836_4#186	EI4124_NLF2	rectal swab	452~	10	no	S	S	S	S	S	S	S	S	S	S	R	S	S
ERS1451318	21836_4#102	EI4457_LF1	rectal swab	46	14	no	S	S	S	R	S	S	S	S	S	R	R	S	S
ERS1451357	21631_1#140	EI4415_NLF	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451341	21836_4#125	EI4309_LF	rectal swab	394	5	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1451373	21836_4#156	EI4379_LF	rectal swab	205	11	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451365	21836_4#148	EI4353_LF2	rectal swab	1196	11	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1451397	21836_4#180	EI4397_LF2	rectal swab	224	11	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451405	21836_4#188	EI4435_NLF	rectal swab	393	5	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451381	21836_4#164	EI4431_LF	rectal swab	10	14	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451389	21836_4#172	EI4475_LF2	rectal swab	10	14	no	S	S	S	I	S	S	S	S	S	S	R	S	S
ERS1451342	21836_4#126	EI4575_NLF	rectal swab	218	14	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451350	21836_4#133	EI4685_NLF	rectal swab	1193	13	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	R
ERS1451358	21836_4#141	EI4689_LF	rectal swab	648	9	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1451374	21836_4#157	EI4708_LF	rectal swab	4684	11	ESBL	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1451352	21836_4#135	EI4927_LF	rectal swab	6778	11	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1451382	21836_4#165	EI4727_LF	rectal swab	354	2	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451375	21836_4#158	EI4824_LF1	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	R
ERS1451406	21836_4#189	EI4747_LF2	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	R
ERS1451391	21836_4#174	EI4863_LF2	rectal swab	131	12	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	R
ERS1451327	21836_4#111	EI4766_LF1	rectal swab	998	10	ESBL	S	R	R	S	R	S	R	S	S	S	R	R	R

ERS1451319	21836_4#103	EI4749_LF	rectal swab	93	14	no	S	S	S	S	S	S	S	S	S	S	S	I	S
ERS1451335	21836_4#119	EI4787_LF	rectal swab	4948~	14	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451390	21836_4#173	EI4737_LF1	rectal swab	4948~	14	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451343	21631_1#127	EI4789_NLF	rectal swab	1193	13	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451398	21836_4#181	EI4741_NLF	rectal swab	4691	14	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451383	21836_4#166	EI4832_LF	rectal swab	38	3	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451321	21836_4#105	EI4946_LF	rectal swab	38	3	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451399	21836_4#182	EI4865_NLF	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R
ERS1451351	21836_4#134	EI4812_LF	rectal swab	5171	5	no	S	S	S	S	S	S	S	S	S	S	S	S	S
ERS1451326	21631_1#110	EI4917_LF	rectal swab	205~	11	no	S	S	S	S	S	S	S	S	S	S	R	I	S
ERS1451393	21836_4#176	EI4948_LF1	rectal swab	10	14	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1451401	21836_4#184	EI4948_LF2	rectal swab	10	14	no	S	S	S	R	S	S	S	S	S	R	R	I	S
ERS1451331	21836_4#115	EI4981_NLF	rectal swab	648	9	ESBL	S	R	R	R	R	S	R	S	S	R	S	R	R
ERS1451334	21836_4#118	EI4923_Bhae	rectal swab	141	10	ESBL	S	R	R	S	R	S	R	S	S	S	S	R	R
ERS1451346	21836_4#130	EI4973_LF2	rectal swab	167	14	ESBL	S	R	R	R	R	S	R	S	S	R	R	R	R

Supplementary table S3. Reference genomes included in this study

No	Reference	Genbank accession
1	Ecoli_O26H11_11368_EHEC	AP010953.1
2	Ecoli_SECEC_SMS35	CP000970.1
3	Escherichia_coli_0127_H6_E2348_69_v1	FM180568.1
4	Escherichia_coli_042_v1	FN554766.1
5	Escherichia_coli_AA86_v1	AFET01000001.1
6	Escherichia_coli_CFT073_v1	AE014075.1
7	Escherichia_coli_D5_v1	CP010145.1
8	Escherichia_coli_ETEC_H10407_v2	FN649414.1
9	Escherichia_coli_IAI1	GCF_000026265_1
10	Escherichia_coli_IAI39	GCF_000026345_1
11	Escherichia_coli_IHE3034_v1	NC_017628.1
12	Escherichia_coli_IMT2125_v1	NZ_HE964769.1
13	Escherichia_coli_JJ1886_v1	CP006784.1
14	Escherichia_coli_NA114_v2	GCF_000214765.2
15	Escherichia_coli_NCTC13351_v1	SAMEA2709031
16	Escherichia_coli_SE11	NC_011419.1
17	Escherichia_coli_SE15_v1	AP009378.1
18	Escherichia_coli_ST131_strain_EC958_v1	HG941718.1
19	Escherichia_coli_UMN026_v2	CU928149.2
20	Escherichia_coli_chi7122_v1	NZ_HE962388.1
21	Escherichia_coli_str_K-12_substr_MG1655_v2	GCF_000005845.2
22	Escherichia_coli_str_K12_substr_W3110_v1	GCF_000010245.2
23	Escherichia_fergusonii_ATCC_35469	GCF_000026225.1
24	Escherichia_albertii_EC0_170	GCF_001549955.1
25	Escherichia_albertii_KF1	GCF_000512125.1

Supplementary table S4. Metadata for paired and longitudinal samples

StudyNo	PatientNo	Sample name	Lane	MLST	Link to blood sample	date_since_ICUadmit	type	Outoome	Antibiotics	Antibiotics_other
03-0048	P1	141214_16448	21528_7#106	1722	141214_16448	0	blood	Discharged home to die	Ceftriaxone	Ceftriaxone--12/12/14
03-0048	P1	EI191_LF	21836_4#24	10	141214_16448	3	longitudinal	Discharged home to die	Ceftriaxone	Ceftriaxone--12/12/14
03-0314	P2	260515-15968	21836_4#57	648	260515-15968	0	blood	Discharged home to die	Ceftriaxone	Other-METRONIDAZOLE-
03-0314	P2	EI1749_NLF	21836_4#36	130	260515-15968	2	carriage	Discharged home to die	Ceftriaxone	Other-METRONIDAZOLE-
03-0314	P2	EI1749_LF2	21836_4#44	648	260515-15968	2	carriage	Discharged home to die	Ceftriaxone	Other-METRONIDAZOLE-
03-0314	P2	EI1787_NLF	21836_4#31	130	260515-15968	4	longitudinal	Discharged home to die	Ceftriaxone	Other-METRONIDAZOLE-
03-0314	P2	EI1787_LF2	21836_4#39	648	260515-15968	4	longitudinal	Discharged home to die	Ceftriaxone	Other-METRONIDAZOLE-
03-0314	P2	EI1830_LF2	21836_4#47	93	260515-15968	8	longitudinal	Discharged home to die	Ceftriaxone	Other-METRONIDAZOLE-
03-0314	P2	EI1888_LF2	21836_4#23	93	260515-15968	11	longitudinal	Discharged home to die	Ceftriaxone	Other-METRONIDAZOLE-
03-0283	P3	EI1537_BhaeNLF	21836_4#14	6698	010615-16405	1	carriage	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	EI1537_LF	21836_4#22	131	010615-16405	1	carriage	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	EI1564_Bhae	21836_4#87	131	010615-16405	2	longitudinal	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	EI1564_LF	21836_4#16	43	010615-16405	2	longitudinal	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	EI1605_NLF	21836_4#55	6698	010615-16405	6	longitudinal	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	EI1605_BhaeLF	21836_4#63	131	010615-16405	6	longitudinal	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	EI1705_LF2	21836_4#95	131	010615-16405	13	longitudinal	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	EI1761_NLF	21836_4#71	6698	010615-16405	16	longitudinal	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	EI1761_LF2	21836_4#8	131	010615-16405	16	longitudinal	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	EI1864_LF	21836_4#79	131	010615-16405	20	longitudinal	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0283	P3	010615-16405	21836_4#82	131	010615-16405	20	blood	Discharged home, stable	Metronidazole	Other-METRONIDAZOLE-
03-0559	P4	050915-16464	21836_4#25	73	050915-16464	0	blood	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0559	P4	EI3399_NLF	21836_4#67	131	050915-16464	0	carriage	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0559	P4	EI3399_LF	21836_4#75	155	050915-16464	0	carriage	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0559	P4	EI3449_NLF	21836_4#56	38	050915-16464	3	longitudinal	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0559	P4	EI3449_LF2	21836_4#72	1722	050915-16464	3	longitudinal	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0559	P4	EI3449_LF1	21836_4#64	88	050915-16464	3	longitudinal	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0710	P5	061115-28103	21836_4#18	1193	061115-28103	0	blood	Transfer to another hospital	Amikacin	Ceftriaxone--
03-0710	P5	EI4341_BhaeNLF	21836_4#13	1193	061115-28103	1	carriage	Transfer to another hospital	Amikacin	Ceftriaxone--
03-0710	P5	EI4381_LF	21836_4#88	69	061115-28103	3	longitudinal	Transfer to another hospital	Amikacin	Ceftriaxone--
03-0710	P5	EI4381_NLF	21631_1#96	648	061115-28103	3	longitudinal	Transfer to another hospital	Amikacin	Ceftriaxone--
03-0209	P6	250315-63665	21836_4#2	69	250315-63665	0	blood	Discharged home, stable	Imipenem	Imipenem--
03-0209	P6	EI950_LF	21836_4#32	354	250315-63665	1	longitudinal	Discharged home, stable	Imipenem	Imipenem--
03-0514	P7	210815-15032	21836_4#10	448	210815-15032	0	blood	Discharged home to die	Ertapenem	Other-NORFLOXACIN-
03-0514	P7	EI3127_LF	21836_4#5	448	210815-15032	8	carriage	Discharged home to die	Ertapenem	Other-NORFLOXACIN-
03-0514	P7	EI3185_Ecoli	21631_1#80	448	210815-15032	11	longitudinal	Discharged home to die	Ertapenem	Other-NORFLOXACIN-
03-0012	P8	131114_64463	21528_7#109	1193	131114-64463	0	blood	Discharged home, stable	Ceftriaxone	
03-0012	P8	EI32_NLF	21836_4#70	1193	131114-64463	1	carriage	Discharged home, stable	Ceftriaxone	
03-0014	P9	141114_14352	21528_7#111	224	141114-14352	0	blood	Discharged home, stable	Amikacin	
03-0014	P9	EI36_LF	21836_4#86	167	141114-14352	0	carriage	Discharged home, stable	Amikacin	
03-0029	P10	201114-14705	21836_4#19	5150	201114-14705	0	blood	Discharged home, stable	Ceftriaxone	

03-0029	P10	EI103_NLF	21836_4#94	5150	201114-14705	8	carriage	Discharged home, stable	Ceftriaxone	
03-0039	P11	101214_16015	21528_7#114	38	101214-16015	0	blood	Discharged home to die	Imipenem	
03-0039	P11	EI138_NLF	21836_4#7	4381	101214-16015	1	carriage	Discharged home to die	Imipenem	
03-0039	P11	EI138_LF	21836_4#15	515	101214-16015	1	carriage	Discharged home to die	Imipenem	
03-0081	P12	100115-14032	21836_4#1	648	100115-14032	0	blood	Discharged home to die	Ceftriaxone	IMIPENEM--10/01/15
03-0081	P12	EI298_NLF	21836_4#27	648	100115-14032	0	carriage	Discharged home to die	Ceftriaxone	IMIPENEM--10/01/15
03-0081	P12	EI298_LF	21836_4#35	38	100115-14032	0	carriage	Discharged home to die	Ceftriaxone	IMIPENEM--10/01/15
03-0131	P13	310115-63575	21836_4#26	3052	310115-63575	0	blood	Discharged home to die	Ceftriaxone	Imipenem--31/01/15
03-0131	P13	EI534_LF	21836_4#21	3052	310115-63575	1	carriage	Discharged home to die	Ceftriaxone	Imipenem--31/01/15
03-0135	P14	010215-15680	21836_4#90	131	010215-15680	0	blood	Transfer to another hospital	Amikacin	Ceftriaxone--
03-0135	P14	EI544_NLF	21836_4#30	131	010215-15680	0	carriage	Transfer to another hospital	Amikacin	Ceftriaxone--
03-0135	P14	EI544_LF	21836_4#38	31	010215-15680	0	carriage	Transfer to another hospital	Amikacin	Ceftriaxone--
03-0151	P15	080215-16203	21836_4#3	410	080215-16203	0	blood	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0151	P15	EI644_LF	21836_4#46	5208	080215-16203	0	carriage	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0173	P16	100315-17879	21836_4#73	354	100315-17879	0	blood	Discharged home, stable	Ceftriaxone	Ceftriaxone--10/03/15
03-0173	P16	EI772_LF2	21836_4#68	354	100315-17879	0	carriage	Discharged home, stable	Ceftriaxone	Ceftriaxone--10/03/15
03-0173	P16	EI746_LF1	21836_4#76	617	100315-17879	0	carriage	Discharged home, stable	Ceftriaxone	Ceftriaxone--10/03/15
03-0185	P17	160315-63225	21836_4#9	394	160315-63225	0	blood	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0185	P17	EI815_LF	21836_4#43	394	160315-63225	0	carriage	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0301	P18	210515-15556	21836_4#66	1193	210515-15556	0	blood	Discharged home to die	Imipenem	Other-METRONIDAZOLE-
03-0301	P18	EI1676_NLF	21836_4#93	1193	210515-15556	0	carriage	Discharged home to die	Imipenem	Other-METRONIDAZOLE-
03-0322	P19	300515-16242	21836_4#65	131	300515-16242	0	blood	Discharged home, stable	Imipenem	Imipenem--
03-0322	P19	EI1808_NLF	21836_4#52	1193	300515-16242	0	carriage	Discharged home, stable	Imipenem	Imipenem--
03-0322	P19	EI1808_LF2	21836_4#60	131	300515-16242	0	carriage	Discharged home, stable	Imipenem	Imipenem--
03-0437	P20	210715-16270	21836_4#49	38	210715-16270	0	blood	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0437	P20	EI2619_LF4	21836_4#20	23	210715-16270	0	carriage	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0437	P20	EI2619_LF2	21836_4#28	38	210715-16270	0	carriage	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0471	P21	010815-34451	21836_4#17	NA	010815-34451	0	blood	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0471	P21	EI2793_NLF	21836_4#51	354	010815-34451	1	carriage	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0471	P21	EI2793_LF2	21836_4#59	167	010815-34451	1	carriage	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0504	P22	160815-14563	21836_4#33	126	160815-14563	0	blood	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0504	P22	EI3097_LF4	21836_4#83	4040	160815-14563	1	carriage	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0504	P22	EI3097_LF1	21836_4#91	58	160815-14563	1	carriage	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0504	P22	EI3097_LF2	21836_4#4	4040	160815-14563	1	carriage	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0525	P23	240815-15273	21836_4#42	69	240815-15273	0	blood	Transfer to another hospital	Ceftriaxone	Other-METRONIDAZOLE-
03-0525	P23	EI3217_LF2	21836_4#53	58	240815-15273	0	carriage	Transfer to another hospital	Ceftriaxone	Other-METRONIDAZOLE-
03-0525	P23	EI3217_LF3	21836_4#61	58	240815-15273	0	carriage	Transfer to another hospital	Ceftriaxone	Other-METRONIDAZOLE-
03-0625	P24	290915-14739	21836_4#81	69	290915-14739	0	blood	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0625	P24	EI3807_LF	21836_4#84	4306	290915-14739	0	carriage	Transfer to another hospital	Ceftriaxone	Ceftriaxone--
03-0634	P25	061015-15322	21836_4#34	131	061015-15322	0	blood	Transfer to another hospital	Imipenem	Imipenem--
03-0634	P25	EI3892_LF	21836_4#29	648	061015-15322	0	carriage	Transfer to another hospital	Imipenem	Imipenem--
03-0634	P25	EI3892_Bhae	21836_4#37	648	061015-15322	0	carriage	Transfer to another hospital	Imipenem	Imipenem--
03-0634	P25	EI3892_LF1	21836_4#45	44	061015-15322	0	carriage	Transfer to another hospital	Imipenem	Imipenem--

03-0699	P26	311015-17508	21836_4#11	1193	311015-17508	0	blood	Discharged home to die	Imipenem	Imipenem--
03-0699	P26	EI4263_LF	21836_4#54	34	311015-17508	0	carriage	Discharged home to die	Imipenem	Imipenem--
03-0699	P26	EI4263_NLF	21836_4#62	1193	311015-17508	0	carriage	Discharged home to die	Imipenem	Imipenem--
03-0714	P27	071115-14091	21836_4#89	602	071115-14091	0	blood	Discharged home, stable	Imipenem	Imipenem--
03-0714	P27	EI4355_LF_m	21836_4#92	602	071115-14091	0	carriage	Discharged home, stable	Imipenem	Imipenem--
03-0770	P28	301115-16266	21836_4#58	1193	301115-16266	0	blood	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0770	P28	EI4625_NLF	21836_4#77	69	301115-16266	0	carriage	Discharged home, stable	Ceftriaxone	Ceftriaxone--
03-0770	P28	EI4625_LF	21836_4#85	5415	301115-16266	0	carriage	Discharged home, stable	Ceftriaxone	Ceftriaxone--

Supplementary table S5. Pairwise SNPs distance between isolates taken from the same patients

Taxon1_name	Taxon2_name	SNPs	ID	taxon1_st	taxon2_st	group	sameST
141214_16448	EI191_LF	67679	71.11229	1722	10	intra-patient	no
131114_64463	EI32_NLF	3	99.99873	1193	1193	intra-patient	yes
141114_14352	EI36_LF	36597	84.27045	224	167	intra-patient	no
101214_16015	EI138_LF	64020	72.06331	38	515	intra-patient	no
101214_16015	EI138_NLF	64582	72.12544	38	4381	intra-patient	no
EI3185_Ecoli	210815-15032	0	100	448	448	intra-patient	yes
EI3185_Ecoli	EI3127_LF	0	100	448	448	intra-patient	yes
EI4381_NLF	EI4341_BhaeNLF	65958	71.98427	648	1193	intra-patient	no
EI4381_NLF	061115-28103	65928	71.99154	648	1193	intra-patient	no
EI4381_NLF	EI4381_LF	58231	75.32428	648	69	intra-patient	no
100115-14032	EI298_NLF	2	99.99915	648	648	intra-patient	yes
100115-14032	EI298_LF	57918	75.32769	648	38	intra-patient	no
210815-15032	EI3127_LF	0	100	448	448	intra-patient	yes
311015-17508	EI4263_LF	70080	69.80117	1193	34	intra-patient	no
311015-17508	EI4263_NLF	1	99.99958	1193	1193	intra-patient	yes
EI4341_BhaeNLF	061115-28103	0	100	1193	1193	intra-patient	yes
EI4341_BhaeNLF	EI4381_LF	73711	68.76441	1193	69	intra-patient	no
EI1537_BhaeNLF	EI1564_LF	76212	67.29127	6698	43	intra-patient	no
EI1537_BhaeNLF	EI1537_LF	32748	85.95176	6698	131	intra-patient	no
EI1537_BhaeNLF	EI1605_NLF	0	100	6698	6698	intra-patient	yes
EI1537_BhaeNLF	EI1605_BhaeLF	32749	85.95139	6698	131	intra-patient	no
EI1537_BhaeNLF	EI1761_NLF	2	99.99914	6698	6698	intra-patient	yes
EI1537_BhaeNLF	EI1864_LF	32749	85.95133	6698	131	intra-patient	no
EI1537_BhaeNLF	EI1761_LF2	32748	85.95182	6698	131	intra-patient	no
EI1537_BhaeNLF	010615-16405	32749	85.95133	6698	131	intra-patient	no
EI1537_BhaeNLF	EI1564_Bhae	32748	85.95176	6698	131	intra-patient	no
EI1537_BhaeNLF	EI1705_LF2	32748	85.95182	6698	131	intra-patient	no
EI138_LF	EI138_NLF	40589	82.40349	515	4381	intra-patient	no
EI1564_LF	EI1537_LF	75022	68.00358	43	131	intra-patient	no
EI1564_LF	EI1605_NLF	76131	67.312	43	6698	intra-patient	no
EI1564_LF	EI1605_BhaeLF	75022	68.00358	43	131	intra-patient	no
EI1564_LF	EI1761_NLF	76111	67.30431	43	6698	intra-patient	no
EI1564_LF	EI1864_LF	75021	68.00401	43	131	intra-patient	no
EI1564_LF	EI1761_LF2	75023	68.00316	43	131	intra-patient	no
EI1564_LF	010615-16405	75022	68.00358	43	131	intra-patient	no
EI1564_LF	EI1564_Bhae	75023	68.00316	43	131	intra-patient	no
EI1564_LF	EI1705_LF2	75022	68.00358	43	131	intra-patient	no
010815-34451	EI2793_NLF	65168	71.96436	NA	354	intra-patient	no
010815-34451	EI2793_LF2	74971	67.84238	NA	167	intra-patient	no
061115-28103	EI4381_LF	73701	68.76256	1193	69	intra-patient	no
201114-14705	EI103_NLF	4	99.99831	5150	5150	intra-patient	yes
250315-63665	EI950_LF	58833	74.89085	69	354	intra-patient	no
EI2619_LF4	EI2619_LF2	61995	73.5816	23	38	intra-patient	no
EI2619_LF4	210715-16270	61995	73.5816	23	38	intra-patient	no
EI534_LF	310115-63575	1	99.99958	3052	3052	intra-patient	yes
EI1537_LF	EI1605_NLF	32670	85.9792	131	6698	intra-patient	no
EI1537_LF	EI1605_BhaeLF	3	99.99873	131	131	intra-patient	yes
EI1537_LF	EI1761_NLF	32648	85.98167	131	6698	intra-patient	no
EI1537_LF	EI1864_LF	1	99.99958	131	131	intra-patient	yes

EI1537_LF	EI1761_LF2	2	99.99915	131	131	intra-patient	yes
EI1537_LF	010615-16405	3	99.99873	131	131	intra-patient	yes
EI1537_LF	EI1564_Bhae	2	99.99915	131	131	intra-patient	yes
EI1537_LF	EI1705_LF2	0	100	131	131	intra-patient	yes
EI1888_LF2	EI1787_NLF	64670	72.23713	93	130	intra-patient	no
EI1888_LF2	EI1749_NLF	64670	72.23713	93	130	intra-patient	no
EI1888_LF2	EI1787_LF2	66560	71.61584	93	648	intra-patient	no
EI1888_LF2	EI1749_LF2	66560	71.6179	93	648	intra-patient	no
EI1888_LF2	EI1830_LF2	1	99.99957	93	93	intra-patient	yes
EI1888_LF2	260515-15968	66561	71.61747	93	648	intra-patient	no
050915-16464	EI3449_NLF	72878	68.98122	73	38	intra-patient	no
050915-16464	EI3449_LF1	75451	67.81375	73	88	intra-patient	no
050915-16464	EI3399_NLF	33547	85.67206	73	131	intra-patient	no
050915-16464	EI3449_LF2	65796	71.99348	73	1722	intra-patient	no
050915-16464	EI3399_LF	75587	67.74514	73	155	intra-patient	no
EI298_NLF	EI298_LF	57896	75.3316	648	38	intra-patient	no
EI2619_LF2	210715-16270	0	100	38	38	intra-patient	yes
EI3892_LF	061015-15322	65654	72.03595	648	131	intra-patient	no
EI3892_LF	EI3892_Bhae	1	99.99958	648	648	intra-patient	yes
EI3892_LF	EI3892_LF1	66451	71.65809	648	44	intra-patient	no
080215-16203	EI644_LF	34393	85.20489	410	5208	intra-patient	no
EI544_NLF	EI544_LF	74625	67.88748	131	31	intra-patient	no
EI544_NLF	010215-15680	2	99.99915	131	131	intra-patient	yes
EI1787_NLF	EI1749_NLF	0	100	130	130	intra-patient	yes
EI1787_NLF	EI1787_LF2	58363	75.08814	130	648	intra-patient	no
EI1787_NLF	EI1749_LF2	58362	75.09038	130	648	intra-patient	no
EI1787_NLF	EI1830_LF2	64656	72.23694	130	93	intra-patient	no
EI1787_NLF	260515-15968	58363	75.08995	130	648	intra-patient	no
160815-14563	EI3097_LF2	74848	68.20431	126	4040	intra-patient	no
160815-14563	EI3097_LF4	74848	68.20431	126	4040	intra-patient	no
160815-14563	EI3097_LF1	76218	67.45949	126	58	intra-patient	no
061015-15322	EI3892_Bhae	65643	72.03444	131	648	intra-patient	no
061015-15322	EI3892_LF1	74745	68.01231	131	44	intra-patient	no
EI1749_NLF	EI1787_LF2	58363	75.08814	130	648	intra-patient	no
EI1749_NLF	EI1749_LF2	58362	75.09038	130	648	intra-patient	no
EI1749_NLF	EI1830_LF2	64656	72.23694	130	93	intra-patient	no
EI1749_NLF	260515-15968	58363	75.08995	130	648	intra-patient	no
EI3892_Bhae	EI3892_LF1	66433	71.65949	648	44	intra-patient	no
EI544_LF	010215-15680	74585	67.88216	31	131	intra-patient	no
EI1787_LF2	EI1749_LF2	2	99.99915	648	648	intra-patient	yes
EI1787_LF2	EI1830_LF2	66570	71.60589	648	93	intra-patient	no
EI1787_LF2	260515-15968	3	99.99873	648	648	intra-patient	yes
EI3097_LF2	EI3097_LF4	0	100	4040	4040	intra-patient	yes
EI3097_LF2	EI3097_LF1	64433	72.55437	4040	58	intra-patient	no
240815-15273	EI3217_LF2	63636	72.93928	69	58	intra-patient	no
240815-15273	EI3217_LF3	63654	72.93714	69	58	intra-patient	no
EI815_LF	160315-63225	0	100	394	394	intra-patient	yes
EI1749_LF2	EI1830_LF2	66570	71.60794	648	93	intra-patient	no
EI1749_LF2	260515-15968	1	99.99958	648	648	intra-patient	yes
EI1830_LF2	260515-15968	66571	71.60752	93	648	intra-patient	no
EI2793_NLF	EI2793_LF2	68040	70.6385	354	167	intra-patient	no
EI1808_NLF	EI1808_LF2	33725	85.63035	1193	131	intra-patient	no

EI1808_NLF	300515-16242	33778	85.62491	1193	131	intra-patient	no
EI3217_LF2	EI3217_LF3	0	100	58	58	intra-patient	yes
EI4263_LF	EI4263_NLF	70083	69.80313	34	1193	intra-patient	no
EI1605_NLF	EI1605_BhaeLF	32671	85.97883	6698	131	intra-patient	no
EI1605_NLF	EI1761_NLF	2	99.99914	6698	6698	intra-patient	yes
EI1605_NLF	EI1864_LF	32671	85.97877	6698	131	intra-patient	no
EI1605_NLF	EI1761_LF2	32670	85.97926	6698	131	intra-patient	no
EI1605_NLF	010615-16405	32671	85.97877	6698	131	intra-patient	no
EI1605_NLF	EI1564_Bhae	32670	85.9792	6698	131	intra-patient	no
EI1605_NLF	EI1705_LF2	32670	85.97926	6698	131	intra-patient	no
EI3449_NLF	EI3449_LF1	65163	72.28746	38	88	intra-patient	no
EI3449_NLF	EI3399_NLF	74236	68.33489	38	131	intra-patient	no
EI3449_NLF	EI3449_LF2	58710	75.06678	38	1722	intra-patient	no
EI3449_NLF	EI3399_LF	64892	72.37768	38	155	intra-patient	no
301115-16266	EI4625_NLF	71486	69.65699	1193	69	intra-patient	no
301115-16266	EI4625_LF	72945	68.99269	1193	5415	intra-patient	no
EI1808_LF2	300515-16242	40	99.98298	131	131	intra-patient	yes
EI1605_BhaeLF	EI1761_NLF	32649	85.9813	131	6698	intra-patient	no
EI1605_BhaeLF	EI1864_LF	2	99.99915	131	131	intra-patient	yes
EI1605_BhaeLF	EI1761_LF2	1	99.99958	131	131	intra-patient	yes
EI1605_BhaeLF	010615-16405	0	100	131	131	intra-patient	yes
EI1605_BhaeLF	EI1564_Bhae	1	99.99958	131	131	intra-patient	yes
EI1605_BhaeLF	EI1705_LF2	3	99.99873	131	131	intra-patient	yes
EI3449_LF1	EI3399_NLF	74810	68.04685	88	131	intra-patient	no
EI3449_LF1	EI3449_LF2	67711	71.1971	88	1722	intra-patient	no
EI3449_LF1	EI3399_LF	24842	89.41962	88	155	intra-patient	no
210515-15556	EI1676_NLF	40	99.98306	1193	1193	intra-patient	yes
EI3399_NLF	EI3449_LF2	65978	71.87051	131	1722	intra-patient	no
EI3399_NLF	EI3399_LF	75429	67.77048	131	155	intra-patient	no
EI772_LF2	100315-17879	4	99.9983	354	354	intra-patient	yes
EI772_LF2	EI746_LF1	65753	71.83543	354	617	intra-patient	no
EI1761_NLF	EI1864_LF	32649	85.98124	6698	131	intra-patient	no
EI1761_NLF	EI1761_LF2	32648	85.98173	6698	131	intra-patient	no
EI1761_NLF	010615-16405	32649	85.98124	6698	131	intra-patient	no
EI1761_NLF	EI1564_Bhae	32648	85.98167	6698	131	intra-patient	no
EI1761_NLF	EI1705_LF2	32648	85.98173	6698	131	intra-patient	no
EI3449_LF2	EI3399_LF	67636	71.22044	1722	155	intra-patient	no
100315-17879	EI746_LF1	65753	71.83604	354	617	intra-patient	no
EI4625_NLF	EI4625_LF	62371	73.47687	69	5415	intra-patient	no
EI1864_LF	EI1761_LF2	3	99.99873	131	131	intra-patient	yes
EI1864_LF	010615-16405	2	99.99915	131	131	intra-patient	yes
EI1864_LF	EI1564_Bhae	3	99.99873	131	131	intra-patient	yes
EI1864_LF	EI1705_LF2	1	99.99958	131	131	intra-patient	yes
EI1761_LF2	010615-16405	1	99.99958	131	131	intra-patient	yes
EI1761_LF2	EI1564_Bhae	0	100	131	131	intra-patient	yes
EI1761_LF2	EI1705_LF2	2	99.99915	131	131	intra-patient	yes
290915-14739	EI3807_LF	43452	81.55429	69	4306	intra-patient	no
010615-16405	EI1564_Bhae	1	99.99958	131	131	intra-patient	yes
010615-16405	EI1705_LF2	3	99.99873	131	131	intra-patient	yes
EI3097_LF4	EI3097_LF1	64431	72.55417	4040	58	intra-patient	no
EI1564_Bhae	EI1705_LF2	2	99.99915	131	131	intra-patient	yes
071115-14091	EI4355_LF_m	0	100	602	602	intra-patient	yes