Understanding within patient diversity of Uropathogenic *Escherichia coli*

Ruqaiah Ismael Bedaiwi

A thesis submitted in partial fulfilment of the requirements of

Nottingham Trent University for the degree of

Doctor of Philosophy

In

Diagnostic Clinical Microbiology & Infection Control

School of Science and Technology

January 2018

Abstract

Urinary tract infections (UTIs) are one of the most common human infections worldwide. UTI is caused by a wide range of Gram-positive and Gram-negative pathogens with Uropathogenic *E. coli* (UPEC) the most common causative pathogen. In clinical microbiology, diagnosis, analysis and treatment are based on single colony selection of a homogenous bacterial population. However, recent work on infections such as cystic fibrosis has highlighted the presence of multiple phenotypic and genotypic variants within a single infected patient.

A study comparing bacteria isolated from urine and blood samples from patients with urosepsis showed the presence of multiple sequence types of *E. coli* within a single patient. Therefore, here we present work investigating the level of within-patient diversity of UPEC at a phenotypic and genotypic level. Forty-two urine samples were collected and antibiotic sensitivity testing performed on each well-isolated colony. Samples are classified based on their sensitivity profiles into three patterns: identical resistance profile, low diversity resistance profiles and highly diverse resistance profiles. Nine urine samples were categorized as having highly diverse resistance profiles. Phenotypic assays of bacteria from the highly diverse group show variation in motility, biofilm formation and association and invasion assays. To determine the phenotypic baseline diversity level, the highly diverse resistance profile samples were compared with samples that were shown to have a homogenous population, and eight randomly selected samples with low diversity patterns. Bacteria from both patterns show no phenotypic variation. We further analysed the levels of genotypic diversity between sample isolates. We compared bacteria from highly diverse resistance profiles using whole-genome sequence data

in order to correlate the phenotypic diversity with genetic changes. Together our data is the most high-resolution snapshot to date-of the levels of extant diversity of UPEC within patients with UTI.

Declaration

I hereby declare that the presented work herein is the result of my original research work, except where references have been made to acknowledge the literature. This work is an intellectual property of the author. You may copy up to 5% of this work for private study, or personal, non-commercial research. Any re-use of the information contained in this document should be fully referenced, quoting the author, title, university, degree level and pagination. Queries or request of any other use, or if a more substantial copy is required, should be directed in the first instance to the owner(s) of the intellectual property rights. Experiments were performed in the pathogen research group at Nottingham Trent University.

Ruqaiah Ismael Bedaiwi

Acknowledgement

First and foremost, praise **Allah**, the Almighty, on whom ultimately, we depend for sustenance and guidance. I would like to thank Allah, the Almighty, for giving me the strength and patience to carry on and fulfil this research and for blessing me with many great people who have been my greatest support in both my personal and professional life.

I would like to seize this opportunity to express my deepest regards and gratitude to my supervisor **Dr. Alan McNally** for his dedication and sincere effort throughout the whole stages of this research. His guidance helped me in all the time of research and writing of this thesis. Besides my supervisor, I would like to thank **Dr. Benjamin Dickins** for his endless cooperation and the rest of my thesis committee: **Prof. Stephen Forsythe**, **Dr. Micheal Loughlin**, **Dr. Jody Winter**, **Dr. Gareth McVicker** and **Dr. Georgina manning** for their insightful comments and encouragement.

My sincerest and deepest thanks and appreciation to the great and almighty **Allah's** gift to me, my beloved family. I would like to express my gratitude for their endless cooperation, their unlimited care and love, and their engagement during the best and difficult moments in my life. My greatest thanks and prayers are for my father, **(Ismael)**, who has inspired me throughout my life, and I would like to ask **Allah**, the almighty, to bless his soul. Further, I have deep gratitude for my mother, **(Gauaher)**, and my younger sister, **(Bshayer)**, for their support, encouragement, and prayers. I would like to

especially thank my dear brother, **(Ali)**, who has spent his life to join me on this journey, offering me a part of his time, supporting, and accompanying me during my sponsorship. Finally, I wish to express my deepest gratitude and appreciation to all my relatives for their prayers, wishes, and support throughout my studies.

I also owe my thanks to all members of **pathogen research group** and my colleges, **Nottingham Trent University**, wishing them further success. The administrative and financial support of the **Saudi Arabia Cultural Bureau**, the **Higher Ministry of Education** and **Tabuk University** are gratefully acknowledged.

This research is dedicated to my beloved parents:

(Ismael and Gauaher)

And my dear siblings:

(Ali and Bshayer)

Poster presentation

Ruqaiah Ismael Bedaiwi, Mathew Diggle and Alan McNally, (2016). Understanding within patient variation of uropathgenic *Escherichia coli. School of Science and Technology Research*. Nottingham Trent University**.**

Ruqaiah Ismael Bedaiwi and Alan McNally, (2016). Understanding within patient variation of uropathgenic *Escherichia coli. Microbe*, Sheffield.

Ruqaiah Ismael Bedaiwi and Alan McNally, (2017). Understanding within patient variation of uropathgenic *Escherichia coli. Microbiology Society*, Edinburgh.

Table of Contents

List of tables

List of figures

Abbreviation

ATP/GTP	Adenosine triphosphate/ Guanosine triphosphate
Bp	Base pairs
***bla*$_{KPC}$**	*Klebsiella pneumoniae* carbapenemase β-lactamase gene
***bla*$_{NDM-1}$**	New Delhi metallo- β-lactamase gene
***bla*$_{VIM}$**	Verona integron-encoded metallo- β-lactamase gene
CFU/ml	Colony forming unit/ milliliter
CO_2	Carbon dioxide
CTX-M	Active on CefoTaXime, first isolated in Munich
DNA	Deoxyribo nucleic acid
g/L	Gram/ litter
HS	High sensitivity
L	Littre
Mg	Milligram
Ml	Milliliter
MLVA	Multiple-Locus Variable number tandem repeat Analysis
Mm	Millimeter
MRSA	Methicillin resistant *Staphylococcus aureus*
MSSA	Methicillin sensitive *Staphylococcus aureus*
ng/µl	Nanogram/ microliter
Nm	Nanometers
PCR	Polymerase chain reaction
PFGE	Pulse field gel electrophoresis

PH	Potential of Hydrogen
Rpm	Rotation per minute
SNPs	Single nucleotide polymorphism
V/V	Volume/ Volume
µg/ml	Microgram/ milliliters
µl	Microliter

Chapter One

Introduction

1 Introduction

1.1 Urinary tract infections

Urinary tract infections (UTIs) are one of the most common infections of humans and are defined as invasion of urinary tract epithelium by pathogenic bacteria. Annually, 150 million people are diagnosed with UTIs worldwide (Foxman, 2003; Hooton, 2012; Lüthje and Brauner, 2016; Flores-Mireles et al., 2015). UTIs can generally be classified into upper UTIs (pyelonephritis) and lower UTIs (cystitis). UTIs cover a wide range of disease presentations varying in severity from asymptomatic bacteriuria to urosepsis and death. Those diseases are categorised based on the pathogen isolation site (Foxman, 2003; Foxman, 2010; Brumbaugh and Mobley, 2012; Flores-Mireles et al., 2015; Kumar et al., 2015). If bacteria are isolated from urine without clinical symptoms this is known as asymptomatic bacteriuria. Cystitis refers to isolation of bacteria from the bladder, while isolation of bacteria from the kidneys is known as pyelonephritis. In addition, urosepsis occurs following the spread of the uropathogen to the bloodstream (Foxman, 2010; Brumbaugh and Mobley, 2012; Flores-Mireles et al., 2015; Kumar et al., 2015).

UTIs also can be classified as either uncomplicated or complicated. Uncomplicated UTIs are defined when the infection occurs in healthy individuals with no structural abnormalities. However, UTIs in individuals with urological abnormalities, inserted instrumentation (catheter) or who are immunocompromised, are classified as complicated infections (Foxman, 2003; Flores-Mireles et al., 2015; Kumar et al., 2015).

Urinary tract infections occur more frequently in women. About 50 % - 80 % of females suffer from UTIs and 20 % - 50 % of those women may experience a

recurrent UTI (Foxman, 1990; Agarwal et al., 2012; Lüthje and Brauner, 2014). Further, infants, pregnant women, elderly people and diabetic patients are categorised as high-risk groups with respect to UTIs (Foxman, 2003; Hannan et al., 2012).

1.1.1 Pathogenesis

UTIs are caused by a wide range of Gram-negative, Gram-positive and fungal pathogens (Hacker, 2002). The most frequent causative agent of uncomplicated UTIs is the group known as Uropathogenic *Escherichia coli*, which account for 70 % - 95 % of community-acquired infections. Uropathogenic *Escherichia coli* are predominant in 50 % of nosocomial infections (Hacker, 2002; Kucheria et al., 2005). However, complicated UTIs can be caused by other Gram-negative pathogens such as *Proteus* species, multi-resistant *Klebsiella pneumoniae* and *Pseudomonas aeruginosa*. These pathogens play a role in infections in catheterised and immunocompromised patients. Gram-positive bacteria such as *Staphylococcus aureus* and *Staphylococcus epidermidis* may also have a role in causing nosocomial UTIs, especially in patients who have undergone organ transplantation or chemotherapy (Hacker, 2002).

Pathogens that cause UTIs are thought to originate from the human intestinal tract. There are also some indications that UTIs may result from the consumption of contaminated food, and other evidence suggests that uropathogens could be transmitted sexually (Pitout, 2012a). Once the uropathogen has migrated to the bladder, colonisation of the urinary tract may begin (Foxman, 2010; Hooton, 2012). Following migration and colonisation by uropathogens, adherence to urinary tract cells is initiated by the action of pili and flagella (Flores-Mireles et al., 2015). The adherence of the uropathogen can result in the invasion of the urinary tract

epithelium cell. Invasion of the urinary tract allows bacteria to survive within the host, and to produce toxins and proteases which release nutrients such as iron allowing their sequestration by siderophores (Flores-Mireles et al., 2015). Uropathogens then can multiply and evade host immunity, so allowing bacteria to ascend from the urethra to the bladder, ureters and kidneys and cause infection (Agarwal et al., 2012; Hooton, 2012; Flores-Mireles et al., 2015). This route of UTI spread is known as the ascending route of infection and it is a common sequence of infection (Agarwal et al., 2012). Efficient colonisation of the periurethral area with uropathogens is dependent on a variety of virulence factors which are required for motility, colonisation and evading the immune system (Foxman, 2010).

1.1.1.1 Catheter associated urinary tract infections

The most common infections in healthcare settings are catheter associated UTIs (Kumar et al., 2015). These infections are considered to be correlated with high morbidity and may contribute to the development of severe infections, such as those of the blood stream (Choong and Whitfield, 2000; Flores-Mireles et al., 2015). In catheterised patients, biofilm producing strains have a significant role in blocking the catheter and this may result in haematuria and pain (Choong and Whitfield, 2000). UTIs are diagnosed in catheterised patients by the presence of a bacterial count >1000 CFU/ml in addition to the symptoms listed above. In the case of a positive urine culture, the catheter should be replaced and an antimicrobial treatment initiated based on the culture antimicrobial sensitivity result (Kumar et al., 2015).

1.1.1.2 Recurrent urinary tract infections

Recurrent UTIs are defined as the occurrence of two or more infections within six months, or more than three infections within a year (Lüthje and Brauner, 2016). Recurrent UTIs are commonly seen in women and it has been reported that 25 % of women have a second UTI within six months of the first infection (Ejrnæs, 2011). Anatomical structure plays a crucial role in bacterial ascension to the urinary tract. Women have short urethras and the close proximity to the periurethral area is the most probable route of infection (Minardi et al., 2011).

Reinfection with the same bacterial strain because of its persistence within the urinary tract is considered one of the reasons for recurrent UTIs (Lüthje and Brauner, 2016). Following adherence, some pathogens invade the epithelium cells to form intracellular bacterial communities. Cells that are invaded by uropathogens undergo apoptosis, are excreted into the urine, and the uropathogen is released from intracellular bacterial communities to invade deeper tissue (Lüthje and Brauner, 2016). Within infected cells, bacteria form dormant cells that do not multiply and are known as the quiescent intracellular reservoir. Once these dormant cells become activated a recurrent infection may result (Lüthje and Brauner, 2016).

1.1.2 Clinical manifestations

UTIs can comprise a wide spectrum of diseases varying in severity. Asymptomatic UTIs can be diagnosed by the presence of a high level of bacteriuria in the absence of symptoms. Symptomatic UTIs are generally classified based on the site of the occurrence of symptoms: lower UTIs and upper UTIs (Hannan et al., 2012). Classical UTI symptoms include pain, frequent urination and urgency. Patients who have these symptoms are diagnosed with cystitis in 96 % of cases (Bien et al., 2012).

Lower UTIs affecting the urethra and bladder may produce symptoms which include frequent urination, a burning sensation and dysuria. Further, lower abdominal discomfort with cloudy or bloody urine that may be foul smelling may occur as symptoms of lower UTIs. In contrast, infection of the upper urinary tract (ureters and kidneys) may associate with flank pain and fever. These symptoms are, in general, a result of inflammation of the urinary tract epithelium following uropathogen colonisation (Hannan et al., 2012).

1.1.3 Diagnosis

UTIs are diagnosed based on urinalysis and the presence of a significant number of pathogen cells in urine samples without or with symptoms. Urinalysis tests are commonly used in the laboratory to determine the nature of the infection. Nitrite and leukocyte esterase assays are dipstick tests that can help in UTI diagnosis. A positive nitrite result reflects the presence of a Gram-negative species. Leukocyte esterase is another marker of UTIs and turns to positive in the presence of an elevated number of leukocytes as a consequence of the presence of uropathogens. Due to the limitations of the dipstick assays and to confirm a positive result, a microscopic examination of urine sediment should be made to confirm the presence of bacteria in the urine. However, the accuracy of microscopic examination is related to number of epithelial cell present within the sample, and there should be < 15 cells per high power field (40 x). UTIs are diagnosed on the basis of positive culture plates with bacterial concentrations exceeding 10^3 – 10^5 CFU/ml (Foxman, 2010; Hilbert, 2011).

1.1.4 Treatment

UTIs are treated based on the patient's condition (exhibiting symptoms or not), the patient's history of previous UTI infections and previous treatment. Patients who have a symptomatic bacteriuria with colony counts $>10^5$ CFU/ml are usually not treated unless those patients are in an at-risk group such as the elderly, pregnant women, young girls or suffering from diabetes (Hannan et al., 2012; Kumar et al., 2015). High-risk group patients should be treated with a seven-day regime of Nitrofurantoin, Trimethoprim-Sulfamethoxazole and Fosfomycin. Patients with infection limited to the bladder are defined as having uncomplicated cystitis and are treated same as the high-risk group of asymptomatic bacteriuria patients. The antimicrobial dose should be continued for seven days as having the treatment for a shorter duration may result in a recurrent infection. Extended use of antibiotics may develop complicated cystitis. In this case, a further seven-day regime of Ciprofloxacin and Levofloxacin is recommended (Kumar et al., 2015).

Patients with fever, chills and flank pain, in addition to a positive laboratory examination, are diagnosed as pyelonephritis cases. The empirical treatment recommended is based on a Gram stain result to prevent any further complication. Empirically, Ciprofloxacin is recommended as an oral dose. However, if regional resistance to ciprofloxacin is more than 10 %, treatment should be initiated with Ceftriaxone or aminoglycoside followed by Ciprofloxacin (Kumar at al., 2015). Complicated pyelonephritis patients need to be treated in hospital with treatment including the use of one of Ceftazidime, Cefepime, Piperacillin-Tazobactam, Aztreonam, Meropenem or Imipenem for a fourteen-day period (Kumar at al., 2015).

1.2 *Escherichia coli* species

Escherichia coli was first known as *Bacterium coli commune*. This name was given by a German physician called Theodor Escherich in 1885 (Welch, 2006; Chaudhuri and Handerson, 2012; Croxen et al., 2013). The name of *Escherichia coli (E. coli)* became recognised in 1954 (Croxen et al., 2013). *E. coli* are, mostly, harmless commensal organisms that colonise the gastrointestinal tract of warm-blooded animals and become part of the normal flora within hours after birth (Kaper et al, 2004; Croxen et al., 2013).

Escherichia coli is a member of the family Enterobacteriaceae, which encompasses Gram negative, mostly oxidase negative facultative anaerobic bacilli. *E. coli* has the ability to grow over a wide range of temperatures with the optimal temperature in a range between 37 °C and 42 °C, and is commonly motile with peritrichous flagella (Welch, 2006; Croxen et al., 2013). *E. coli* isolates are antigenically variable, with those variations occurring because of the variability in somatic (O), capsular (K) and flagellar (H) antigens (Welch, 2006; Chaudhuri and Handerson, 2012; Croxen et al., 2013).

1.2.1 Pathogenic *Escherichia coli*

E. coli were considered to be non-pathogenic bacteria until they were isolated from several outbreaks of infantile diarrhoea in the 1940s (Chaudhuri and Handerson, 2012). Since then, *E. coli* have been classified based on genetic differences and various clinical presentations into commensal strains, intestinal pathogenic strains and extraintestinal pathogenic strains (Russo and Johnson, 2000). Pathogenic *E. coli* infections may cause diseases ranging from diarrhoea to meningitis, and from asymptomatic bacteriuria to lethal urosepsis (Kaper et al., 2004; Pitout, 2012a). The

intestinal pathogenic *E. coli* group comprises Entero-pathogenic, Entero-invasive and Entero-aggregative pathogenic *E. coli* strains while the Extraintestinal pathogenic *E. coli* include uropathogenic and neonatal meningitis *E. coli* strains (Kaper et al., 2004; Pitout, 2012a). *E. coli* have been classified into four phylogenetic groups based on the genetic relatedness indicated A, B_1, B_2 and D. The commensal and intestinal *E. coli* strains commonly belong to groups A and B_1. However, Extraintestinal pathogenic *E. coli* isolates mainly belong to B_2 and D (Pitout, 2012b).

The pathogenic strains differ from the non-pathogenic ones by possessing virulence factor genes which enable them to cause intestinal and extraintestinal diseases (Russo and Johnson, 2000; Croxen et al., 2013). There are around one million base pairs difference in genome size between commensal and pathogenic *E. coli* strains (Dobrindt et al., 2003). This variation in genome size is thought to be linked to the presence of diverse virulence factors which are required for pathogenic strains to survive and cause infection within the specific host (Wiles, 2008; Croxen et al., 2013). In general, the *E. coli* genome is composed of a conserved core and a flexible gene pool. The conserved core provides the backbone of genetic information that is responsible for the essential bacterial processes whilst the flexible gene pool provides the bacteria with the ability to adapt to new environments and to exhibit pathogenicity (Pitout, 2012a; Croxen et al., 2013). The flexible gene pool varies among various *E. coli* clones and contributes to intra-species variability and its size is dependent on the gain and loss of extra-genomic DNA (Pitout, 2012a).

1.2.2 Extraintestinal pathogenic *Escherichia coli*

Extraintestinal Pathogenic *E. coli* (ExPEC) are defined as a group of *E. coli* strains which have the ability to cause disease outside of the gastrointestinal tract. It

comprises two pathotypes: uropathogenic *E. coli* and neonatal meningitis *E. coli* (Kucheria et al., 2005; Lloyd et al., 2007). ExPEC also have the ability to exist in the gut as normal flora without causing disease (Russo and Johnson, 2003). However, ExPEC exhibit genome diversity and possess a wide range of virulence factors that are important for host colonization and can cause infection via dissemination to extraintestinal areas such as the blood, central nervous system (CNS) and urinary tract (Kucheria et al, 2005; Pitout, 2012a). ExPEC may acquire diverse virulence factors enabling them to establish a variety of infections such as bacteremia, nosocomial pneumonia, cholecystitis, cholangitis, peritonitis, cellulitis, osteomyelitis and arthritis. Thus, ExPEC may result in infection of the biliary and central nervous systems (Russo and Johnson, 2003; Hussain et al, 2012; Pitout, 2012b).

1.2.3 Uropathogenic *Escherichia coli*

The Uropathogenic *E. coli* (UPEC) are opportunistic intracellular organisms, and were first recognized in the 1970s as a cause of UTIs (Wiles et al., 2008; Agarwal et al., 2012). UPEC is the most common causative organism of UTIs and Extraintestinal human disease (Wiles et al., 2008). UPEC accounts for about 90 % of community-acquired infections and up to 50 % of nosocomial ones (Wiles et al., 2008; Toval et al., 2014). UPEC strains can colonise uroepithelial cells and indwelling medical surfaces resulting in a range of diseases such as asymptomatic bacteriuria, cystitis, pyelonephritis and urosepsis (Wiles et al., 2008; Spurbeck and Mobley, 2013). UPEC isolates exhibit variability in virulence gene repertoire and expression level which results in heterogeneity in bacterial growth and persistence within the urinary tract (Wiles et al., 2008).

1.2.3.1 Uropathogenic *Escherichia coli* virulence factors

The ability of UPEC to cause symptomatic UTIs is strictly related to the expression of various virulence factors (Bien et al., 2012). Virulence factor genes vary among UPEC isolates resulting in UPEC heterogeneity (Bien et al, 2012). Common UPEC virulence factors are linked to adhesion, the production of toxins, siderophore systems, capsule and lipopolysaccharide. These virulence factors are encoded by pathogenicity island and mobile DNA elements (Brzuszkiewick et al., 2006; Pitout, 2012a). UPEC virulence factors are divided into surface associated and secreted virulence factors. The surface associated virulence factors are attached to the surface of UPEC cells while the other virulence factors are secreted at the site of infection (Bien et al., 2012).

1.2.3.1.1 Adhesion

Attachment of UPEC to, and invasion of, bladder superficial cells is an important determinant of UPEC pathogenicity and initiation of UTIs (Mulvey et al., 2001; Bien et al., 2012). Adhesion is a common characteristic of pathogenic organisms to overcome fluid flow. In the case of UPEC, this fluid is normally urine (Johnson, 1991). In addition, adhesion is thought to trigger host and bacterial cell signalling directly, and to facilitate delivery of bacterial content to the host cells (Bien et al., 2012). Adherence of UPEC to urinary tract epithelial cells is mediated by fimbriae. Fimbriae are rod-shaped structures comprised of several subunits and adhesive tips. Promotion of adhesion and invasion may be associated with other structures and molecules such as flagella and toxins (Lüthje and Brauner, 2014). The most common adhesive organelles among UPEC are type 1 fimbriae and P fimbriae and 80 % of UPEC isolates express both or either type of these fimbriae (Mabbett et al., 2009).

UPEC type 1 fimbriae have the ability to bind to mannose receptors expressed by umbrella cells at the luminal surface of the urinary tract. This binding is mediated by *FimH* at the fimbrial tip to enhance UPEC survival, stimulate mucosal inflammation, mediate adhesion, invasion and promote intracellular bacterial community formation (Bien et al., 2012; Lüthje and Brauner, 2014). Binding of UPEC to host surfaces results in phosphorylation events that are required for the signalling stimulation pathway (Bien et al., 2012). *FimH* receptors are solubilised by the Tamm-Horsfall protein which is produced by the kidney and, because of this, the binding ability between UPEC and the urinary tract is limited (Bien et al., 2012). This may explain the higher frequency of type 1 fimbriae in strains isolated from the lower urinary tract (Lüthje and Brauner, 2014).

UPEC may also express P fimbriae that have a role in ascending infections and adhesion of uropathogens to the upper urinary tract, and also have the ability to induce the production of cytokines. P fimbriae are commonly expressed by UPEC in pyelonephritis patients (Bien et al., 2012; Lüthje and Brauner, 2014). P fimbriae are heteropolymeric fibres of different subunits encoded by the *Pap*A to K gene operon. P fimbriae have the ability to bind with the Galα of glycosphingolipid in kidneys. The attachment of P fimbriae with receptors in the kidney may result in the release of cermides that act against Toll-like receptor 4 (TLR4) leading to immune system activation (Bien et al., 2012).

There are other types of adhesive organelles in UPEC such as S fimbriae and Dr adhesions (Lüthje and Brauner, 2014). S fimbriae facilitate the binding of the uropathogen to epithelial and endothelial cells of the lower urinary tract and kidneys which results in the dissemination of the infection. This type of fimbriae is mostly

present in isolates from patients with meningitis, sepsis and ascending UTIs. However, Dr adhesin binds to type IV collagen in the kidney. This adhesion organelle is frequently found in chronic pyelonephritis isolates (Bien et al., 2012).

1.2.3.1.2 Flagella

Flagella (H antigen) are the motility organelles of UPEC and 70 % to 90 % of UPEC causing UTIs are flagellated. Although bacterial motility is not required for successful colonisation, flagella may contribute to UPEC migration to the upper urinary tract and allow further interaction between bacterial cell and host epithelium. It is thought that flagella may have a role in invasion of the renal collecting duct in pyelonephritis isolates and enable UPEC to reach the renal epithelium barrier to the blood stream (Subashchandrabose and Mobley, 2015). It was reported that cystitis isolates are motile than pyelonephritis isolates (Johnson, 1991).

1.2.3.1.3 Capsule and lipopolysaccharide

The capsule (K antigen) is a polysaccharide structure which surrounds the bacterial cell protecting the uropathogen from the immune system and phagocytic engulfment. There are more than 80 types of K antigen within the species *E. coli*. Although K_1, K_2, K_3, K_5, K_{12}, K_{13}, K_{20} and K_{51} are commonly seen in cystitis and pyelonephritis isolates, the most common K antigen among urinary tract isolates is K_1. Capsular types K_1 and K_5 are represented 63 % of isolates from women with pyelonephritis infections and these two antigens are also thought to have a role in protection from the humoral immune response. The presence of K antigens is significantly related to the ability of *E. coli* strains to resist serum. This ability can help bacteria to evade the complement cascade system, by blocking its activation (Johnson, 1991; Bien *et al*., 2012). Lipopolysaccharide (LPS) is an integral component

of the Gram-negative bacterial cell wall and has a role in pathogenicity. The endotoxin is well-known for its role as an activator of the host response, in addition to its ability to induce nitric acid and produce cytokines (Bien *et al*., 2012).

1.2.3.1.4 Toxins

Toxins are defined as secreted virulence factors that are produced by UPEC. Toxins are believed to be essential in the deep spreading of infections into underlying tissue. As the urinary tract environment has a limited source of nutrients, the secretion of toxins can damage the host tissue providing bacteria with access to nutrients and to evade the immune system (Bien et al., 2012; Lüthje and Brauner, 2014; Flores-Mireles et al., 2015).

The toxin most often secreted by UPEC is α-haemolysin. It is a lipoprotein and thought to be associated with upper UTIs such as pyelonephritis. A high level of α-haemolysin production may lead to erythrocyte and nucleated cell lysis which allows UPEC to cross the mucosal barrier and results in damaging of the effector immune cells and gaining host nutrients and iron. However, secretion of α-haemolysin at low levels can induce apoptosis of neutrophils, T lymphocytes and renal cells, and also promotes their exfoliation (Bien et al, 2012).

One third of pyelonephritis strains produce the cytotoxic necrotising factor 1 (CNF1). CNF1 is a protein secreted by *E. coli* and able to stimulate actin stress fibre and membrane ruffle formation leading to pathogen entry into the cell. There is another toxin secreted by pyelonephritis strains which is the secreted autotransporter toxin (SAT). This toxin is an important product of bacteria as it has toxic activity against bladder and kidney cells (Bien et al., 2012).

1.2.3.1.5 Iron uptake system

The iron uptake system is important for bacterial virulence. The utilization of siderophores for iron scavenging is an essential mechanism allowing uropathogens to grow and survive extraintestinally (Wiles et al., 2008; Hannan et al., 2012; Flores-Mireles et al., 2015). One iron acquisition system is expressed by both commensal and pathogenic *E. coli* strains and is known as enterobactin. Enterobactin competes with transferrin for iron binding (Wiles et al., 2008). There are additional iron acquisition systems which are encoded by UPEC: aerobactin and yersiniabactin, which are two siderophore mechanisms (Wiles et al., 2008; Flores-Mireles et al., 2015). Aerobactin is highly expressed by UPEC and characterised by its ability to bind with iron molecules and stability at low pH. Yersiniabactin has a role in protecting against intracellular killing and copper stress, and it is important in biofilm formation (Flores-Mireles et al., 2015). A comparison study by Lloyd et al, (2007) illustrated that enterobactin (*ent/feb*), enterobactin-like (*iro*), aerobactin (*iuc/iut*), yersiniabactin (*fyu*), iron transport (*sit*) and heme (*chu*) systems have a clear role in UPEC survival within the urinary tract. A recent study developing a vaccine against UTIs found seven different proteins related to the iron acquisition system that conferred immune protection. Four of these proteins (*IreA*, *Hma*, *IutA* and *FyuA*) may work as antigens that could develop the necessary protection against UTIs (Mobley and Alteri, 2016).

1.2.3.1.6 Biofilm formation and extracellular matrix components

Numerous pathogenic bacteria have the ability to form biofilms and this may contribute to pathogen protection from antibacterial treatments and host defence mechanisms. Biofilm formation is generally related to infection persistence (Lüthje

and Brauner, 2014; Flores-Mireles et al., 2015). In UPEC, having intracellular bacterial communities protects the bacteria from host immunity, antibiotic treatment and other environmental stresses are defined as biofilm formation and mediated by type 1 pili, antigen 43 and other adhesive organelles. In catheterised patients, biofilm formation is linked to the presence of type 1 pili (Flores-Mireles et al., 2015).

1.3 *Escherichia coli* antimicrobial resistance

Clinical diseases that are caused by ExPEC are widely treated with first line agents including cephalosporins, fluoroquinolones and trimethoprim-sulfamethoxazole. However, increasing resistance to these agents, which was first observed in the late 1990s, makes infection management challenging. β-lactamases are common enzymes that cause inactivation of β-lactam molecules. β-lactamases include Amp C β-lactamases, extended spectrum β-lactamases and carbapenemases. Increasingly common production of these enzymes is responsible for the development of resistance to all β-lactam antibiotics used in treating *E. coli* infections (Pitout, 2012b). Extended spectrum β-lactamases (ESBLs) are commonly seen in community-acquired isolates of *E. coli* and characterised by resistance to all β-lactam antibiotics apart from carbapenems and cephamycins (Pitout, 2012b). ESBLs have increased significantly in community isolates across Europe (Bevan et al., 2017). Between the 1980s and 1990s, two types of ESBLs were identified: the sulfhydryl reagent variable (SHV) and Temoneira (TEM). Then, ESBLs encoded by genes on plasmids (CTX-M) were described in 1983 and identified as different β-lactamases (Pitout, 2012b). CTX-M β-lactamase enzymes spread and became distributed worldwide, especially in *E. coli*. Extraintestinal pathogenic *E. coli* expressing CTX-M β-lactamases are commonly seen in UTIs, bacteraemia and intra-abdominal cases. CTX-M is usually carried on a

resistance plasmid that confers resistance to β-lactam, fluoroquinolones, aminoglycosides and trimethoprim-sulfamethoxazole antibiotics (Johnson et al, 2010). Global spreading of CTX-M producing bacteria may limit treatment options and could contribute to the emergence of carbapenemase-producing Enterobacteriaceae (Bevan et al., 2017). The type of CTX-M varies according to the geographical region with most types worldwide being CTX-M-14 and CTX-M-15 (Pitout, 2012b). Bevan et al. (2017) proposed factors which might be related to global dissemination of ESBLs such as plasmid content and food-animal transmission. The ability to home *bla*$_{\text{CTX-M}}$ on a plasmid has held great importance in the field of evolution, due to the relocation of the antimicrobial resistance gene via horizontal gene transfer in Enterobacteriaceae. *Inc*F plasmids predominantly carry *bla*$_{\text{CTX-M-15}}$ while various types of plasmid carry by *bla*$_{\text{CTX-M-14}}$ including *Inc*F and *Inc*K (Bevan et al., 2017). Spreading of a virulent clonal strain is also important, such as *E. coli* ST131, which is isolated worldwide and commonly carries CTX-M-15. In addition, other factors may contribute to the global spreading of ESBLs such as the existence of *bla*$_{\text{CTX-M}}$ in a variety of reservoirs in environment and food-producing animals (Bevan et al., 2017).

1.4 *Escherichia coli* Sequence types

UTIs are caused by UPEC isolates belonging to a number of multi locus sequence types and this is thought to reflect the pathogenic potential and likelihood of the establishment of infection (Croxall et al., 2011; Toval et al., 2014). Multi locus sequence typing (MLST) is a nucleotide-based grouping system based on the sequencing of internal fragments of selected genes (Tartof et al., 2005). It is a sequence typing tool that obtains the nucleotide sequence of ~450 bp fragments

derived from seven housekeeping genes in order to identify their allelic profiles (Chaudhuri and Henderson, 2012; Pitout, 2012b). In *E. coli*, there are seven housekeeping genes used for obtaining sequence types which are adenylate kinase (*adk*), fumarate hydratase (*fumC*), isocitrate/isopropylmalate dehydrogenase (*icd*), adenylosuccinate dehydrogenase (*purA*), DNA gyrase (*gyrB*), ATP/GTP binding motif (*recA*) and malate dehydrogenase (*mdh*). These housekeeping genes were selected based on the idea that they evolve slowly (Tartof et al., 2005). A previous epidemiological study by Croxall et al. (2011) demonstrated that the UPEC population was diverse and 52 sequence types were involved in causing UTIs in the Nottingham area, UK. Another study carried out by Lau et al. (2008) reported that the most common UPEC sequence type that caused UTIs is sequence type 131 followed by sequence types 69, 73 and 95. A study on ExPEC isolates reveals that *E. coli* bloods stream isolates generally belong to ST73, ST131, ST95, ST69 and ST12 respectively (Kallonen et al., 2017).

1.4.1 Sequence type 131

Sequence type 131 (ST131) is a pandemic clone of ExPEC and the dominant strain associated with extraintestinal disease. ST131 is commonly associated with UTIs and bloodstream infections especially in community isolates (Rogers et al., 2011; Croxall et al., 2011; Adams-Sapper et al., 2013). In addition, ST131 is highly prevalent among antimicrobial and multidrug resistant isolates (Johnson et al., 2010). A study by Clark et al. (2012) in Nottingham, UK reported that ST131 is the most common sequence type among ExPEC isolates. Transmission of ST131 between family members may occur to cause pyelonephritis and septic shock (Hannan et al., 2012).

ST131 belongs to phylogenetic group B2 and all ST131 isolates belong to the O25 serogroup. Further, ST131 is significantly associated with carrying CTX-M and OXA (oxacillinase) genes and characterised by frequent expression of CTX-M-15 which results in resistance to ciprofloxacin, trimethoprim, ampicillin and cefotaxime (Lau et al., 2008; Johnson et al., 2010; Croxall et al., 2011). Since 2007, ST131 has been associated significantly with an increase in ESBL-producing clinical *E. coli* isolates (Pitout, 2012b). In addition, ST131 isolates resistant to carbapenems are now extant through acquisition of bla_{KPC}, bla_{VIM} and bla_{NDM-1} genes (Adams-Sapper et al., 2013).

ST131 can be divided into sub-clusters and isolates of this sequence type may differ in genetic background, virulence profiles, antimicrobial resistance and ability to spread (Johnson et al., 2010; Croxall et al., 2011; Pitout, 2012b; Adams-Sapper et al., 2013). Martinez-Medina et al. (2009) showed that ST131 isolates may harbour a variety of virulence genes such as *PapC*, *PapGIII*, *ibeA*, *cnf1* and *hlyA*, and ST131 isolates possess one of five different types of *FimH* (Adams-Sapper et al., 2013).

1.4.2 Sequence type 69

Sequence type 69 (ST69) is widely disseminated in North America and belongs to phylogenetic group D. Isolates of this sequence type are commonly associated with blood stream infections (Leflon-Guibout et al., 2008; Adams-Sapper et al., 2013). ST69 in San Francisco, USA, was highly associated with students with UTIs and these isolates were often multidrug resistant, especially the clonal group A (CgA) which is a common clone of ST69 (Adams-Sapper et al., 2013; Skjøt-Rasmussen et al., 2013).

A study carried out in California reported that CgA isolates were commonly isolated from community acquired UTIs. In this study, 81 % of CgA isolates were carrying O1, O2, O4, O6, O7, O16, O18, O25 or O75 (Manges et al., 2001). However, in Europe, CgA

is commonly isolated from women with UTIs involving bacteraemia and characterised by sulfamethoxazole-trimethoprim resistance (Skjøt-Rasmussen et al., 2013). A study published within the UK by Gibreel et al. (2012) showed that ST69 isolates were more often isolated from females than males. ST69 was commonly associated with trimethoprim resistance and more associated with hospital isolates (Gibreel et al., 2012).

1.4.3 Sequence type 73

Sequence type 73 (ST73) is characterised by having a smaller genome size than other UPEC sequence types (Zdziarski et al., 2008). ST73 complex is highly diverse, and is a commonly isolated UPEC and belongs to phylogenetic group B2 (Gibreel et al., 2012). ST73 complex isolates showed more susceptibility to antibiotics than ST131 and ST69 isolates in a study carried out on patients with ExPEC bloodstream infections in San Francisco, USA (Adams-Sapper et al., 2013). Also, a study by Kallonen et al., (2017) showed a similar finding in the UK.

1.4.4 Sequence type 95

Sequence type 95 (ST95) belongs to phylogenetic group B2 and contains isolates from serogroups O1, O2 and O18 with polysaccharide antigen K1 commonly expressed among isolates of this sequence type. ST95 is commonly associated with community onset infections (Mora et al., 2009; Adams-Sapper et al., 2013). ST95 strains have been isolated from avian pathogenic *E. coli* (APEC), human uropathogenic *E. coli* (UPEC) and neonatal meningitis *E. coli* (NMEC) members (Johnson and Russo, 2005). ST95 isolates showed the lowest resistance to antibiotics of UPEC sequence types (Gibreel et al., 2012; Adams-Sapper et al., 2013). ST95 subgroups commonly possess *pap* genes that encode P fimbriae, a phenotype which

enhances colonisation of avian and human epithelium cells which is then thought to promote zoonotic transmission (Adams-Sapper et al., 2013; Maluta et al., 2014).

1.4.5 Sequence type 10

Sequence type 10 (ST10) is a complex of sequence types characterised by the carriage of five different ESBLS, comprising CTX-M-14, SHV-12, CTX-M-9, CTX-M-15 and CTX-M-32 (Oteo et al., 2009). ST10 complex isolates are commonly seen among faecal sample isolates of healthy *E. coli* carriers (Leflon-Guibout et al., 2008). In Brazil, ST10 clonal complex strains have been isolated from avian pathogenic *E. coli,* human UTI and sepsis cases. ST10 isolates were also thought to be associated with zoonotic infections worldwide (Maluta et al., 2014). This sequence type may carry the bla_{TEM} resistant gene (Adam-Sapper et a., 2013).

1.4.6 Sequence type 12

Sequence type 12 (ST12) is a clonal complex characterised by multidrug resistance to two or more classes of antibiotics (Adams-Sapper et al., 2013) and belongs to phylogenetic group B2 (Kallonen et al., 2017). ST12 is one of five predominant sequence types of ExPEC isolates (Kallonen et al., 2017). ST12 isolates can carry the bla_{TEM}, $bla_{CTX\text{-}M\text{-}14}$ and bla_{KPC} resistance genes (Adam-Sapper et a., 2013).

1.5 Within host bacterial diversity

Isolates within a bacterial species vary in antimicrobial resistance and virulence potential. Both these factors are considered to be formally adaptive mechanisms allowing bacteria to survive within the host and escape host immunity (Martinez and Baquero, 2002; Didelot et al., 2016). Variation within bacterial species is a result of evolution, and the evolutionary process in term of mutation rate accumulation per

site can happen over years (long-term evolution) or months, the latter being described as the short-term evolution (Didelot et al., 2016). Gene acquisition, gene loss or other genomic alterations are the key determinants of bacterial evolution and may result in the formation of a new strain (Dobrindt, 2005; Vejborg et al., 2011). Bacterial evolution mechanisms include point mutations, genetic rearrangement or horizontal gene transfer (HGT). Further, extreme bacterial evolution is highly associated with HGT in which long sequences are transferred via plasmids or pathogenicity islands. The acquired DNA sequences are thought to be responsible for transforming non-virulent strains to virulent ones and are linked with the transfer of antimicrobial resistance plasmids and virulence factors. Those genomic variations are commonly seen in the flexible gene pool and result in interspecies or intraspecies bacterial diversity. Closely related strains can acquire various horizontally transferred elements which may provide a specific trait to a particular strain (Dobrindt, 2005). Within species evolution is most easily detected by applying whole-genome sequencing. Homologous gene gain and non-homologous gene loss are evolutionary processes which can be measured. Also, the mutation rate per species can be evaluated (Didelot et al., 2016).

1.5.1 *Staphylococcus epidermidis* within-host diversity

Over time, polyclonal (varied sequence types within a single host) and monoclonal diversity (genetic variation within same sequence type) were demonstrated in patients infected with *Staphylococcus epidermidis*. Monoclonal diversity leads to phenotypic variation and can result in diverse drug resistance profiles in isolates from the same sequence type from a single patient. The reason behind such monoclonal diversity is thought to be related to mutations, DNA rearrangement or

loss. Further, polyclonal diversity associated with multi sequence types of *S. epidermidis* was also determined in four out of fourteen studied patients (Galdbart et al., 1999).

1.5.2 *Helicobacter pylori* within-host diversity

Genetic diversity was determined in a patient infected with *Helicobacter pylori* over a period of six years. Genetic variation within the flexible gene pool was found in the more recent isolates which became more tolerant to antibiotic treatment (Israel et al., 2001). Genetic variability was determined within and between isolates from a mother and her children who were all infected with *H. pylori*. Although the clonal isolate was the dominant strain between family members, bacterial change was shown between individuals and thought to have resulted from pathogen adaption to the newly colonised host. In addition, the mother was infected with two distinct strains of *H. pylori* (Kivi et al., 2007).

1.5.3 *Staphylococcus aureus* within-host diversity

Within-host diversity was evaluated among isolates from samples obtained over a year and half from individuals infected with *Staphylococcus aureus*. Samples were categorized into two groups: a group colonized with different strains of *S. aureus* and group containing very similar strains with minor genetic variation (Cespedes et al., 2005). *S. aureus* has been shown to exhibit phenotypic variation resulting in small colony variants to be produced in a single patient, and other changes were found in their virulence factor content and a reduction in susceptibility to antibiotics (Tuchscherr et al., 2011). This study correlated with data from McAdam et al. (2011) where, although three isolates of *S. aureus* from a single host were of similar sequence type, 31 point mutations were identified between them. The recognized

mutations were thought to be related to genes responsible for growth, virulence and antibiotic resistance (McAdam et al., 2011). In addition, 131 colonies from 13 individuals were sequenced to evaluate within host diversity of nasal *S. aureus* isolates. These isolates commonly exhibited microevolution (mutations) and, to a lesser extent, mobile genetic element variation (Golubchik et al., 2013).

1.5.4 *Burkholderia dolsa* within-host diversity

An epidemiological study using whole genome sequencing was carried out on 114 isolates of *Burkholderia dolsa* from 14 patients. Most of the isolates were from patients' airways and there were some blood isolates. Dissemination of infection to bloodstream was reported to be associated with multiple strains of *B. dolsa* within the same individual (Lieberman et al., 2011). Dozens of colonies were sequenced at a single time point from a patient infected with *B. dolsa* and deep sequencing was also performed on isolates from five patients. Results of this research concluded that the collected samples followed the diverse community model which is thought to be derived from adaption of the pathogen to the environment under selective pressure, so generating polymorphic mutations (Lieberman et al., 2014).

1.5.5 *Mycobacterium tuberculosis* within-host diversity

Multiple infections may occur in patients infected with *Mycobacterium tuberculosis*, with two strains of *M. tuberculosis* found in 19 % of patients examined. This was seen more frequently in patients with a previous infection (Warren et al., 2004). Microdiversity of a single strain within a single host was shown and intraspecies diversity was indicated in patients infected with *M. tuberculosis*. The reason was thought to be related to an elevation of the mutation rate in persistent infections resulting in the development of multidrug resistant strains (Warner et al., 2015).

O'Neill et al. (2015) reported that most within and between host variation of isolates from people infected with *M. tuberculosis* occurred in genes responsible for pathogenicity.

1.5.6 *Pseudomonas aeruginosa* within-host diversity

The evolution of *Pseudomonas aeruginosa* was assessed in one cystic fibrosis patient over a period of time. The first sample was taken after six months and the second sample was taken after 96 months. Isolates from the second sample had accumulated mutations and it was thought that those mutations affected protein function. Isolates from the second sample varied phenotypically by the loss of motility, pyoverdine production, secreted protease and biofilm formation. Mutations also affected the multidrug efflux pump genes, making the later isolate more resistant to treatment. It had become resistant to aminoglycosides, amikacin, gentamicin and tobramycin when compared with the six-month isolate which was resistant to β-lactam agents only (Smith et al., 2006). In another study, two colonies from three patients infected with *P. aeruginosa* were selected for genome sequencing to assess genetic variation. Colonies from samples displayed genetic variation thought to have an impact on cell physiology and gene expression (Chung et al., 2012). Another study was carried to determine the evolution of *P. aeruginosa* within a single host over a period of time, and showed that long-term colonisation led to increased mutation, especially in genes responsible for pathogen adaption to the host (Feliziani et al., 2014). Within-host diversity was also determined in four Italian patients infected with *P. aeruginosa* over time with similar findings to the previous study (Marvig et al., 2015). Moreover, bacterial diversity within a single host at single time point was evaluated by Darch et al. (2015) where 44 colonies of *P.*

aeruginosa were compared phenotypically and genetically within a single patient. The result of this study determined that the isolates varied in antibiotic resistance profile, pyocyanin, LasA, LasB and quorum sensing production. In addition, whole genome sequence was carried out on 22 colonies and it was determined that phenotypic variation may be related to recombination events (Darch et al., 2015).

1.5.7 *Escherichia coli* within-host diversity

Comparative genome sequence analysis reveals that ExPEC exhibit genetic diversity underlying their phenotypic diversity (Dobrindt, 2005). ExPEC strains belonging to phylogenetic groups B2 and D have a larger genome size with a higher percentage of virulence genes than isolates of phylogenetic groups A and B1 (Zdziarski et al., 2008). Horizontal gene transfer, gene loss and gain have a played role in the evolution of UPEC and drive phenotypic variation (Brzuszkiewicz et al., 2006).

Levert et al. (2010) reported within-host phenotypic and genetic diversity of *E. coli* among patients with extraintestinal infections. Three outcomes were obtained in this research; patients with no diversity, patients showing within sequence type heterogeneity (monoclonal diversity) and patients with infections caused by a number of sequence types of *E. coli* (polyclonal diversity) (Levert et al., 2010). Further, molecular analysis for a family infected with shiga toxigenic *E. coli* (STEC) revealed that family members were infected with multiple STEC strains. The strains detected were thought to have pathogenic potential and potentially could cause the development of symptoms (Staples et al., 2012). In 2013, isolates from one patient in a genomic analysis study revealed the presence of two different sequence types, ST131 in his blood samples and ST10 in his urine samples. Reanalysis of urine culture plates demonstrated the presence of two other UPEC sequence types (ST131 and

ST127). In that study, a genetic analysis was conducted to compare the blood and urine isolates from patients who had similar sequence types, and some mutations were detected in the two isolates from a single patient, and there was a difference in serum resistance (McNally et al., 2013). In another study, sixteen single colonies were sequenced from eight fecal samples in order to evaluate within-host diversity. One individual had multiple sequence types of *E. coli*. Extensive heterogeneity was also observed in the accessory genome between individual strains (Stoesser et al., 2015).

1.6 Aim of this research

Urinary tract infection is commonly caused by *Uropathogenic Escherichia coli* (UPEC). Clinical diagnosis and the choice of the most effective treatment in the clinical laboratories are based on single colony selection, assuming that identical colonies of a homogenous bacterial culture are similar. However, within-host diversity is well recognized in many organisms including *E. coli*. Evolution of diversity within a patient at a single time point may have a crucial impact on the correct diagnosis of the UTI and its treatment. This PhD research aims to investigate UPEC diversity within a single host. Determining within-host diversity was achieved through three stages:

- Sensitivity testing with eight different classes of antibiotics for a maximum of forty-eight well-isolated colonies on collected UPEC culture plates.
- Comparing selected colonies within samples via various phenotypic assays using motility assay, biofilm formation assay using crystal violet and association and invasion assay to T_{24} epithelial cell line.
- Investigation of genetic diversity within collected samples using whole-genome sequencing technology.

Chapter Two

Materials and Methods

2 .1 Materials and Methods

2.1.1 Collection of bacterial strains

A total forty-two uropathogenic *E. coli* pure culture plates were used in this research from patients with urinary tract infections. Patients' bacterial slopes were collected from the Clinical Microbiology Department at Queen Medical Centre (QMC), Nottingham, UK. Collected bacterial slopes were selected randomly with no identifiable information except the sex of the patients. Samples were assigned as UTI with given serial number containing the letter 'F' for female patients or letter 'M' for male patients. Samples were collected over a six months period in 2014. Table 2.1 shows details of the samples details that used in this research.

2.1.2 Bacterial cultures maintenance and storage

Uropathogenic *E. coli* samples were collected as nutrient agar slopes and sub-cultured on LB (Sigma-Aldrich, UK) agar plates in order to have well-isolated colonies. The initial streak of each of the collected samples were maintained in 1 ml of Lysogeny Broth (LB) containing 20 % glycerol (Fisher Scientific) and stored at – 80 ˚C. The susceptibility patterns of the varied well-isolated colonies, within each sample of the high and low diversity groups, were maintained. A single representative colony (bacterial isolate) of the identical group samples, with identical sensitivity patterns, were also stored. The saved bacterial isolates were categorised by colony number, sample number and the letter 'F' or 'M'.

Table 2.1 Patient samples used in this research:

Sample ID	Source	Gender
UTI (F1)	Urine	Female
UTI (F2)	Urine	Female
UTI (F3)	Urine	Female
UTI (F4)	Urine	Female
UTI (F5)	Urine	Female
UTI (F6)	Urine	Female
UTI (F7)	Urine	Female
UTI (F8)	Urine	Female
UTI (F9)	Urine	Female
UTI (F10)	Urine	Female
UTI (F11)	Urine	Female
UTI (F12)	Urine	Female
UTI (F13)	Urine	Female
UTI (F14)	Urine	Female
UTI (F15)	Urine	Female
UTI (F16)	Urine	Female
UTI (F17)	Urine	Female
UTI (F18)	Urine	Female
UTI (F19)	Urine	Female
UTI (F20)	Urine	Female
UTI (F21)	Urine	Female
UTI (M1)	Urine	Male
UTI (M2)	Urine	Male
UTI (M3)	Urine	Male
UTI (M4)	Urine	Male
UTI (M5)	Urine	Male
UTI (M6)	Urine	Male
UTI (M7)	Urine	Male
UTI (M8)	Urine	Male
UTI (M9)	Urine	Male
UTI (M10)	Urine	Male
UTI (M11)	Urine	Male
UTI (M12)	Urine	Male
UTI (M13)	Urine	Male
UTI (M14)	Urine	Male
UTI (M15)	Urine	Male
UTI (M16)	Urine	Male
UTI (M17)	Urine	Male
UTI (M18)	Urine	Male
UTI (M19)	Urine	Male
UTI (M20)	Urine	Male
UTI (M21)	Urine	Male

2.2 Culture media

2.2.1 Lysogeny Broth (LB) Agar

Media was ordered from Sigma-Aldrich, UK. It consists of 5 g/L yeast extract, 10 g/L sodium chloride, 10 g/L tryptone, and 15 g/L granulated agar. 40 g of LB agar was dissolved in 1 L distilled water (dH_2O) and sterilised by autoclaving for 15 minutes at 121 °C.

2.2.2 Lysogeny Broth (LB)

Media was ordered from Sigma-Aldrich, UK. It consists of 5 g/L yeast extract, 10 g/L sodium chloride and 10 g/L tryptone. 20 g of LB broth was dissolved in 1L distilled water (dH_2O) and sterilised by autoclaving for 15 minutes at 121 °C.

2.2.3 Cystine Lactose Electrolyte Deficient (CLED) agar

Media was ordered from Oxoid Limited, UK. It consists of 0.128 g/L L-cystine, 10 g/L lactose, 4 g/L tryptone, 0.02 g/L bromothymol blue and 4 g/L peptone. 36.2 g was dissolved in 1 L distilled water (dH_2O) and sterilised by autoclaving for 15 minutes at 121 °C.

2.2.4 Iso-Sensitest agar

Media was ordered from Oxoid Limited, UK. It consists of 11 g/L hydrolysed Casein, 3 g/L peptones, 2 g/L glucose, 3 g/L sodium chloride, 1 g/L soluble starch, 2 g/L disodium hydrogen phosphate, 1 g/L sodium acetate, 0.2 g/L magnesium glycerophosphate, 0.1 g/L calcium gluconate, 0.001 g/L cobaltous sulphate, 0.001 g/L cupric sulphate, 0.001 g/L zink sulphate, 0.001 g/L ferrous sulphate, 0.002 g/L manganous chloride, 0.001 g/L menadione, 0.001 g/L cyanocobalamin, 0.02 g/L L-cystine hydrochloride, 0.02 g/L L-tryptophan, 0.003 g/L pyridoxine, 0.003 g/L pantothenate, 0.003 g/L nicotinamide, 0.0003 g/L biotin, 0.00004 g/L thiamine, 0.01

g/L adenine, 0.01 g/L guanine, 0.01 g/L xanthine, 0.01 g/L uracil and 8 g/L agar. 31.4 g was dissolved in 1 L distilled water (dH_2O) and sterilised by autoclaving for 15 minutes at 121 °C.

2.2.5 Technical agar

Media was ordered from Oxoid limited, UK. A 3 g of technical agar used with 1 L of LB broth for motility assay followed by autoclaving 15 minutes at 121 °C.

2.3 Antibiotic used for susceptibility testing

Antibiotic resistance tests of all isolates were performed with reference to the protocol for standardized disk diffusion method provided by The British Society for Antimicrobial Chemotherapy (BSAC) (BSAC, 2013). A panel of eight antibiotics were selected and are described in table 2.2.

Table 2.2 Antibiotic disks used in chapter three

Antibiotic	Experiment	Supplier
Ampicillin (10 µg) **Cefepime (30 µg)** **Cefpodxime (10 µg)** **Ceftazidime (30 µg)** **Ciprofloxacin (1 µg)** **Gentamicin (10 µg)** **Meropenem (10 µg)** **Trimethoprim (2.5 µg)**	Sensitivity Testing	Mast Company limited, UK

2.4 General media, buffers and reagents

2.4.1 Saline solution

Saline solution was used as a diluent of bacterial culture (Oxoid Limited, UK). One saline tablet was added to 500 ml dH_2O followed by autoclaving at 121 ˚C for 15 minutes.

2.4.2 Glycerol solution

80 % of glycerol solution was used to make bacterial stocks (Fisher Scientific, UK) and was prepared by adding 80 ml of glycerol into 20 ml dH_2O.

2.4.3 Crystal violet solution

1 % of crystal violet solution was used to strain biofilm in biofilm formation experiment (Sigma Aldrich, UK). It composed of 2.3 % certified crystal violet, 0.1 % ammonium oxalate and 20 % ethyl alcohol.

2.4.4 McCoy's 5A modified medium

McCoy's 5A modified medium was used as a minimal defined medium in bacterial growth studies and also as a growth and infection medium in cell culture assays (Sigma Aldrich, UK). It is ready to use and contains 2.2 g/L sodium bicarbonate.

2.4.5 Dulbecco's phosphate buffer saline (PBS)

Dulbecco's phosphate buffer saline (PBS) was used to wash cells during tissue culture work (Sigma Aldrich, UK). It is composed of sodium and magnesium chloride solution.

2.4.6 Penicillin-Streptomycin Antibiotics

Penicillin-Streptomycin antibiotic was purchased as a sterile liquid bio reagent (Sigma Aldrich, UK). It was manufactured by mixing 10000 units of penicillin with 10 mg streptomycin/ml. 5 % (5 ml) of penicillin-streptomycin antibiotics was added to

the McCoy's 5A modified medium to create the growth medium for T24 cell line. Penicillin- Streptomycin antibiotics were stored at 4 ˚c.

2.4.7 Gentamicin Antibiotic Solution

Gentamicin (Sigma Aldrich, UK) was used in the invasion assay to kill the extracellular bacteria with a concentration of 100 µg/ml. Gentamicin antibiotic solution was stored at 4 ˚C.

2.4.8 Triton X-100

Triton X-100 was used during cell culture studies (Sigma Aldrich, UK). 1 % triton X-100 was prepared by adding 100µl of Triton X-100 to 10 ml of dH_2O (V/V).

2.4.9 Nextera XT DNA library prep reagents

	Reagents (Illumina, Inc)
Tagment Genomic DNA (Illumina, Inc)	
Nextera XT library prep kit (Box 1 of 2)	(ATM) Amplicon tagment mix
	(TD) Tagment DNA buffer
Nextera XT library prep kit (Box 2 of 2)	(NT) Neutralise tagment buffer
Library amplify (illumine, Inc)	
Nextera XT library prep kit (Box 1 of 2)	(NPM) Nextera PCR master mix
Nextera XT index kit	Index 1 primers (N7) and index 2 primers
PCR clean up (Illumina, Inc)	
Nextera XT library prep kit (Box 1 of 2)	(RSP) Resuspension buffer
AMPure XP beads (Beckman Coulter, Inc)	
80% freshly prepared ethanol	
Pool library	

2.5 Methods

2.5.1 Disk diffusion method

Isolates were screened for sensitivity testing to eight different antibiotics that belong to the aminoglycoside, penicillin, cephalosporin, carbapenems, quinolones classes and others such as trimethoprim. This was performed on each well-isolated colony (maximum 48 colonies per sample) based on the British Society for Antimicrobial Chemotherapy (BSAC, 2013) using the disk diffusion method. Commercial antibiotics used in this research project were: ampicillin (10 µg), cefepime (30 µg), cefpodoxime (10 µg), ceftazidime (30 µg), ciprofloxacin (1 µg), gentamicin (10 µg), meropenem (10 µg) and trimethoprim (2.5 µg). A cotton swab was dipped in 0.5 McFarland solution of bacterial isolate, rolled around the tube and then spread over the surface of an iso-sensitest agar plate. *E. coli* ATCC 10418 and ATCC 11560 were used as sensitive and resistant control for quality purposes. Plates were incubated at 37 °C overnight. The zone of inhibition was measured in mm and compared with the defined standards on British Society of Antimicrobial Chemotherapy (BSAC; BSAC, 2013).

2.5.2 Motility assay using semi solid agar plates

For the motility assay, all isolates from - 80 ˚C stock were incubated in Lysogeny Broth (LB) broth at 37 ˚C for 18 hours with shaking (200 rpm). A 3 µl drop was applied onto a motility agar plate containing LB broth and 0.3 % Bacto agar. The plates were incubated at 37 ˚C for 18 hours. The motility assay was assessed by measuring the zone of migration through the agar. The motility ability of each isolate was measured, diameter in mm as the mean of three independent experiments. *Salmonella enteritidis* and *Staphylococcus aureus* were used as positive and negative controls respectively.

2.5.3 Biofilm formation assay using crystal violet

Biofilm formation was analysed in a 96-well plate assay at 37 °C for 24, 48, 72 and 96 hours. First of all, all variants were grown at 37 °C for 18 hours without shaking. The optical density was adjusted to 0.5 - 0.7 using fresh LB broth and measured in a spectrophotometer at 600 nm. Two hundred microliters was transferred in triplicate to a 96 well micro titre plates and incubated. At set time points the wells were washed three times with saline and then stained with two hundred microliters (0.1 % V/V) crystal violet for 15 minutes at room temperature. Two hundred microliters of ethanol was added to each well after stain removal and incubated at room temperature for 10 minutes. Finally, one hundred microliters from each well was transferred to a new micro titre plate. The optical density at 600 nm measured using a micro titre plate reader. The biofilm formation of each isolate was measured as the mean of absorbance of three independent experiments. The negative control was LB broth.

2.5.4 Tissue association and invasion assay

2.5.4.1 Culture cell line

The human bladder epithelial T24 cell line was grown in a growth medium (Table 2.3) at 37 °C in humidified condition under 5 % CO_2 incubator. Cells were subcultured every 2 to 3 days. The confluent monolayer was then de-attached using Trypsin–EDTA and incubated in a CO_2 incubator for 5 minutes. The cell suspension was transferred to a 15ml falcon tube and centrifuged for 5 minutes at 1200 rpm. The pellet was then suspended in 3-5 ml of fresh growth media. T24 cell line was then used to subcultured in 24 well micro plates at concentration of 5×10^4 cell/ml two days before association and invasion assay.

2.5.4.2 Association and invasion assay

Bacteria were grown overnight in 1ml LB broth and incubated at 37 ˚C with shaking (200 rpm). Cultures were centrifuged at 15000 x G for 1 minute and pellets suspended in infection media (1:100). The 24 well micro plates that contain T_{24} cell were infected with 500 µl of bacterial suspension in duplicate and incubated at 37 ˚C in CO_2 incubator for 3 hours. For association assays, cells were washed twice with Dulbecco's phosphate buffered saline and lysed with 1 % triton x-100 solution. Serial dilutions of lysed cells were plated in duplicate on LB agar and incubated overnight at 37 ˚C. For invasion assay, after 3 hours incubation, the 24 well micro plates were washed twice with Dulbecco's phosphate buffered saline and fresh infection media containing 100 µg/ml of gentamicin solution was added and incubated for 2 hours at 37 ˚C in CO_2 incubator. After that, cells were washed with Dulbecco's phosphate buffered saline, lysed with 1% triton x-100, diluted serially with saline and finally plated out in duplicate on LB agar plates.

Colonies were counted using the Miles & Misra technique to determine the level of bacterial association and invasion. These assays were measured in CFU/ml as the mean of three independent experiments. In these assays, DH5α and CFT073 were used as negative and positive controls respectively.

The growth and infection media were prepared as described in the table 2.3.

Table 2.3 Growth and infection media preparation

Growth Media	Infection Media
Addition of 10 % of Foetal bovine serum, 5 % of Penicillin-streptomycin and 5 % of non-essential amino acid solution to Macoy's 5A Medium.	Addition of 10% of Foetal bovine serum and 5 % of non-essential amino acid solution to Macoy's 5A Medium.

2.6 Next-generation sequencing

2.6.1 Genomic DNA extraction

Genomic DNA was extracted from uropathogenic *E. coli* strains pre-grown in LB agar cultures. A uropathogenic *E. coli* colony was inoculated into sterile LB broth at 37 ˚C for 18 hours. The genomic DNA was extracted from 1 ml of the LB broth cells suspension using the GenEluteTM Bacterial Genomic DNA kit (Sigma Aldrich, UK) according to the procedure described in the manufacturer's user guide booklet.

2.6.2 Nanodrop

The purity and concentration for all extracted genomic DNA was measured using a Nanodrop 2000 (Thermo Scientific, UK). Extracted DNA samples only accepted with a minimum 260/280 nm values of 1.8.

2.6.3 Qubit

Qubit was used to quantify the extracted DNA and normalised to 0.2 ng/µl using High Sensitivity (HS) Qubit kit (Qubit 3.0 flourometer) (Thermo Fisher Scientific, Inc) with range between 0.2- 100 ng.

2.6.4 Nextera XT DNA library

Next generation sequencing was applied on extracted samples of bacterial isolates using Miseq platform (Illumina, Inc). Next generation sequence was performed according to the protocol described in the Nextera XT DNA library prep reference guide. Following to PCR amplification, PCR products were cleaned up using AMPure XP beads (Beckman Coulter, Inc). Libraries then were normalised to 4 nm after checking the DNA concentration using Qubit HS assay. Libraries were loaded onto Miseq reagent Kit V2, paired ends, 251 cycles.

2.7 Bioinformatics

2.7.1 Genome assembly

De novo assemblies were performed using SPAdes (Version 3.5.0; Bankevich et al., 2012). SPAdes provides computational efficiency and strong performance for small genome. Further, SPAdes provides (careful) flag which reduce errors and short indels. SPAdes assembly was performed using the command line:

```
SPAdes.py --pe1-1 /file location.FASTQR1 --pe1-2 /file location.
FASTQR2 -careful –o /file name
```

2.7.2 QUAST

Genome assembly quality was assessed using QUAST version 2.0. QUAST is easy, representative and informative software that provides a range of metrics presented to assess the quality of assembly (Gurevich et al., 2013). QUAST was performed using the command line:

```
QUAST.py /file name –o /file name
```

2.7.3 Genome annotation

Prokka is a rapid and accurate tool used for bacterial genome annotation (Seemann, 2014). Assembled FASTA files are the only acceptable format for this software and are required to start the Prokka script where it utilizes a built in database to annotate the genome and produces output files in GFF formats which are compatible annotation formats ready for further analysis (version 1.12; Seemann, 2014). Prokka annotation was performed using the command line:

```
Prokka –outdir /file location –usegenus /file name --locustag *.FASTA
```

2.7.4 Genome mapping

All sequenced uropathogenic *E. coli* raw FASTQ files were aligned to their respective *de novo* assembled genome (fasta files) using CSI phylogeny, CGE website (https://cge.cbs.dtu.dk/services/CSIPhylogeny/) (version 1.4). CSI phylogeny is a web-based used for calling and filtering SNPs using BWA and SAMtools. BWA software was developed and optimised by Li and Durbin, (2009). BWA version 3.0 provides a more accurate and fast alignment of paired ends sequencing reads produced by Illumina (Li and Durbin, 2009). SAMtools version 1.4 was used to extract the required information from SAM files generated by the BWA alignment process. SAMtools software is used to manipulate the data in SAM format files such as indexing, merging, sorting and other post processing requirements (Li et al., 2009; Evolution and Genomics, 2011). This results in VCF files containing all the SNP data for each strain against their reference genome.

2.7.5 Artemis ACT

Artemis (version 2) is a genomic tool used for visualisation the annotated genome (Rutherford et al., 2000).

Chapter Three

Multiple resistance profiles present in similar morphological colonies within a single Uropathogenic *Escherichia coli* patient

3.1 Introduction

3.1.1 Microbiological diagnosis of positive UTIs

In microbiological laboratories, clinical diagnosis is based on single colony selection especially if identical morphological colonies are obtained within a bacterial culture plate. This assumes that a single colony is representative of the entire bacterial population. Based on that assumption, bacterial isolate identification, antibiotic susceptibility and further biochemical testing are carried out on the positive urine culture plate in dedicated laboratories (Szczepura, 1991; Johnson et al., 1995).

3.1.1.1 Antimicrobial susceptibility testing using disk diffusion method

Antimicrobial susceptibility testing is a common technique to validate the sensitivity of bacterial isolates to an antimicrobial agent (Szczepura, 1991). This is to be able to select the most effective treatment for the bacterial infection (BSAC, 2012). The method of choice by both the European Committee on Antimicrobial Susceptibility Testing (EUCAST) and the British Society of Antimicrobial Chemotherapy (BSAC) is the disk diffusion method (BSAC, 2012). The disk diffusion method provides an accurate and reproducible way of examining antimicrobial sensitivity (BSAC, 2012) and as such is the most routinely applicable method in clinical laboratories (NCCLS, 1992). The method allows clinicians to examine different antimicrobial agents on clinical isolates at a single time point (CLSI, 2012). The results of disk diffusion tests can be easily interpreted by measuring the zone of inhibition surrounding the antibiotic disk and based on the measured diameter, an isolate is categorised as sensitive, intermediate or resistant to the tested antimicrobial agent (Jenkins and Schuetz, 2012).

3.1.1.2 Antibiotics used for treating urinary tract infections

Many classes of antimicrobial agents are used to treat urinary tract infections, these agents vary in their mode of actions. Currently, urinary tract infections are treated with β-lactams, aminoglycosides, trimethoprim-sulfamethoxazole and fluoroquinolone antibiotics (Gupta et al., 2011). Since the 1990s, antibiotic resistance has emerged that makes management of urinary tract infections challenging (Pitout and Laupland, 2008). Due to the emergence of antimicrobial resistance among clinical isolates, the disk diffusion method should be employed using a wide range of antibiotic classes to guide clinicians as to the most effective antimicrobial agent. Therefore, BSAC guidelines recommend that when testing the susceptibility of uropathogens, screening should be performed using six classes of antibiotics disks which are aminoglycosides, penicillin, cephalosporins, quinolones, carbapenems and others such as trimethoprim (BSAC, 2013).

3.1.2 Variation in susceptibility profiles within a single patient

Several studies have revealed that there might be phenotypic and genotypic variation between bacteria isolated from a single host. Studies have determined the variation between bacteria isolated from a single patient in morphology and susceptibility to antibiotics over time. One of the earliest studies by Thomassen (1979) reported that mucoid and non-mucoid colonies of *P. aeruginosa* from cystic fibrosis patients have a varied susceptibility profile. For a single patient and over a six-month period, it was noticed that isolated bacterial colonies varied in their morphological appearance, lost their mucoid ability, and became more resistant to treatment. As morphological changes in colony appearance might be related to

varied resistance profiles, Wolter et al. (1995) recommended mixing various morphological colonies to provide an efficient result for susceptibility testing. This was carried out on bronchial secretions of patients colonised with *P. aeruginosa* in cystic fibrosis cases (Wolter at el, 1995). Due to complications in treating *Pseudomonas* infection and in some cases, the most susceptible agent may not improve the patient condition, the possibility of the presence of diverse isolates within a patient was presumed. Therefore, the idea of applying sensitivity testing on four identical colonies isolated from single cystic fibrosis patients infected with *P. aeruginosa* was demonstrated by Foweraker and co-workers in 2005. This study revealed that over time only one patient showed an identical resistance profile, 51 % of collected samples patients had mixed colony morphotypes and different sensitivity results while 15 % of isolated colonies were similar morphologically but still varied in their resistance patterns (Foweraker et al., 2005).

Further variation in resistance profile was recognised in patients infected with *S. aureus*. Bacteria isolated from blood of the same patient at different time points became tolerant to vancomycin treatment (Sieradzke et al., 2003). In 2007, Mwangi et al., carried out a PCR sequence analysis on two isolates from the previous study and 35-point mutations were identified in the resistant isolate which were thought to be related to vancomycin resistance. Another study was carried out on three *S. aureus* isolates that were collected at three different time points, from cystic fibrosis patients in order to determine phenotypic and genetic variation. Applying whole-genome sequencing determined that these isolates were genetically varied and that this could affect antibiotic sensitivity (McAdam et al., 2011).

Variation within hosts infected with ExPEC has also been recognised. In 2010, a molecular study was carried out on 126 isolates from nineteen patients. The determined types of diversity (polyclonal and monoclonal) where shown to have an impact on bacterial phenotypic features such as antibiotic resistance and pathogenic potential (Levert et al., 2010).

Recently, within-host variation of bacterial pathogens has been determined at a single time point. Darch et al. (2015) determined the level of phenotypic and genotypic diversity of forty-four identical colonies from a single cystic fibrosis patient infected with *P. aeruginosa* and variation in antibiotic sensitivity were determined between colonies from a single cystic fibrosis patient infected with *P. aeruginosa* (Darch et al., 2015).

3.1.3 Aim of the chapter

Uropathogenic *E. coli* is the most common causative agent of urinary tract infections (Kucheria et al., 2005) and the presence of more than one sequence type of uropathogenic *E. coli* within a single urosepsis patient has been reported (McNally et al., 2013). As such the aim of this research was to determine the population diversity level of uropathogenic *E. coli* within multiple patients at a single time point. To investigate this, we started by examining antimicrobial sensitivity using the disk diffusion method on a maximum of forty-eight well-isolated colonies from a given urine sample. The rationale being that changes in sensitivity profile can be a sign for bacterial diversity and can provide us with an initial prediction of the levels of diversity present in our collected samples.

3.1.4 Strains used in this chapter for disk diffusion method

The disk diffusion method was performed on bacteria isolated from 21 female and 21 male samples collected from Queen Medical Centre Hospital (QMC) in Nottingham. A maximum of forty-eight well-isolated colonies from each sample were compared using 8 different classes of antibiotics that belong to the aminoglycoside, penicillin, cephalosporin, carbapenem, quinolones classes and others such as trimethoprim as referred in table 2.2 in Chapter 2. Isolated colonies from each collected sample were streaked out on LB agar plates and assigned with a specific colony number including the original patient sample number. A swab from the entire original LB plate and colonies that exhibited varied resistance profiles within patient samples were also stored in Lysogeny Broth (LB) broth containing 20 % (V/V) glycerol at - 80 °C.

3.2 Results

3.2.1 Uropathogenic *E.coli* varied resistance profile patterns

The antimicrobial sensitivity of screened colonies within each sample was determined based on the BSAC disk diffusion test, 2013. Due to the presence of differing levels of variation in resistance patterns, samples were classified into three resistance profile patterns table 3.1. We found patients where all isolated colonies had identical susceptibility profiles for all examined antibiotics. We defined this pattern as an identical resistance profile pattern as shown in (A) in table 3.1. Some patients had colonies exhibiting changes in their resistance profile between an intermediate zone of inhibition and either a resistant or sensitive pattern for a given antibiotic(s). We defined this pattern as a low diverse resistance profile patterns as

shown in (B) in table 3.1. The last pattern we found was the highly diverse resistance profile pattern. This profile pattern is defined when isolated colonies within each sample vary in their susceptibility profile between sensitive and resistant for one or more antibiotic as shown in (C) in table 3.1.

Table 3.1 Examples of the three defined patterns among collected samples based on variation in their resistance profile. Three separate resistance profile patterns were obtained based on variation in antibacterial susceptibility between isolated colonies within each sample. (1) represents an identical resistance profile pattern and was defined when all examined colonies displayed an identical susceptibility pattern for the examined antibiotics. (2) represents low diversity resistance profile patterns. These patterns were defined when examined colonies varied in their susceptibility between resistances (R) to intermediate (I) zones of inhibition, or sensitive (S) to intermediate (I) zones of inhibition, for an examined antibiotic. Finally, (3) represents highly diverse resistance profile patterns. Highly diverse patterns were defined when the examined colonies varied in their susceptibility between sensitive (S) and resistance (R) zones of inhibition for a single antibiotic. Zone of inhibition measurements were interpreted based on BSAC, 2013.

Sample number	Number of colonies have been examined	Number of different resistance profile	Number of colonies that have the same pattern	Aminoglycoside	Penicillin	Cephalosporin			Carbapenems	Quinolones	Miscellaneous
				GM	AP	CPM	CPD	CAZ	MM	CIP	TM
(1) UTI (M19)	48 colonies	1 profile	48	R	R	R	R	R	S	R	R
(2) UTI (F8)	48 colonies	5 profiles	20	S	S	S	S	S	S	S	R
			19	I	S	S	S	S	S	S	R
			7	I	S	I	S	S	S	S	R
			1	I	S	S	S	I	S	S	R
			1	I	S	I	S	I	S	S	R
(3) UTI (F3)	48 colonies	3 profiles	45	R	R	R	R	S	S	R	R
			2	R	R	R	S	S	S	R	R
			1	R	S	I	S	S	S	R	R

3.2.2 Percentages of resistance profile patterns of collected samples

The highly diverse resistance profile patterns accounted for 21 % of collected samples and most of these samples were from female patients. The low diverse resistance profile patterns accounted for 66.67 % and the majority of these samples were from male patients. The last pattern is the identical resistance profile and it accounted for 11.9 % of collected samples. The majority of collected patient samples belong to the low diverse resistance patterns and the minority to the identical resistance profile.

Figure 3.1 shows the number of samples belonging to the various resistance patterns (high, low and identical) for all collected female and male patient samples. Highly diverse resistance pattern samples accounted for 29 % (6 samples) of all studied female patient samples and for 14.29 % (3 samples) of all studied male patient samples. Furthermore, 66.67 % (28 samples) of collected samples showed low-level diversity to antimicrobials of which 46.4% were from female patients and 53.57 % from male patients. The remaining five samples were found with colonies that had identical patterns of antibiotic susceptibility. Two female samples and three male samples are classified as having an identical resistance profile pattern. Together these data show that female UTI samples are slightly more varied in their sensitivity profiles than male samples, particularly with respect to highly diverse resistance profile samples. With respect to male patients, an identical number of samples (3 patients) exhibited highly diverse resistance profile patterns and identical resistance patterns.

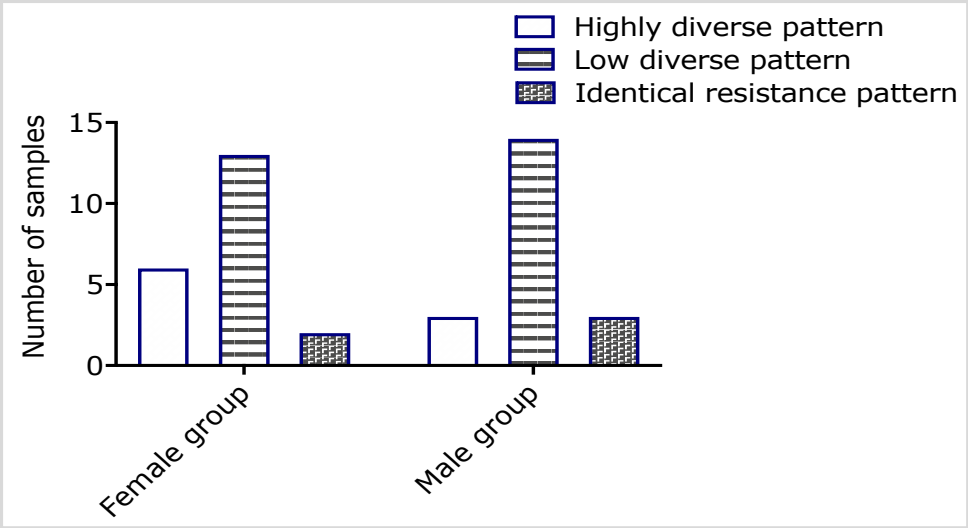


Figure 3.1 The percentage of varied resistance profiles patterns among collected female and male patients samples. The bar charts represent number of samples within high, low and identical resistance profile patterns of female and male groups. The highest number of samples was shown in low diverse resistance profile patterns followed by highly diverse resistance profile patterns and the lowest number were determined in the identical resistance profile pattern.

3.2.3 Highly diverse resistance profile patterns

Table 3.2 and table 3.3 show the antibiotic resistance profile patterns obtained within the six female and three male samples exhibiting high diversity. Three of these female samples and one male sample contain an *E. coli* classified as multi-drug resistant, as shown in table 3.2 and 3.3, due to the isolates being resistant to three or more classes of antibiotics.

Variation between sensitivity and resistance was observed for all antibiotics except Meropenem as shown in figure 3.2. The highest levels of variation between sensitive and resistant bacterial isolates within samples were seen with Ampicillin followed by Gentamicin and Cefpodoxime. For female patients' samples, apart from Ciprofloxacin and Meropenem, variation between sensitive and resistant bacterial isolates was seen in all other examined antibiotics. For male patients' samples, variation in resistance profiles between sensitivity and resistance was seen in Gentamicin, Ampicillin, Ciprofloxacin and Trimethoprim.

Table 3.2 Female highly diverse resistance profile patterns samples. The highly diverse resistance profile patterns among female samples including the number of examined colonies within each sample, number of colonies that have a similar resistance patterns and the resistance profile for all examined colonies. The highlighted samples represent multidrug resistance samples. Patterns with (R) indicate that the bacterial isolate is resistant to the examined antibiotic. Similarly, (I) indicates that the bacterial isolate has an intermediate zone of inhibition to the examined antibiotic and (S) indicates that the bacterial isolate is sensitive to the examined antibiotic. Zones of inhibition measurements were interpreted based on BSAC, 2013.

Sample number	Number of colonies have been examined	Number of different resistance profile	Number of colonies that have the same patterns	Aminoglycoside	Penicillin	Cephalosporin			Carbapenems	Quinolones	Miscellaneous
				GM	AP	CPM	CPD	CAZ	MM	CIP	TM
(A) UTI (F1)	26 colonies	4 profiles	15	S	R	S	S	S	S	S	R
			9	S	S	S	S	S	S	S	R
			1	S	R	S	S	S	S	S	S
			1	S	R	S	S	S	S	S	I
(B) UTI (F2)	40 colonies	9 profiles	17	I	R	I	S	I	S	R	R
			7	I	R	I	S	I	I	R	R
			5	R	R	R	S	I	I	R	R
			4	S	R	S	S	S	S	R	R
			2	R	R	I	S	I	I	R	R
			2	R	R	R	S	I	S	R	R
			1	R	R	I	S	R	I	R	R
			1	R	R	R	R	R	I	R	R
			1	I	R	I	S	S	S	R	R
(C) UTI (F3)	48 colonies	3 profiles	45	R	R	R	R	S	S	R	R
			2	R	R	R	S	S	S	R	R
			1	R	S	I	S	S	S	R	R
(D) UTI (F4)	22 colonies	3 profiles	17	I	R	S	S	S	S	S	S
			4	S	R	S	S	S	S	S	S
			1	S	S	S	S	S	S	S	S

Multiple resistance profiles present in similar morphological colonies within a single Uropathogenic *Escherichia coli* patient

Sample number	Number of colonies have been examined	Number of different resistance profile	Number of colonies that have the same patterns	Aminoglycoside	Penicillin	Cephalosporin			Carbapenems	Quinolones	Miscellaneous
				GM	AP	CPM	CPD	CAZ	MM	CIP	TM
(E) UTI (F5)	48 colonies	9 profiles	16	S	R	I	S	S	S	S	S
			14	S	R	I	S	I	S	S	S
			9	I	R	R	S	I	S	S	S
			2	I	R	R	S	R	I	S	S
			2	I	R	R	S	R	I	S	S
			2	I	R	R	S	I	I	S	S
			1	S	R	R	S	I	S	S	S
			1	S	R	R	S	R	I	S	S
			1	I	R	I	S	I	S	S	S
			1	I	R	I	S	S	S	S	S
(F) UTI (F6)	40 colonies	5 profiles	35	R	R	R	R	S	S	R	R
			2	R	R	I	S	S	S	R	R
			1	R	R	R	R	I	S	R	R
			1	R	R	I	R	S	S	R	R
			1	S	R	R	S	S	S	R	R

Table 3.3 Male Highly diverse resistance profile patterns samples. The highly diverse resistance profile patterns among male samples including the number of examined colonies within each sample, number of colonies that have a similar resistance patterns and the resistance profile for all examined colonies. The highlighted samples represent multidrug resistance samples. Patterns with (R) indicate that the bacterial isolate is resistant to the examined antibiotic. Similarly, (I) indicates that the bacterial isolate has an intermediate zone of inhibition to the examined antibiotic and (S) indicates that the bacterial isolate is sensitive to the examined antibiotic. Zones of inhibition measurements were interpreted based on BSAC, 2013.

Sample number	Number of colonies have been examined	Number of different resistance profile	Number of colonies that have the same patterns	Aminoglycoside	Penicillin	Cephalosporin			Carbapenems	Quinolones	Miscellaneous
				GM	AP	CPM	CPD	CAZ	MM	CIP	TM
(H) UTI (M1)	36 colonies	11 profiles	8	I	R	I	S	S	S	I	R
			7	S	R	S	S	S	S	S	R
			7	S	R	S	S	S	S	I	R
			6	I	R	S	S	S	S	I	R
			2	I	R	I	S	I	S	I	R
			1	I	R	S	S	S	S	S	R
			1	I	R	S	S	I	I	I	R
			1	I	R	I	S	I	I	I	R
			1	I	R	I	S	S	S	S	R
			1	S	R	I	S	I	S	I	R
			1	I	R	I	S	I	S	R	R
(I) UTI (M2)	48 colonies	6 profiles	33	S	R	S	S	S	S	S	R
			7	I	R	S	S	S	S	S	R
			3	I	R	I	S	S	S	S	R
			3	I	R	I	S	I	S	S	R
			1	S	R	I	S	S	S	S	R
			1	S	S	S	S	S	S	S	R
(G) UTI (M3)	40 colonies	2 profiles	39	R	R	R	R	R	S	R	R
			1	S	R	R	R	R	S	R	S

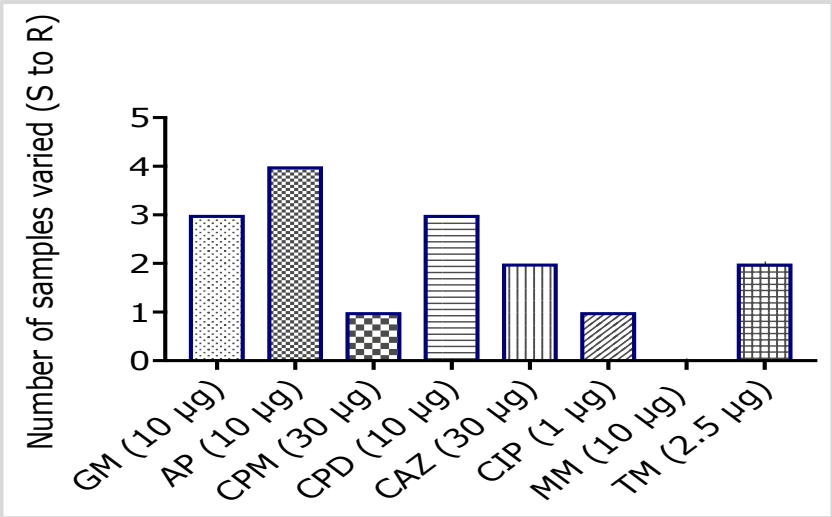


Figure 3.2 Number of samples that varied in resistance among examined antibiotics between sensitive and resistance. Number of samples that changed in their resistance profiles between sensitive (S) to resistant (R) for an examined antibiotics of all collected highly diverse samples. Meropenem is the only antibiotic that does not show any resistance zone of inhibition. Gentamicin (GM), Ampicillin (AP), Cefepime (CPM), Cefpodoxime (CPD), Ceftazidime (CAZ), Ciprofloxacin (CIP), Meropenem (MM) and Trimethoprim (TM).

3.2.4 Low diverse and identical resistance patterns samples

The low diversity resistance profile pattern was the most commonly observed profile in samples from both males and females. In this resistance profile, there is a change from either resistant or sensitive to an intermediate level of resistance between the examined colonies. In the female low diverse resistance profile patterns (table 3.4), the examined colonies within samples varied in their susceptibilities to Gentamicin, Cefepime, Ceftazidime, Meropenem, Ciprofloxacin and Trimethoprim. Whereas male patient samples (table 3.5) show a variation between isolated colonies in all examined antibiotics apart from Ampicillin, Cefpodoxime and Ciprofloxacin. The highest level of variation among collected samples was noticed in Gentamicin and was distributed equally between female and male patients. This was followed by Cefepime and Ceftazidime, respectively (Figure 3.3). The lowest level of variation was in Ciprofloxacin, seen only in two female samples.

Five samples were categorised as having identical resistance profile patterns (table 3.6). Only three male patient samples (14.29 %) contained an identical pattern, in comparison with two female patients (10 %). All but one (male sample) of the identical resistance profile pattern samples are categorised as multi-drug resistant. One female sample bacterial isolate is resistant to all examined antibiotics except ciprofloxacin. However, the male samples are resistant to all examined antibiotics apart from Meropenem. The remaining one male sample is sensitive to all examined antibiotics.

Table 3.4 Female low diverse resistance profile patterns samples. Low diverse resistance profile patterns among female samples including the number of examined colonies within each sample, number of colonies that have a similar resistance patterns and the resistance profile for all examined colonies. The highlighted samples represent multidrug resistance samples Patterns with (R) indicate that the bacterial isolate is resistant to the examined antibiotic. Similarly, (I) indicates that the bacterial isolate has an intermediate zone of inhibition to the examined antibiotic and (S) indicates that the bacterial isolate is sensitive to the examined antibiotic. Zones of inhibition measurements were interpreted based on BSAC, 2013.

Sample number	Number of colonies have been examined	Number of different resistance profile	Number of colonies that have the same patterns	Aminoglycoside	Penicillin	Cephalosporin			Carbapenems	Quinolones	Miscellaneous
				GM	AP	CPM	CPD	CAZ	MM	CIP	TM
UTI (F7)	32 colonies	5 profiles	25	S	R	S	S	S	S	S	R
			4	I	R	S	S	S	S	S	R
			1	I	R	I	S	I	I	S	R
			1	I	R	I	S	S	I	S	R
			1	S	R	I	S	S	I	S	R
UTI (F8)	48 colonies	5 profiles	20	S	S	S	S	S	S	S	R
			19	I	S	S	S	S	S	S	R
			7	I	S	I	S	S	S	S	R
			1	I	S	S	S	I	S	S	R
			1	I	S	I	S	I	S	S	R
UTI (F9)	47 colonies	4 profiles	18	S	S	S	S	S	S	S	R
			15	S	S	S	S	S	S	I	R
			8	S	S	S	S	S	S	S	I
			6	S	S	S	S	S	S	I	I
UTI (F10)	48 colonies	4 profiles	41	S	S	S	S	S	S	S	S
			5	I	S	S	S	S	S	S	S
			1	I	S	I	S	S	S	S	S
			1	S	S	I	S	S	S	S	S
UTI (F11)	48 colonies	4 profiles	22	I	R	S	S	S	S	S	R
			17	I	R	I	S	S	S	S	R
			5	S	R	S	S	S	S	S	R
			4	I	R	I	S	I	S	S	R
UTI (F12)	32 colonies	3 profiles	21	S	R	R	S	S	S	S	R
			9	S	R	I	S	S	S	S	R
			2	S	R	R	S	I	S	S	R

Multiple resistance profiles present in similar morphological colonies within a single Uropathogenic *Escherichia coli* patient

Sample number	Number of colonies have been examined	Number of different resistance profile	Number of colonies that have the same patterns	Aminoglycoside	Penicillin	Cephalosporin			Carbapenems	Quinolones	Miscellaneous
				GM	AP	CPM	CPD	CAZ	MM	CIP	TM
UTI (F13)	48 colonies	3 profiles	44	S	S	S	S	S	S	S	S
			3	I	S	S	S	S	S	S	S
			1	I	S	S	S	S	S	S	I
UTI (F14)	38 colonies	3 profiles	26	I	R	S	R	R	S	R	R
			8	I	R	I	R	R	S	R	R
			4	S	R	S	R	R	S	R	R
UTI (F15)	28 colonies	3 profiles	13	S	R	S	S	S	S	S	R
			8	I	R	S	S	S	S	S	R
			7	I	R	I	S	S	S	S	R
UTI (F16)	48 colonies	2 profiles	47	S	R	S	S	S	S	S	R
			1	I	R	S	S	S	S	S	R
UTI (F17)	36 colonies	2 profiles	32	S	S	S	S	S	S	I	R
			4	S	S	S	S	S	S	S	R
UTI (F18)	48 colonies	2 profiles	47	S	R	S	S	S	S	S	S
			1	I	R	S	S	S	S	S	S
UTI (F19)	48 colonies	2 profiles	32	R	R	R	R	R	S	R	R
			16	I	R	R	R	R	S	R	R

Table 3.5 Male low diverse resistance profile patterns samples. The low diverse resistance profile patterns among male samples including the number of examined colonies within each sample, number of colonies that have a similar resistance patterns and the resistance profile for all examined colonies. The highlighted samples represent multidrug resistance samples. Patterns with (R) indicate that the bacterial isolate is resistant to the examined antibiotic. Similarly, (I) indicates that the bacterial isolate has an intermediate zone of inhibition to the examined antibiotic and (S) indicates that the bacterial isolate is sensitive to the examined antibiotic. Zones of inhibition measurements were interpreted based on BSAC, 2013.

Sample number	Number of colonies have been examined	Number of different resistance profile	Number of colonies that have the same patterns	Aminoglycoside	Penicillin	Cephalosporin			Carbapenems	Quinolones	Miscellaneous
				GM	AP	CPM	CPD	CAZ	MM	CIP	TM
UTI (M4)	36 colonies	7 profiles	13	I	S	S	S	S	S	S	S
			6	S	S	S	S	S	S	S	S
			5	I	S	I	S	I	S	S	S
			5	I	S	S	S	I	S	S	S
			4	I	S	I	S	S	S	S	S
			2	I	S	I	S	I	S	S	I
			1	I	S	I	S	I	I	S	I
UTI (M5)	48 colonies	6 profiles	27	S	S	S	S	S	S	S	S
			13	I	S	S	S	S	S	S	S
			3	I	S	I	S	S	S	S	S
			2	S	S	I	S	S	S	S	S
			2	S	S	S	S	S	I	S	S
			1	I	S	I	S	I	S	S	S
UTI (M6)	48 colonies	5 profiles	39	S	R	S	S	S	S	S	R
			4	I	R	S	S	S	S	S	R
			2	S	R	I	S	S	S	S	R
			2	I	R	I	S	S	S	S	R
			1	I	R	I	S	I	S	S	R
UTI (M7)	18 colonies	4 profiles	9	I	S	I	S	I	S	R	R
			6	S	S	S	S	S	S	R	R
			2	S	S	S	S	I	S	R	R
			1	S	S	I	S	S	S	R	R
UTI (M8)	48 colonies	3 profiles	34	S	S	S	S	S	S	S	S
			11	I	S	S	S	S	S	S	S
			3	I	S	I	S	S	S	S	S
UTI (M9)	48 colonies	2 profiles	34	R	R	R	R	R	S	R	R
			14	R	R	R	R	R	I	R	R

Multiple resistance profiles present in similar morphological colonies within a single Uropathogenic *Escherichia coli* patient

Sample number	Number of colonies have been examined	Number of different resistance profile	Number of colonies that have the same pattern	Aminoglycoside	Penicillin	Cephalosporin			Carbapenems	Quinolones	Miscellaneous
				GM	AP	CPM	CPD	CAZ	MM	CIP	TM
UTI (M10)	48 colonies	2 profiles	26	R	R	S	S	S	S	S	R
			2	R	R	I	S	S	S	S	R
UTI (M11)	40 colonies	2 profiles	32	R	R	R	R	R	S	R	R
			8	R	R	R	R	R	I	R	R
UTI (M12)	40 colonies	2 profiles	28	S	S	S	S	S	S	S	S
			12	I	S	S	S	S	S	S	S
UTI (M13)	32 colonies	2 profiles	31	I	R	S	S	S	S	S	S
			1	S	R	S	S	S	S	S	S
UTI (M14)	48 colonies	2 profiles	47	S	S	S	S	S	S	S	S
			1	I	S	S	S	S	S	S	S
UTI (M15)	48 colonies	2 profiles	47	S	S	S	S	S	S	S	S
			1	I	S	S	S	S	S	S	S
UTI (M16)	48 colonies	2 profiles	38	I	R	S	S	S	S	S	S
			10	S	R	S	S	S	S	S	S
UTI (M17)	28 colonies	2 profiles	25	I	R	I	S	S	S	S	R
			3	S	R	I	S	S	S	S	R
UTI (M18)	32 colonies	2 profiles	31	S	S	S	S	S	S	S	S
			1	S	S	I	S	S	S	S	S

Table 3.6 Female and Male identical resistance profile patterns samples. The identical resistance profile pattern among female and male samples including the number of examined colonies within each sample, number of colonies that have a similar resistance patterns and the resistance profile for all examined colonies. The highlighted samples represent multidrug resistance samples. Patterns with (R) indicate that the bacterial isolate is resistant to the examined antibiotic. Similarly, (I) indicates that the bacterial isolate has an intermediate zone of inhibition to the examined antibiotic and (S) indicates that the bacterial isolate is sensitive to the examined antibiotic. Zones of inhibition measurements were interpreted based on BSAC, 2013.

Sample number	Number of colonies have been examined	Number of different resistance profile	Number of colonies that have the same patterns	Aminoglycoside	Penicillin	Cephalosporin			Carbapenems	Quinolones	Miscellaneous
				GM	AP	CPM	CPD	CAZ	MM	CIP	TM
Female samples											
(J) UTI (F20)	40 colonies	1 profile	40	R	R	R	R	R	R	S	R
(K) UTI (F21)	22 colonies	1 profile	22	R	R	S	S	S	S	S	R
Male samples											
(L) UTI (M19)	48 colonies	1 profile	48	R	R	R	R	R	S	R	R
(M) UTI (M20)	32 colonies	1 profile	32	S	S	S	S	S	S	S	S
(N) UTI (M21)	44 colonies	1 profile	44	R	R	R	R	R	S	R	R

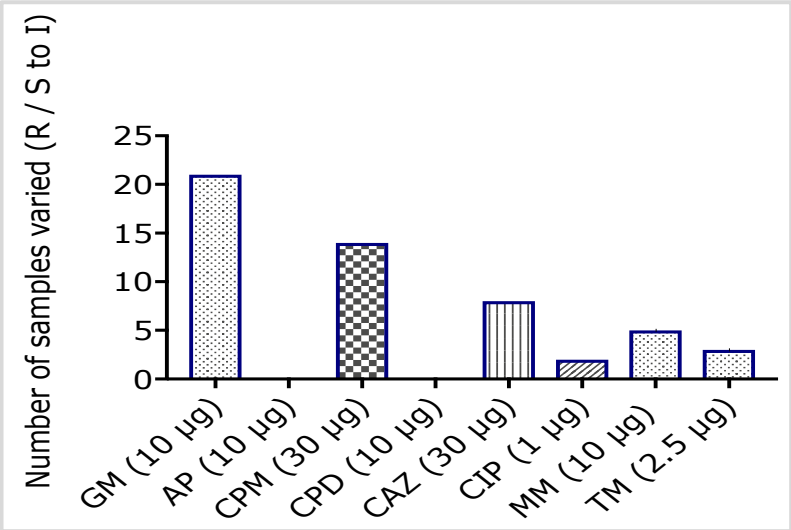


Figure 3.3 Number of samples that varied in susceptibility among examined antibiotics. The figure represents number of samples that changed to intermediate (I) zone of inhibition from either sensitive (S) or resistance (R) for examined antibiotics of all collected low diverse samples. Ampicillin and Cefpodxime antibiotics are not included in this pattern**.** Gentamicin (GM), Ampicillin (AP), Cefepime (CPM), Cefpodoxime (CPD), Ceftazidime (CAZ),, Ciprofloxacin (CIP), Meropenem (MM) and Trimethoprim (TM).

3.3 Discussion

A study published in 2013 reported the presence of two sequence types of UPEC within a single patient with urinary tract infection (McNally et al., 2013). A body of literature proposes that bacterial populations within a single host may differ (Sieradzke et al., 2003; Mwangi et al., 2007; McAdam et al., 2011; Lieberman et al., 2011) and existence of genetic variation between bacterial isolates from clinical infections has been proven (Sieradzke et al., 2003; Mwangi et al., 2007). In addition, multiclonal infections have also been identified in some cases (McAdam et al., 2011; Lieberman et al., 2011). We would like to determine the entire of diversity within a single patient infected with UPEC.

Based on the rationale that variation in susceptibility profiles may be a sign of bacterial within-host diversity, in our research we have applied the disk diffusion method on 48 isolated colonies from patients infected with UPEC to assess population diversity levels. The experiment was performed on forty-two samples. The samples were then classified based on the obtained resistance pattern into three groups (table 3.1). The observed variation in the resistance profiles of our collected UPEC samples could be linked with the phenomenon termed heteroresistance, when the susceptibility profiles within a single population exhibit variation (El-Halfawy and Valvano, 2015). There are two possibilities that may explain the presence of heteoresistance in our samples.

Having hetroresistance within bacterial populations could be due to the presence of polyclonal diversity, where two or more sequence types of UPEC are present within a single patient. This has been shown previously in patients infected with multiple *E.*

coli sequence types. Levert and co-workers (2010) reported that 21 % of patients infected with ExPEC could be categorised as having a polyclonal infection. UPEC is a heterogeneous group of organisms comprised of multiple different sequence types that possess varied resistance profiles, and so it is possible that we have multiple sequence types within our highly diverse resistance profiles group. As an example, ST95 is a highly susceptible isolate whereas ST131 generally multi-drug resistant (Gibreel et al., 2012).

Alternatively, development of heteroresistance within our samples may also be a result of single nucleotide polymorphisms or insertions and deletions, leading to monoclonal diversity. This situation is increasingly being reported in recurrent infections and chronic infections (El-Halfawy and Valvano, 2015). A further study on patients infected with *H. pylori* reported that the presence of heteroresistant isolates might be due to highly dynamic DNA, with resistant colonies in that study thought to be derived from a single pre-existing strain rather than infection with multiple bacterial strains (Kao at al., 2014a). In addition, host diversity was examined in patients infected with ExPEC by Levert et al. (2010) and it was found that most patients could be categorised as demonstrating monoclonal diversity due to microheterogeneity of mutations. Furthermore, prolonged infection can drive sensitive *E. coli* to become antibiotic resistant through the gain of mutations, such as occurs for gentamicin resistance (El-Halfawy and Valvano, 2015). This possibility could be linked with the variation observed in our samples that display highly diverse and low diversity resistance profiles.

We suspect that the highly diverse resistance profiles most likely originated due to infection with multiple sequence types of UPEC and we hypothesise this will be accompanied by the presence of other phenotypic and genetic variations. Furthermore, we presume that the variation observed in our low diversity resistance patterns originates from a monoclonal infection showing low-level mutation and phenotypic diversity. In total, 11.9 % of our samples were classified as possessing an identical pattern and can therefore be considered a homogenous population.

The highly diverse resistance profile was observed in 66 % of samples came from female patients, making this sex appear to be more susceptible to population diversity. However, this could be linked to the fact that women are more susceptible to urinary tract infections (Foxman, 2003), as females have a shorter urethra that allows uropathogens to access the bladder more readily (Ronald, 2003). In addition, females are more susceptible to recurrent urinary tract infections, and prolonged vaginal colonisation with uropathogenic *E. coli* has been reported (Ronald, 2003; Hooton, 2012). Our female patients may have an infection with mixed strains of UPEC or have a recurrent infection with the same strain which is accumulating genetic variation over time. As reported by Hooton (2012) two thirds of women having a previous history of UTIs become reinfected with the same bacterial strain causing the initial infection (Hooton, 2012; El-Halfawy and Valvano, 2015). The clinical history of the patient, of whom the samples were taken, may have provided further details about the collected samples, however, this information was not available. Additionally, due to the smaller sample size used within the investigation,

it cannot be generalised that female patients are more susceptible to the hetero resistance phenomenon.

Variations in antibiotic resistance were mostly seen for ampicillin and gentamicin, followed by cephalosporins and trimethoprim. These classes of antibiotics have seen an increase in resistance in recent years, and resistance to β-lactams and aminoglycosides has been reported to be between 20 % and 45 % according to surveillance data from Europe and North and South America (Pitout and Laupland, 2008; Foxman, 2010). Therefore high variation in ampicillin and gentamicin resistance could be supported by the overall increase in resistance to these antibiotics.

Overall, most classical diagnoses performed in a clinical laboratory are carried out based on the assumption that a single colony from a homologous bacterial culture is representative of an entire bacterial population (Pappas, 1991). This was supported by a study published in 2014, Willner and co-workers examined urinary tract isolates and reported that single isolates were representative of urinary tract infections at the genus and strain levels. In their research 50 samples were examined in order to determine the genus level of urinary tract infection communities, while examination at the strain level was performed for 27 *E. coli* samples (Willner et al., 2014). DNA was also extracted from urine samples and single cultured colonies from eight patients and metagenomics analysis conducted. However, our research was based on comparing isolated colonies from pure UPEC cultures from patients with UTIs at a single time point in order to determine population diversity level. This approach was previously taken for *P. aeruginosa* to determine the level of phenotypic and

genotypic variation (Darch et al., 2015). In our research, a different experimental design for phenotypically comparing colonies from pure UPEC samples was applied compared to those utilised by Willner et al. (2014). The obtained variation in susceptibility testing provided a prediction of the diversity of our samples and allowed the next stage of this research to be designed. The results of the following chapters will expand our knowledge about strain-level diversity by either confirming that single isolates are representative of an entire bacterial population or proving that there is diversity at the strain level and estimating the level of this diversity. Further, if there is a potential for the presence of a missed bacterial isolate during diagnosis, which could have implications on providing the most efficient treatment and infection persistence (Foweraker et al., 2005). This may impact the treatment efficacy of urinary tract infections and lead to further complications.

Chapter Four

Phenotypic diversity within apparently identical isolates of Uropathogenic *Escherichia coli* from a single patient

4.1 Introduction

4.1.1 Phenotypic diversity of Uropathogenic *E. coli* isolates

Strains of UPEC are primarily confined to four defined phylogenetic groups, namely A, B_1, B_2, and D, with the most virulent isolates belonging to groups B_2 and D (Boyd and Hartl, 1998). UPEC isolates can vary phenotypically on the basis of the virulence factors that they possess (Johnson, 1991). A study performed on 72 ExPEC strains determined that the variable virulence factors within the phylogenetic groups were not unique to those specific groups (Johnson et al., 2001). Within isolates from individual patients, phenotypic diversity was demonstrated in eight patients infected with ExPEC. These isolates exhibited wide variation in antibiotic resistance, outer-membrane permeability, growth rate under standard conditions, and virulence properties (Levert et al., 2010).

4.1.2 Role of motility in *E. coli* uropathogenesis

Motility in flagellated strains of UPEC plays a role in efficient colonisation and migration of infection to the upper urinary tract, and also in the spread of uropathogens to the blood stream, resulting in sepsis (Simms and Mobley, 2008; Lane et al., 2005; Lane et al., 2007). Flagellated UPEC isolates are responsible for 70 % – 90 % of urinary-tract infections (Bien et al., 2012). UPEC isolates vary in their motility, and can be classified as either hypo-motile or hyper-motile (Kao et al., 2014b). As an illustration of that, isolates from the urine of pyelonephritis patients are less motile than isolates from the urine of cystitis patients (Herrmann and Burman, 1985).

Flagella organelles consist of flagellin, which can be composed of one of 56 different proteins, the H antigens (Erdem et al., 2007), and these proteins are thought to have

a role in physiological mechanisms such as adhesion, invasion and biofilm formation. Parthasarathy et al. (2007) suggested that the flagella of meningitis-associated *E. coli* could mediate the invasion of brain epithelium. In addition, H6 and H7 flagellin antigens in Enteropathogenic and Enterohaemorrhagic *E. coli* are involved in adhesion to colon epithelium cells (Erdem et al., 2007). Furthermore, motility is thought to be related to biofilm formation. Wood et al. (2006) demonstrated, in a comparative study among eight *E. coli* strains, that differences in flagella expression lead to differences in biofilm architecture. Also, deletion of *fliA* or *qseB* genes not only affected cell attachment and bacterial motility, but also decreased biofilm formation.

4.1.3 Role of biofilm formation in uropathogens

Invasion by UPEC of superficial urinary tract cells is the first stage in the formation of intracellular bacterial communities (Wright et al., 2005). These communities are highly organised multicellular adherent colonies and are surrounded by a matrix (Traunter and Darouiche, 2004). The biofilm matrix is formed of a variety of molecules such as proteins, exopolysaccharides and nucleic acids (Branda et al., 2005). During biofilm formation, bacterial cells are surrounded by a thick layer of exopolysaccharides, which is responsible for non-specific adhesion, and fimbriae which protrude from the bacterial cell surface and are responsible for specific adherence to the host surface. Once bacterial cells adhere to the surface, biofilm formation can take place (Costerton, 1999). Other factors such as the curli fimbriae, F conjugative pilus, and flagella are also involved in the formation of a mature biofilm. The curli fimbriae and F conjugative pilus are involved in cell-surface interaction, and biofilm stabilisation and maturation, respectively (Soto et al., 2011).

Biofilm formation protects the pathogen from antibiotic treatment and phagocytic cells, and allows the pathogen to survive within the host (Traunter and Darouiche, 2004). Ponnuusamy et al. (2012) reported that there is a significant relationship between biofilm formation and antibiotic resistance. In that study, biofilm-producing Uropathogenic *E. coli* isolates were recognised to be less sensitive to ampicillin, β-lactam and aminoglycoside antibiotics. Biofilm formation may also be the main determinant of recurrent urinary tract infection (Soto et al., 2011) and is thought to be the factor that makes the infection persist (Subashchandrabose and Mobley, 2015). In addition, Uropathogenic *E.coli* cystitis isolates are highly motile and seem to be very efficient at biofilm formation (Herrmann and Burman, 1985; Tabasi et al., 2015).

4.1.4 Role of association and invasion in uropathogens

UPEC attachment is a key factor in urinary tract colonisation, infection development, and persistence. To establish a urinary tract infection, UPEC need to overcome several factors such as urine flow, host immunity, and antimicrobial treatment, and adherence of UPEC to the urinary tract epithelium may help to overcome those obstacles (Justice et al., 2004). The ability of UPEC to bind to host cells by adhesive organelles is the main determinant in urinary tract pathogenicity and promotes host cell invasion (Mulvey et al., 2001). There are many types of pilus involved in the adhesion of UPEC to the host epithelial cells. Type 1 pili, P pili, S pili and the Dr adhesin family have all been found to be involved in attachment and invasion (Johnson, 1991).

Type 1 pili are encoded by most UPEC isolates (Russell and Orndorff, 1992; Jones et al., 1995). They are composed of a helical rod of *Fim* A - repeated subunits, with a

wide tip containing the adhesive protein *Fim* H and two adaptor proteins, F and G (Russell and Orndorff, 1992; Jones et al., 1995). Attachment to bladder epithelium occurs via binding of FimH to mannose-containing glycoprotein receptors on bladder cell surfaces (Mulvey et al., 2001). Type 1 fimbriae are not only expressed by the most common UPEC, but their role in biofilm formation has also been shown (Pratt and Kolter, 1998; Schembri et al., 2001). Type 1 fimbriae also have a role in uroepithelial colonisation, and they help in bacterial attachment to the bladder epithelium and subsequent invasion (Sokurenko et al., 1998; Jacobsen et al., 2008). P pili adhesion organelles are found commonly in Uropathogenic *E. coli*. P pili are composed of repeating *Pap* A subunits and a tip which is composed of *Pap* – G, E, F, and K (Mulvey, 2002). P pili are thought to have a role in the early colonisation of the urinary tract (Orskov et al., 1982). Pap G is highly efficient at binding with glycolipids on erythrocytes and the host kidney (Leffler and Svanbrog-Eden, 1980; Lund et al., 1987) and is often present in isolates from pyelonephritis patients (Roberts et al., 1994).

S pili are another type of adhesive organelle which are expressed by UPEC. These pili have a role in the interaction between UPEC and kidney epithelium and vascular cells. S pili are composed of *Sfa* A, the major subunit, and three smaller subunits, *Sfa* G, *SfaH* and *SfaS* (Schmoll et al., 1989), and they attach to sialosyloligosaccharide (Goldhar, 1996). S pili are highly expressed in sepsis and meningitis isolates, and seem to have a role in bacterial dissemination (Mulvey, 2002). Dr family adhesins are another group of adhesive organelles composed of adhesive fimbriae *AFA* - I and *AFA* - II (Goldhar, 1996). They are expressed in most recurrent UPEC (Korotkova et al.,

2008) and have a role in the ascension of uropathogens to the upper urinary tract (Goldhar, 1996).

4.1.5 Aim of the chapter

In chapter 3, three resistance level profiles were described based on variation in susceptibility results obtained from apparently identical colonies isolated from the same patient. It was suggested that those variations may be related to population diversity. This pattern of observing a diverse population among the isolates could reflect further phenotypic variation, especially where there was a highly diverse pattern. To determine the degree of phenotypic diversity within various resistance profile patterns additional phenotypic assays were performed on the isolates from within each sample exhibiting a variable resistance profile. The baseline standard level of phenotypic variation was determined via comparison of isolates from highly diverse samples with isolates from patients thought to have a homogenous population (isolates of an identical resistance profile pattern).

4.1.6 Bacterial isolates used in this chapter

Bacterial isolates (variants) were obtained from UPEC samples with varied resistance profiles as described in Chapter 3, were used in this chapter and described in table 4.1, In addition to the control cultures used in the phenotypic assays.

Phenotypic diversity within apparently identical isolates of Uropathogenic *Escherichia coli* from a single patient

Table 4.1 Bacterial strains used in this chapter

Highly diverse resistance profile pattern		
Sample ID	**Samples bacterial isolates (Variants)**	**Resistance profile**
A. UTI (F1)	Colony 2	Resistance to AP and TM
	Colony 6	Resistance to TM
	Colony 14	Resistance to AP
	Colony 18	Resistance to AP, Intermediate to TM
B. UTI (F2)	Colony 24	Resistance to AP, GM, CIP and TM
	Colony 30	Resistance to AP, CIP and TM
	Colony 31	Resistance to all except MM
C. UTI (F3)	Colony 3	Resistance to AP and CPM
	Colony 7	Sensitive to AP, Intermediate to CPM
D. UTI (F4)	Colony 2	Resistance to AP
	Colony 20	Sensitive to AP
E. UTI (F5)	Colony 19	Resistance to CAZ
	Colony 28	Sensitive to CAZ
F. UTI (F6)	Colony 2	Resistance to GM
	Colony 40	Sensitive to GM
G. UTI (M3)	Colony 5	Resistance to GM and TM
	Colony 35	Sensitive to GM and TM
H. UTI (M1)	Colony 2	Resistance to CIP
	Colony 25	Sensitive to CIP
I. UTI (M2)	Colony 18	Resistance to AP
	Colony 31	Sensitive to AP
Identical resistance profile pattern		
J. UTI (F20)	Five random colonies were selected assigned with (A,B.C,D and E)	Sensitive to MM only
K. UTI (F21)		Resistance to GM, AP and TM
L. UTI (M21)		Sensitive to MM only
M. UTI (M20)		Sensitive to all examined antibiotics
N. UTI (M19)		Sensitive to MM only

Controls used in this chapter		
Staphylococcus aureus	_	Motility assay
Salmonella enteritidis	_	Motility assay
***E. coli* K12**	_	Biofilm formation assay
***E. coli* DH5α**	_	Association and invasion assay
***E. coli* CFT07**	_	Association and invasion assay

4.1.7 Statistical analysis

Graph Pad Prism software (version 7 XML) was used to perform statistical analyses of the data. One-way ANOVA was applied to determine the statistical significance for motility assay. Multiple group comparison of two-way ANOVA was used to determine the statistical differences of biofilm formation and association and invasion assays between variants within each sample of highly diverse and identical resistance profile patterns.

4.2 Results

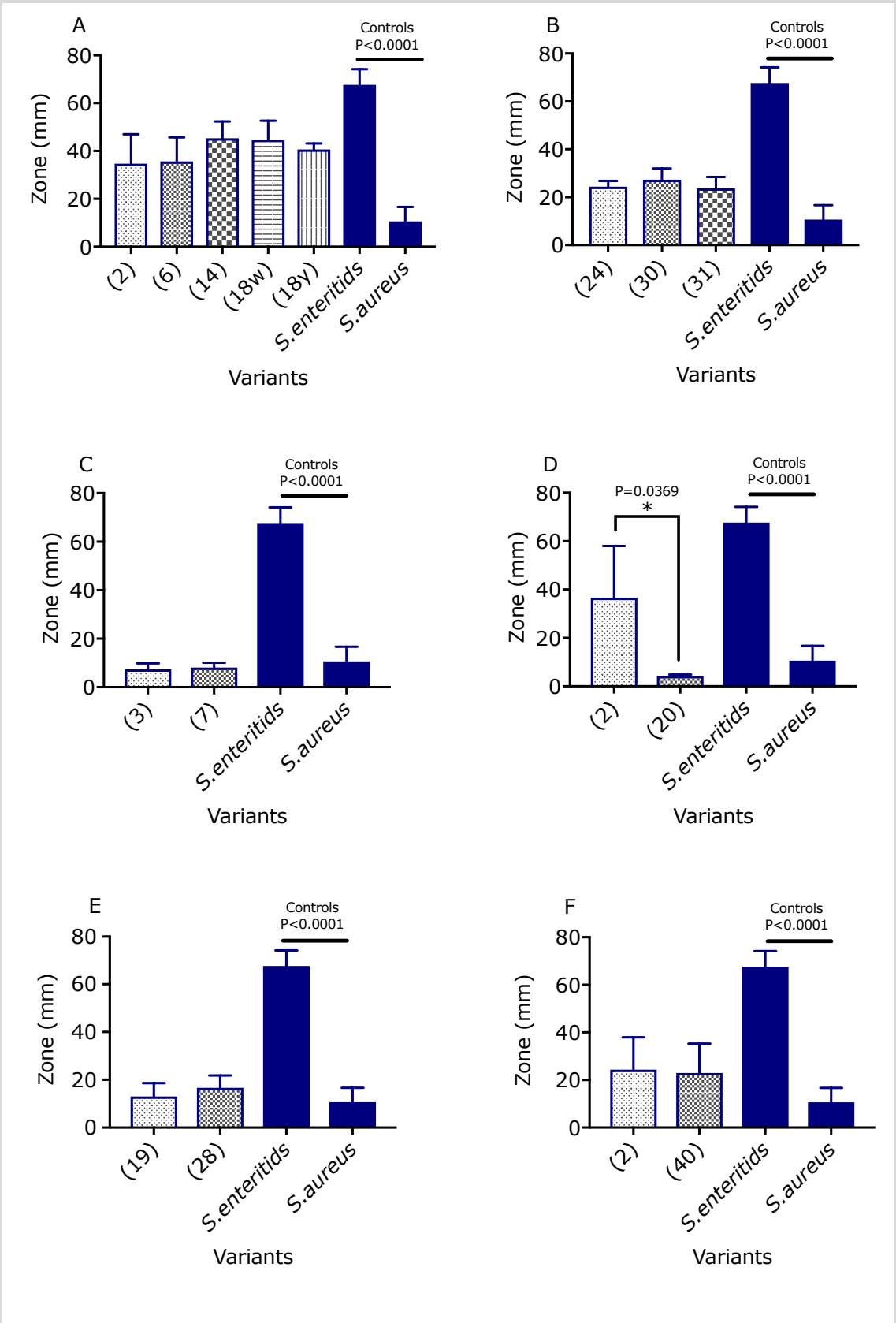
4.2.1 Motility ability among samples variants

A motility assay using semi-solid agar plates was performed on the highly diverse resistance profile variants and five apparently identical colonies of identical resistance profile, and the zone of migration measured and compared. Data from the latter group was taken to reflect the base line variation level of motility. Bar charts in figures 4.1, 4.2 and 4.3 represent the mean diameter of migration zones in mm. *S. aureu*s and *S. enteritidis* were used as negative and positive controls, respectively.

4.2.1.1 Highly diverse resistance profile sample variants

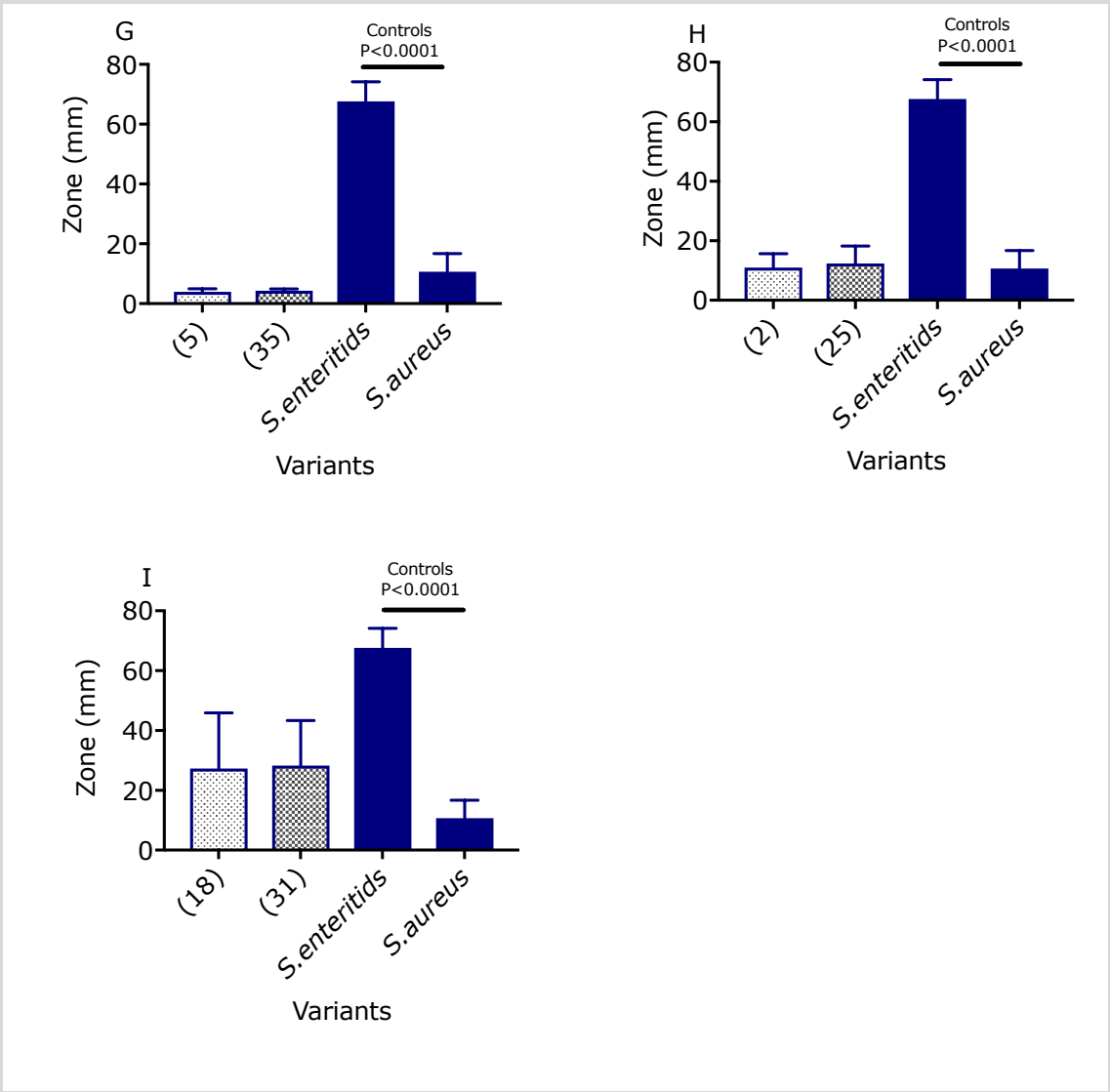
Figures 4.1 and 4.2 show the variation of motility within each isolate from highly diverse samples. Figure 4.1 shows the zone of migration for colonies from female samples. The only significant variation was noticed in sample D where isolates behaved differently in their motility (P value = 0.0369; F value = 18.61). The isolate assigned (2) from sample D is a motile isolate while the other isolate (20) from the same sample is non-motile. Isolates from the other female patient samples showed similar behaviour in their motility when comparisons were made for bacterial isolates obtained from the same sample. No significant differences in motility were found in isolates from male patients (figure 4.2). Overall, isolates from male patients were less motile than those from female patients. Apart from sample G, where the isolates were non-motile, all other isolates were motile. The overall motility pattern in all isolates was approximately > 20 mm as a migration zone. However, the migration zone of variants from sample C was about 7 mm and this defines them as hypo-motile isolates.

A. Highly diverse resistance profile female samples



Figures 4.1 Comparison of motile ability among bacterial isolates (variants) belong to female highly diverse resistance profile pattern samples. Motility assays, using semi-solid agar plates, were applied to six different female samples. Bacterial isolates of each sample were grown on semi-solid agar and incubated at 37 °C for 18 hours. The bacterial isolates of each sample were assigned with a number, each represented the variation between sensitivity and resistance to an antibacterial agent based on susceptibility testing. The corresponding bar chart has been listed with six letters, in order to represent each of the six female samples, as follows: (A) : UTI (F1), (B) : UTI (F2), (C) : UTI (F3), (D) : UTI (F4), (E) : UTI (F5) and (F) : UTI (F6). The bar charts display a representation of the mean of the migration zone across three independent experiments. Furthermore, they are based on three replicates with error bars to indicate the standard deviation of each variable. *S.enteritidis* and *S.aureus* were each used as a positive and negative control respectively, and were also visually represented through documented bar charts. The only significant is shown in sample (D) whereas sample bacterial isolates (variants) has a P value < 0.05.

B. Highly diverse resistance profile male samples

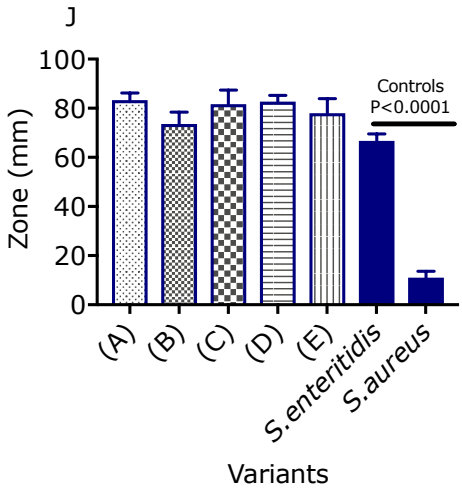


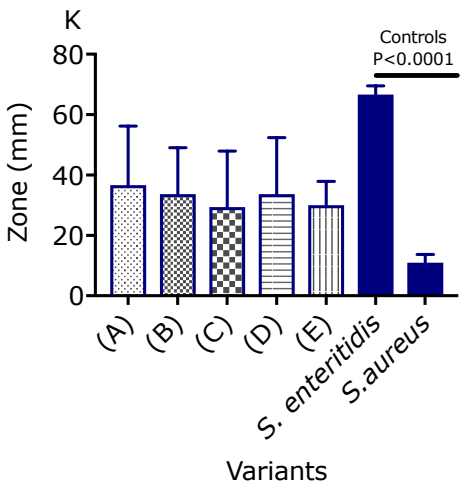
Figures 4.2 Comparison of motile ability between bacterial isolates (variants) belong to male high diverse resistance profile pattern samples. Motility assays, using semi-solid agar plates, were applied to three different male samples. Bacterial isolates of each sample were grown on semi-solid agar and incubated at 37 °C for 18 hours. The bacterial isolates of each sample were assigned with a number, each represented the variation between sensitivity and resistance to an antibacterial agent based on susceptibility testing. The corresponding bar chart has been listed with three letters, in order to represent each of the three male samples, as follows: (G) : UTI (M3), (H) : UTI (M1) and (I) : UTI (M2). The bar charts display a representation of the mean of the migration zone across three independent experiments. Furthermore, they are based on three replicates with error bars to indicate the standard deviation of each variable. *S.enteritidis* and *S.aureus* were each used as a positive and negative control respectively, and were also visually represented through documented bar charts. There is no significant difference in zone of migration in (mm) between variants of each sample.

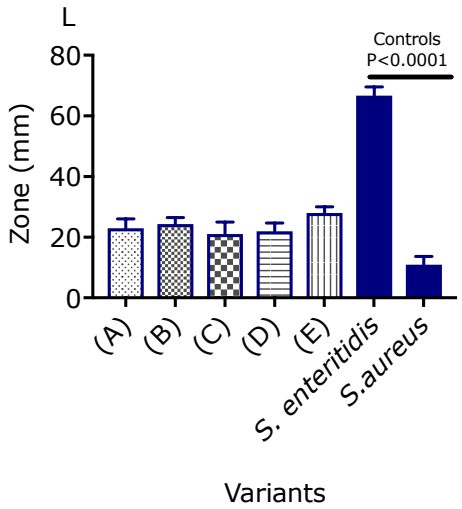
4.2.1.2 Identical resistance profile samples

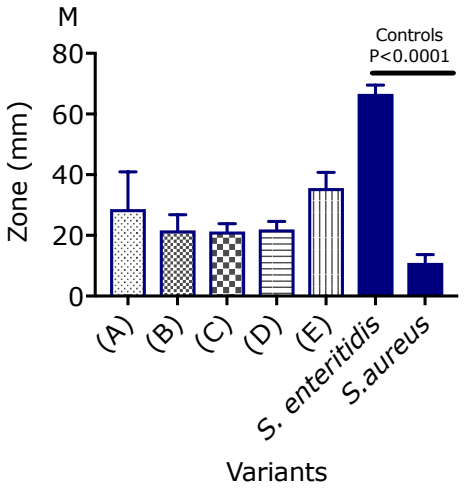
Five random colonies of identical resistance pattern were selected to determine if motility varied between colonies that were thought to be derived from a homogenous population. The motility assay was applied to five colonies assigned A, B, C, D and E from each sample as illustrated in figure 4.3. Isolates from female samples J and K, and male samples L, M and N in figure 4.3 show no significant differences in motility. All isolates from within a sample were motile and behaved similarly.

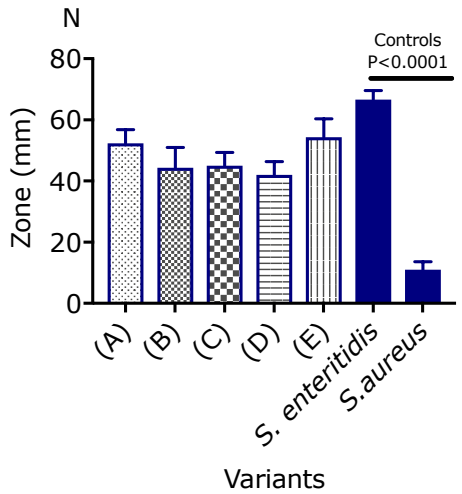
Identical resistance pattern samples











Figures 4.3 Comparison of motile ability of bacterial isolates belong to female and male identical resistance profile pattern samples. Motility assays, using semi-solid agar plates, were applied to identical profile pattern samples. Bacterial isolates of each sample were grown on semi-solid agar and incubated at 37 °C for 18 hours. The bacterial isolates of each sample were assigned with a letters (A to E), each represented the identical susceptibility pattern to an antibacterial agent based on susceptibility testing. The corresponding bar chart has been listed with five letters, in order to represent each of the five samples of identical group, as follows: (J) : UTI (F20), (K) : UTI (F21), (L) : UTI (M21), (M) : UTI (M20) and (N) : UTI (M19). The bar charts display a representation of the mean of the migration zone across three independent experiments. Furthermore, they are based on three replicates with error bars to indicate the standard deviation of each variable. *S.enteritidis* and *S.aureus* were each used as a positive and negative control respectively, and were also visually represented through documented bar charts. There is no significant difference in zone of migration in (mm) between variants of each sample.

4.2.2 Biofilm formation using crystal violet among samples variants

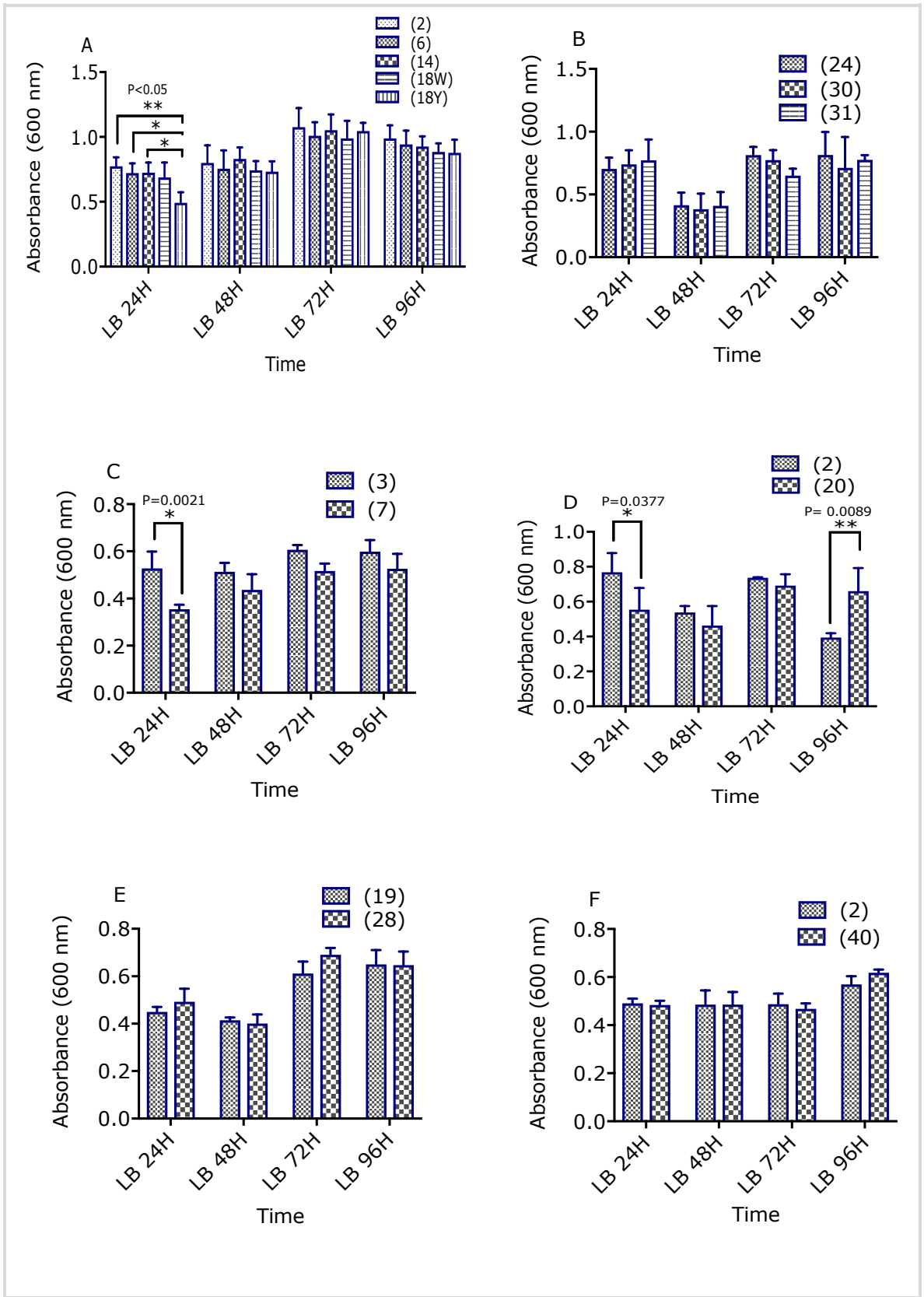
Biofilm formation was measured using the crystal violet assay on the highly diverse resistance profile sample isolates and five apparently identical colonies of identical resistance profile. The data for the five colonies of identical resistance pattern was determined to evaluate the base line variation level of biofilm formation. Bar charts in figures 4.4, 4.5 and 4.6 show the mean absorbance for various examined isolates within samples at 600 nm.

4.2.2.1 Highly diverse resistance profile sample variants

The biofilm-forming ability of isolates from highly diverse resistance samples was compared. Bar charts in figures 4.4 and 4.5 represent the mean of measured optical density for variants within each sample. Isolates from three female samples A, C and D varied in biofilm formation (Figure 4.4), with the variation observed after 24 hours. In sample (D) variation in biofilm production ability was observed after 96 hours incubation. In sample A, a significant variation in biofilm formation ability was noticed between variants 18Y and 2 (P value < 0.001), and also with variants 6 and 14 (P value < 0.05) (F value = 0.7538). variant 18Y was less proficient in biofilm production than the others. Sample C variants also showed a significant difference in biofilm formation, with variant 3 observed to be highly able to produce biofilm when compared to isolate 7 (P value = 0.0021; F value = 26.6). In addition, in sample D, variants behaved differently in biofilm formation; variant 2 being able to produce biofilm after 24 hours (P value = 0.0377) while isolate 20 was increasingly able to form biofilm after 96 hours (P value = 0.0089) (F value = 7.691). The other three female samples showed no significant differences between within sample isolates

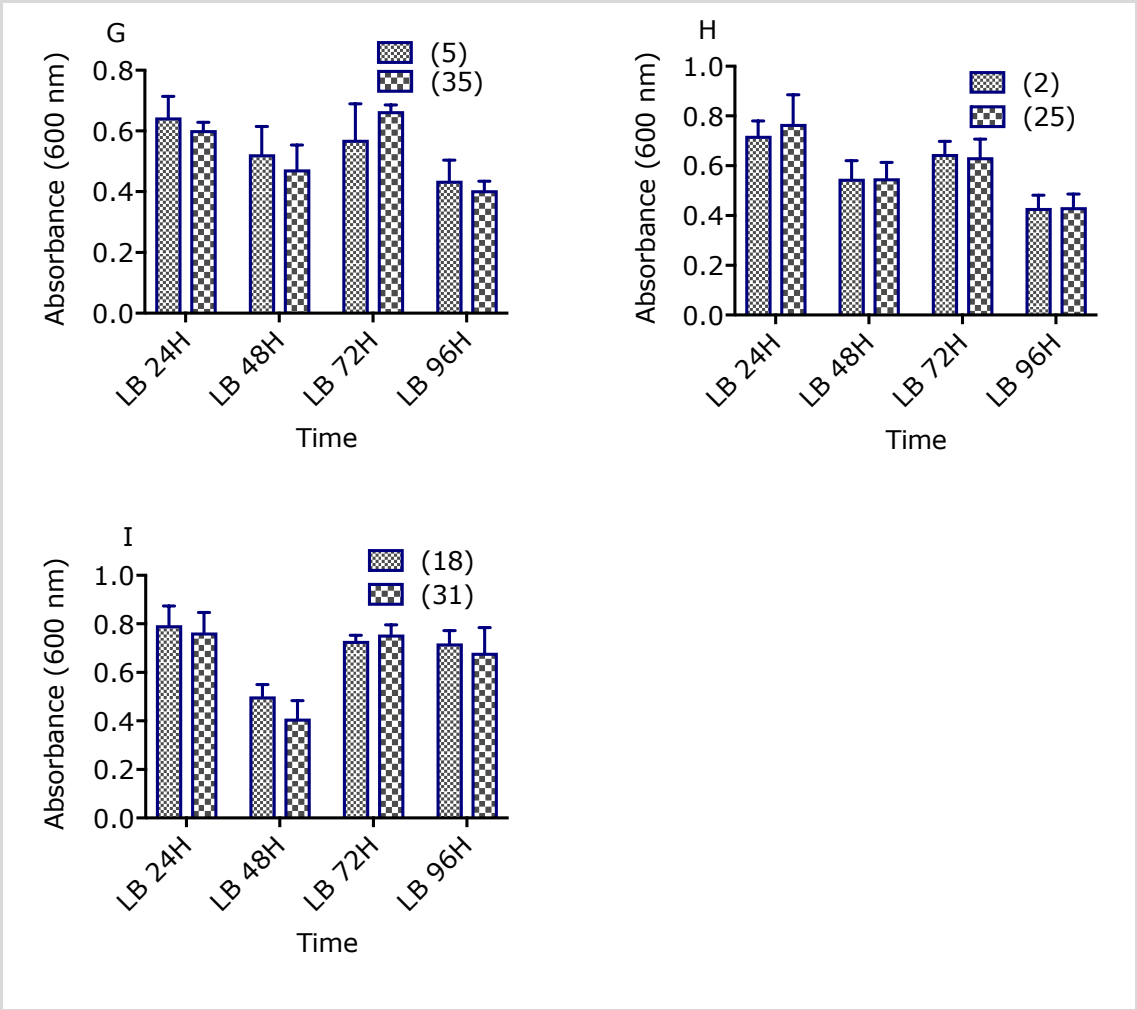
(Figure 4.4). In addition to that, no significant variation was noticed between isolates from male samples as shown in figure 4.5.

A. Highly diverse resistance profile female samples



Figures 4.4 Comparison of biofilm formation ability using crystal violet assay among bacterial isolates (variants) belong to female highly diverse resistance profile pattern samples. Biofilm formation using crystal violet was applied to six different female samples. Bacterial isolate of each sample were grown in 5 ml LB liquid media and incubated at 37 °C for 18 hours without shaking. A 1:100 dilution was incubated in triplicate using 96 well microtitre plates at 37 °C for 24 hours, 48 hours, 72 hours and 96 hours. The bacterial isolates of each sample were assigned with a number, each represented the variation between sensitivity and resistance to an antibacterial agent based on susceptibility testing. The corresponding bar chart has been listed with six letters, in order to represent each of the six female samples, as follows: (A) : UTI (F1), (B) : UTI (F2), (C) : UTI (F3), (D) : UTI (F4), (E) : UTI (F5) and (F) : UTI (F6). The bar charts display a representation of the mean of absorption across three independent experiments. Furthermore, they are based on three replicates with error bars to indicate the standard deviation of each variable. There was no significant difference in absorbance between bacterial isolates (variants) of sample (B), (E) and (F) while a significant variation was noticed between bacterial isolates (variants) of sample (A), (C) and (D). (*) Indicates that P value is less than 0.05 and (**) indicates that P value is less than 0.01. Further, K_{12} and LB were used for control purposes.

B. Highly diverse resistance profile male samples

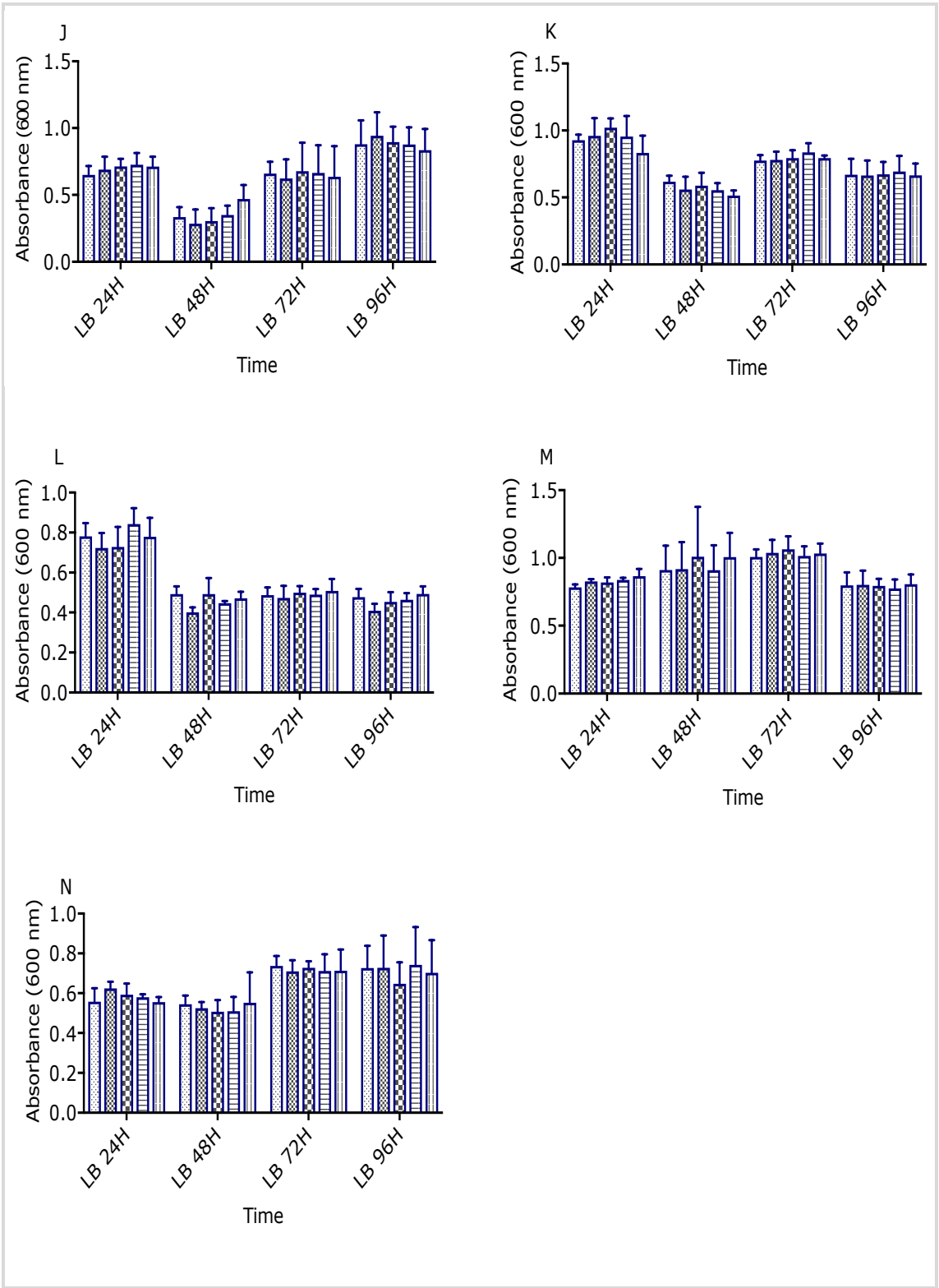


Figures 4.5 Comparison of biofilm formation ability using crystal violet assay among bacterial isolates (variants) belong to male highly diverse resistance profile pattern samples. Biofilm formation using crystal violet was applied to three different male samples. Bacterial isolate of each sample were grown in 5 ml LB liquid media and incubated at 37 °C for 18 hours without shaking. A 1:100 dilution was incubated in triplicate using 96 well microtitre plates at 37 °C for 24 hours, 48 hours, 72 hours and 96 hours. The bacterial isolates of each sample were assigned with a number, each represented the variation between sensitivity and resistance to an antibacterial agent based on susceptibility testing. The corresponding bar chart has been listed with three letters, in order to represent each of the three male samples, as follows: (G) : UTI (M3), (H) : UTI (M1) and (I) : UTI (M2). The bar charts display a representation of the mean of absorption across three independent experiments. Furthermore, they are based on three replicates with error bars to indicate the standard deviation of each variable. There is no significant difference in biofilm formation between variants of each sample. Further, K_{12} and LB were used for control purposes.

4.2.2.2 Identical resistance profile samples

Five random colonies of identical resistance pattern were selected to determine if biofilm formation level varies between colonies that thought to be derived from a homogenous population. Five isolates (A - E) from three male samples (L, M and N) and two female samples (J and K) were tested and there were no significant differences in biofilm formation for isolates from within the same clinical sample.

Identical resistance profile samples



Figures 4.6 Comparison of biofilm formation ability using crystal violet assay of bacterial isolates belong to female and male identical resistance profile pattern samples. Biofilm formation using crystal violet was applied to the identical resistance profile samples. Bacterial isolate of each sample were grown in 5 ml LB liquid media and incubated at 37 °C for 18 hours without shaking. A 1:100 dilution was incubated in triplicate using 96 well microtitre plates at 37 °C for 24 hours, 48 hours, 72 hours and 96 hours. The bacterial isolates of each sample were assigned with a letters (A to E), each represented identical sensitivity pattern to an antibacterial agent based on susceptibility testing. The corresponding bar chart has been listed with five letters, in order to represent each of the five samples of the identical group, as follows: (J) : UTI (F20), (K) : UTI (F21), (L) : UTI (M21), (M) : UTI (M20) and (N) : UTI (M19). The bar charts display a representation of the mean of absorption across three independent experiments. Furthermore, they are based on three replicates with error bars to indicate the standard deviation of each variable. There is no significant difference in biofilm formation between variants of each sample. Further, K_{12} and LB were used for control purposes.

4.2.3 Association and invasion of uropathogenic *E.coli* isolates to T24 among resistance profile patterns

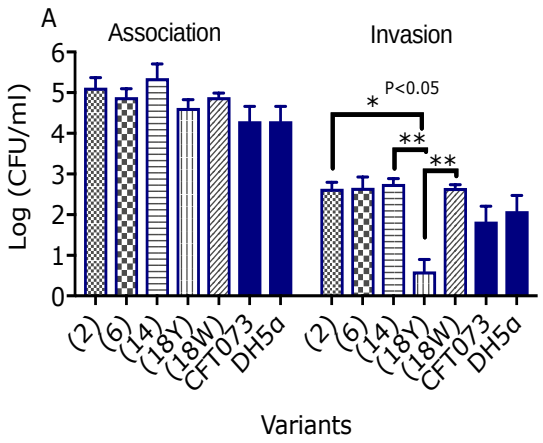
Comparisons of association and invasion of UPEC strains to T24 cells were performed on highly diverse resistance profile sample variants and five apparently identical colonies of identical resistance profile. Data for five colonies of identical resistance pattern were also determined to evaluate the baseline variation of association and invasion using the Miles and Misra technique (Miles et al., 1938). Bar charts in figures 4.7, 4.8 and 4.9 represent the mean log of calculated CFU/ml for isolates within samples.

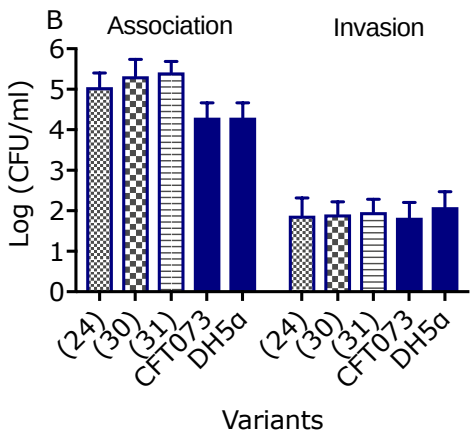
4.2.3.1 Highly diverse resistance profile sample variants

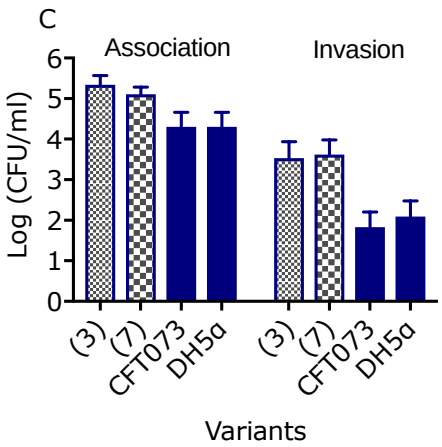
Bar charts in figure 4.7 and 4.8 represent the log of CFU/ml in association and invasion assays of different within sample variants to T24 cells. With regard to the association assay in figure 4.7, all isolates from female samples behaved identically. However variants from sample D showed a variation in association ability (P value < 0.0001; F value = 9.84). Isolate 2 showed about a 1000 fold increase in association ability compared to the other variant. For isolates from samples from female patients, the invasion ability in samples A and F was varied. In sample A, strain 18Y invaded T24 cells less than the other strains and was signifcantly less invasive than strain 2 (P value < 0.05), and strains 14 and 18W (P value < 0.001) (F vlaue = 1.285). The other signifcant variation was noticed in sample F, where variant 2 was about 10000 fold more invasive than isolate 40 (P value = 0.0054, F value = 4.019). On the other hand, male samples in figure 4.8 displayed no significant variation between isolates in association ability. Isolates from only one male sample show a significant variation in invasion ability where sample G showed significant variation between

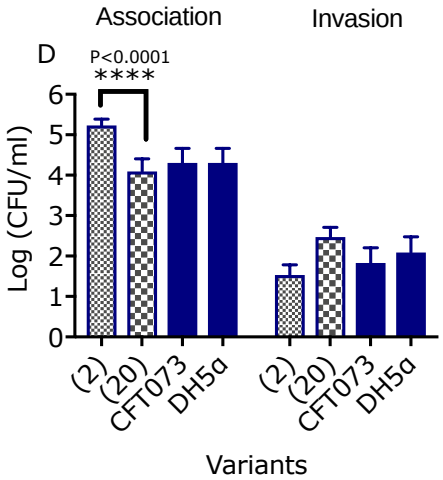
the two variants, with strains 5 showing higher invasion of T24 cells than isolate 35 (P vlaue < 0.05; F value = 1.958).

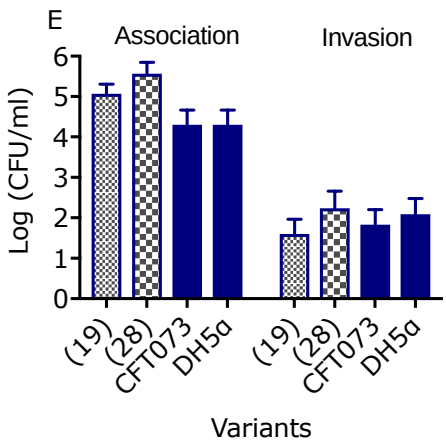
A. Highly diverse resistance profile female samples

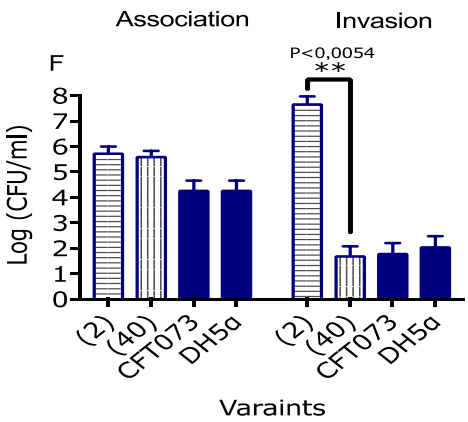






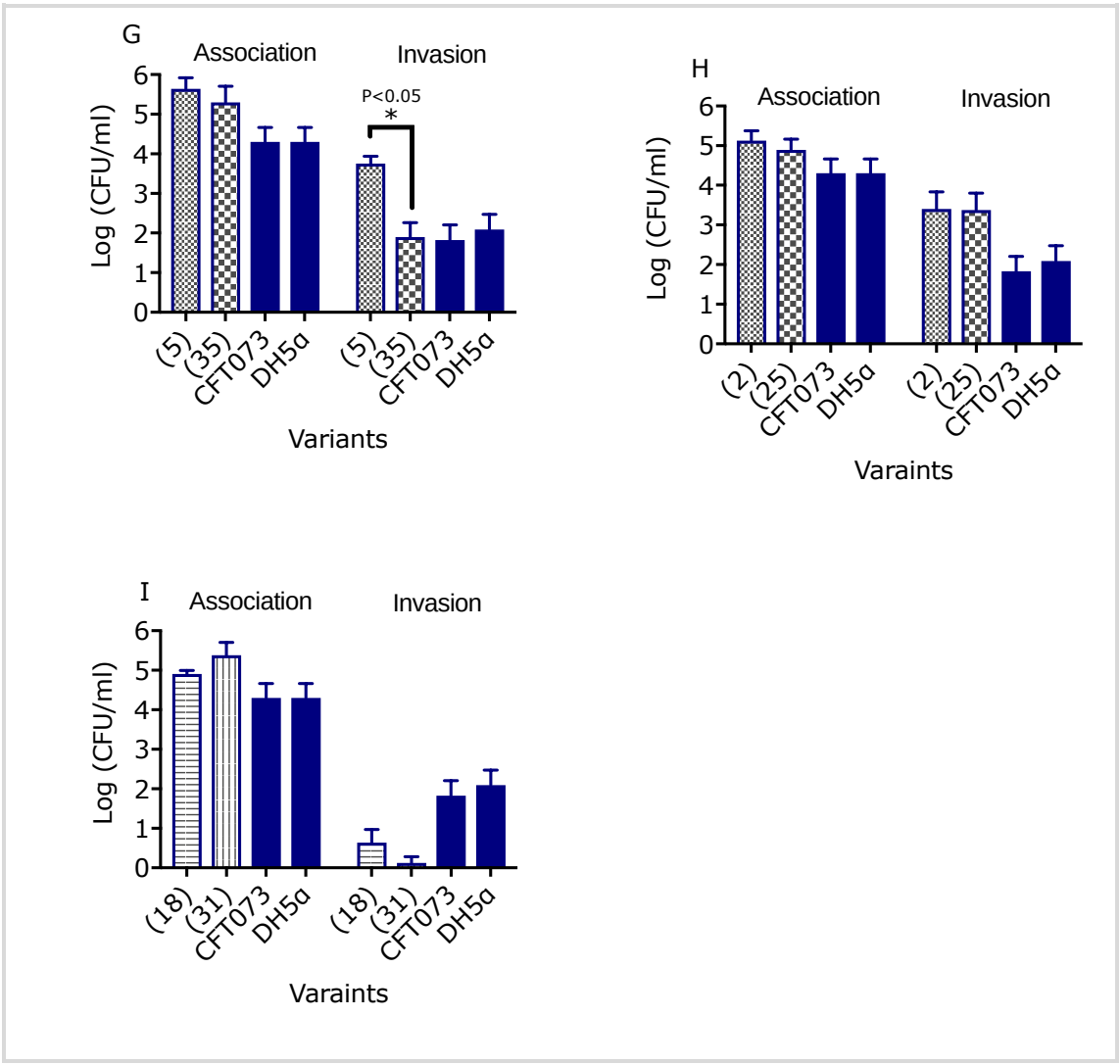






Figures 4.7 Comparison of association and invasion ability to T24 among bacterial isolates (variants) belong to female highly diverse resistance profile pattern samples. Association and invasion assay were applied to six different female samples. Bacterial isolate of each sample were grown in 5 ml LB liquid media and incubated at 37 °C for 18 hours with shaking. A 1:100 dilution was incubated in duplicate using 24 well microtitre plates at 37 °C with CO_2 for 3 hours for the association assay and after the addition of Gentamicin were incubated at 37 °C with CO_2 for 2 hours. The bacterial isolates of each sample were assigned with a number, each represented the variation between sensitivity and resistance to an antibacterial agent based on susceptibility testing. The corresponding bar chart has been listed with six letters, in order to represent each of the six female samples, as follows: (A) : UTI (F1), (B) : UTI (F2), (C) : UTI (F3), (D) : UTI (F4), (E) : UTI (F5) and (F) : UTI (F6). The bar charts display a representation of the mean of log CFU/ml of three independent experiments. Furthermore, they are based on three replicates with error bars to indicate the standard deviation of each variable. CFT073 and DH5α were used as a positive and negative control in duplicate and represented as documented bar charts. Significant variation were shown in association between bacterial isolates (variants) of sample (D) and in invasion in sample (A) and (F). (*) Indicates that the P value is less than 0.05, (**) (P < 0.001) and (****) indicates (P value < 0.0001).

B. Highly diverse resistance profile male samples

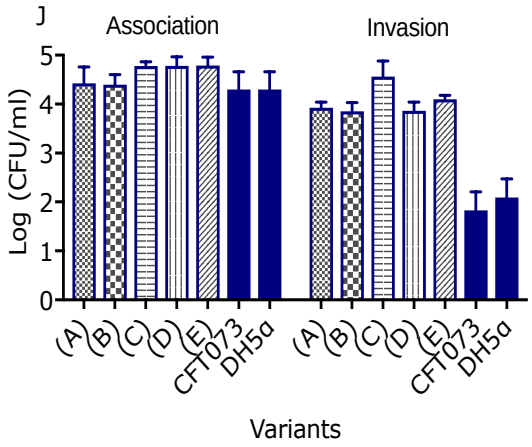


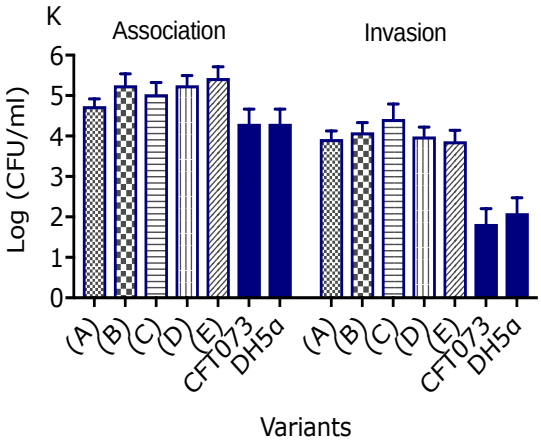
Figures 4.8 Comparison of association and invasion ability to T24 among bacterial isolates (variants) belong to male highly diverse resistance profile pattern samples. Association and invasion assay were applied to three different male samples. Bacterial isolate of each sample were grown in 5 ml LB liquid media and incubated at 37 °C for 18 hours with shaking. A 1:100 dilution was incubated in duplicate using 24 well microtitre plates at 37 °C with CO_2 for 3 hours for the association assay and after the addition of Gentamicin were incubated at 37 °C with CO_2 for 2 hours. The bacterial isolates of each sample were assigned with a number, each represented the variation between sensitivity and resistance to an antibacterial agent based on susceptibility testing. The corresponding bar chart has been listed with three letters, in order to represent each of the three male samples, as follows: (G) : UTI (M3), (H) : UTI (M1) and (I) : UTI (M2). The bar charts display a representation of the mean of log CFU/ml of three independent experiments. Furthermore, they are based on three replicates with error bars to indicate the standard deviation of each variable. CFT073 and DH5α were used as a positive and negative control in duplicate and represented as documented bar charts. There is significant variation in invasion between bacterial isolates (variants) of sample (G). (*) Indicates that the P value is less than 0.05.

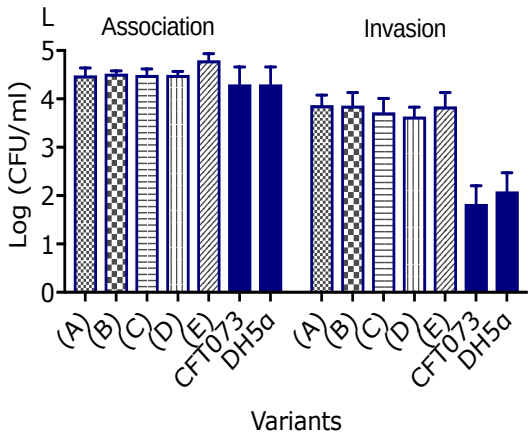
4.2.3.2 Identical resistance profile samples

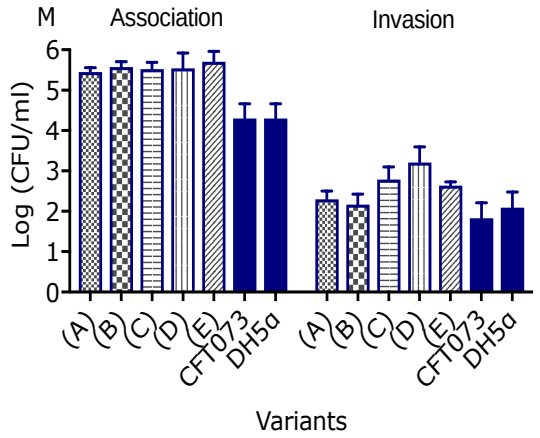
Five random colonies of identical resistance pattern were selected to determine if the association and invasion level varied between colonies that were thought to be derived from a homogenous population. Association and invasion assays were applied on five identical pattern isolates (A, B, C, D and E) obtained from female samples (J) and (K) and male samples (L), (M) and (N), as shown in figure 4.9. There was no significant variation between examined colonies. Further, these samples seemed to have higher invasion capabilities than those from the highly diverse samples in general.

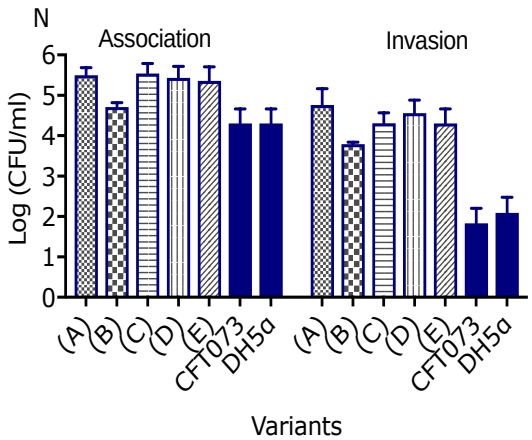
Identical resistance profile samples











Figures 4.9 Comparison of association and invasion ability to ability to T24 among bacterial isolates belong to female and male identical resistance profile pattern samples. Association and invasion assay were applied to five identical resistance profile samples. Bacterial isolate of each sample were grown in 5 ml LB liquid media and incubated at 37 °C for 18 hours with shaking. A 1:100 dilution was incubated in duplicate using 24 well microtitre plates at 37 °C with CO_2 for 3 hours for the association assay and after the addition of Gentamicin were incubated at 37 °C with CO_2 for 2 hours. The bacterial isolates of each sample were assigned with letters (A to E), each represented identical sensitivity pattern to an antibacterial agent based on susceptibility testing. The corresponding bar chart has been listed with five letters, in order to represent each of the five samples of identical group, as follows: (J) : UTI (F20), (K) : UTI (F21), (L) : UTI (M21), (M) : UTI (M20) and (N) : UTI (M19). The bar charts display a representation of the mean of log CFU/ml of three independent experiments. Furthermore, they are based on three replicates with error bars to indicate the standard deviation of each variable. CFT073 and DH5α were used as a positive and negative control in duplicate and represented as documented bar charts.

4.2.4 Summary table of varied phenotypic abilities among highly diverse samples

About 55% of samples yielding isolates with differing resistance profiles were found to contain isolates which differed when assessed by additional phenotypic assays, as illustrated in table 4.2. Resistant bacterial isolates of various samples were noted to be the most highly motile, best produces of biofilms, most adherent and invasive to T24 cells. Most of the Ampicillin resistant isolates showed a high ability to form biofilm. However, those phenotypic variations among highly diverse pattern variants indicated no specific relationship with antibiotic class. Samples from female patients yielded the majority of highly diverse resistance pattern isolates and also the most significant phenotypic variation between variants from within samples.

Table 4.2 Summary of the highly diverse resistance pattern samples. The patterns observed were dependent upon the variants (bacterial isolates) of highly diverse samples, the variants present within each sample, the antibiotics that bacterial isolates were assorted in, and where those variants have a significant phenotypic difference. MDR referred to samples with multi drug resistance ability. (*) represents where the significant variation was found.

Sample ID Variants		MDR	Antibiotic resistance profile		Phenotypic assay		
					Motility assay	Biofilm production	Association and invasion Assay
Highly diverse resistance profiles							
(A) UTI (F1)	**2**		AP/TM	AP/TM resistance		*	*
	6			TM resistance			
	14			AP resistance			
	18W			AP/ TM resistance			
	18Y						
(B) UTI (F2)	**24**	Yes	GM/CAZ/CPM/CPD	GM resistance			
	30			GM sensitive			
	31			GM/CAZ/CPM/CPD resistance			
(C) UTI (F3)	**3**	Yes	AP	Resistance variant		*	
	7			Sensitive variant			
(D) UTI (F4)	**2**		AP	Resistance variant	*	*	*
	20			Sensitive variant			
(E) UTI (F5)	**19**		CAZ	Resistance variant			
	28			Sensitive variant			
(F) UTI (F6)	**2**	Yes	GM	Resistance variant			*
	40			Sensitive variant			
(G) UTI (M3)	**5**	Yes	GM/TM	Resistance variant			*
	35			Sensitive variant			
(H) UTI (M1)	**2**		CIP	Sensitive variant			
	25			Resistance variant			
(I) UTI (M2)	**18**		AP	Resistance variant			
	31			Sensitive variant			

Phenotypic assays was applied also on selection of the low diversity resistance profile pattern isolates. Isolates from four female and four male samples were selected and the same phenotypic assays were applied. There were no significant phenotypic differences recognised between the examined isolates as illustrated in table 4.3.

With regard to the overall phenotypic variation between the three resistance profile patterns (high, low and identical), significant phenotypic variation was only noticed in isolates from the highly diverse resistance pattern samples. No phenotypic variation was observed in the low and identical resistance pattern samples.

Table 4.3: Phenotypic variation of some low diverse resistance profile pattern samples. The patterns observed were dependent upon the variants (bacterial isolates) of low diverse samples, the variants present within each sample, the antibiotics that bacterial isolates were assorted in, and where those variants have a significant phenotypic difference. MDR referred to samples with multi drug resistance ability.

Sample ID	Variants	MDR	Antibiotic resistance profile		Phenotypic assay		
					Motility assay	Biofilm production	Association and invasion Assay
Low diverse resistance profiles							
UTI (F13)	**17**		GM/TM	All sensitive	None		
	38			GM intermediate			
	43			GM/TM intermediate			
UTI (F16)	**1**		GM	GM sensitive	None		
	5			GM intermediate			
UTI (F17)	**13**		CIP	CIP intermediate	None		
	20			CIP sensitive			
UTI (F19)	**2**	Yes	GM	GM resistance	None		
	12			GM intermediate			
UTI (M11)	**5**	Yes	MM	MM intermediate	None		
	12			MM sensitive			
UTI (M12)	**2**		GM	GM sensitive	None		
	16			GM intermediate			
UTI (M16)	**10**		GM	GM intermediate	None		
	25			GM sensitive			
UTI (M10)	**2**		CPM	CPM sensitive	None		
	14			CPM intermediate			

4.3 Discussion

The development of a urinary tract infection within a patient is thought to be related to the presence of a combination of virulence factors in the pathogen (Dobrindt, 2005). Bacterial motility, attachment and invasion of the host, alongside the ability to form biofilm, are all factors which allow UPEC to survive within the urinary tract and cause infection. There is a natural heterogeneity within UPEC in terms of motility, host adherence and invasion (Naves et al, 2008). This variability has been recognised in comparisons between *E. coli* isolates collected from patient's urine samples. *E. coli* isolates obtained over time from a symptomatic patient with a urinary tract infection proved to be genetically heterogeneous and that heterogeneity included metabolic functions such as iron uptake and stress resistance (Zdziarski et al., 2010). In addition, within-host phenotypic variation was demonstrated in eight patients infected with ExPEC by Levert and his group (Levert et al., 2010). Therefore, as with the resistance variation obtained in Chapter 3, colonies that have varied antibiotic susceptibility profiles could have variability to other phenotypic characteristic.

In this chapter, isolates from within each sample showing a highly diverse resistance profile were compared using a motility assay based on spreading on a semi-solid agar plate. Flagella and chemotaxis are thought to be expressed at specific stages, and at specific sites for efficient colonisation (Lane et al., 2005). Flagella are also thought to be highly active in the upper urinary tract leading to the colonisation of the kidneys (Lane et al., 2007). It was noticed that there was a significant variation in motility between variants obtained from the sample form one female case; one variant was motile while the other one was not. Interestingly, the motile bacterial isolate is a resistant colony from the sample. This patient could have a mixed

infection with two different sequence types or with a strain persisting from an old infection, or may have been reinfected with another strain of *E. coli*. Those two strains may behave differently in biofilm formation as illustrated by Pratt and Kolter (1998), where non-motile strains of *E. coli* were deficient in the early stages of biofilm formation (Pratt and Kolter, 1998).

Ponnusamy et al. (2012) reported that most antibiotic-resistant strains were characterized by their ability to form biofilm. It was also shown that biofilm producing isolates are difficult to eradicate form the host and a prolonged antibiotic treatment is needed (Graham and Galloway, 2001). As a correlation with that, variation in biofilm formation was noticed in our samples. Our data showed variation in biofilm formation between the resistant and sensitive isolates obtained from three female patients after 24 hours of incubation. In samples C and D, resistant strains were able to form biofilm and these data agreed with Ponnusamy study. In addition, the non-motile variant from sample D showed a significant difference in its ability to form biofilm after 96 hours incubation and the significance here may due to the resistant strain not being able to survive after 96 hours of incubation. This correlates with a study reporting that the optimal time to examine biofilm production in *E. coli* strains was after 24 hours of incubation (Adamus-Białek et al., 2015). A study on women with recurrent urinary tract infections found that isolates from 74% of recurrent infection cases were able to form biofilm (Soto et al., 2007). Therefore, our female samples having a varied biofilm production level may have a previous history of having UTIs. Generally, all sample isolates were able to form biofilm when compared with the *E. coli* K12 control. Those isolates may show a high

attachment level, and Soto et al. (2007) determined that most biofilm producing strains express type 1 fimbriae.

The ability of isolates from samples with highly diverse resistance profiles to attach to T24 cells and invade them was tested using the gentamicin protection assay. Attachment of UPEC is considered to be the early stage in colonization of the urinary tract leading to subsequent disease (Martinez et al., 2000; Bahrani-Mougeot et al., 2002; Connell et al., 1996). Generally, all isolates attached to T24 cells at a high level and this could reflect the role of adhesive organelles which are important in biofilm formation (Cegelski et al., 2009). Sample D isolates exhibit a significant difference in attachment level to T24 cells and this can be linked with presence of difference in motility and biofilm production between variants from this female patient. Invasion of bladder cells generally has an important role in urinary tract infections persistence (Martinez et al., 2000). Variation in invasion ability between isolates was noticed in samples from two female and one male case. In sample A, one isolate was characterized by low invasion ability and was previously shown to have a low level of biofilm formation ability. The resistant isolates in the other two samples showed high invasion levels in T24 cells, but no significant responses were noticed in the other phenotypic assays. Isolates from these two samples exhibited the crucial factor for infection persistence, the ability of bacterial cells to invade the urinary tract epithelium followed by intracellular bacterial community development (Blango and Mulvey, 2010). Especially for sample F, persistence of infection with possible dissemination to other areas may be relevant in this patient.

Obtained phenotypic variation in the highly diverse resistance profile samples might have arisen from the presence of two or more different sequence types of UPEC.

Further, recurrent infections could have an impact on patients in term of population diversity. Recurrent infection could be as a result of either relapsed infection with the original strains or reinfection by different strains of *E. coli* (Soto et al., 2011). Those two possibilities are relevant because most of the samples yielded isolates which tended to form biofilm. Presence of two different sequence types may be more relevant in sample D, as all examined phenotypic traits varied. This can result from an infection with two different sequence types or a recurrent infection with persistence of the previous strain. Generally, 55 % of highly diverse resistance profile samples are varied phenotypically among examined phenotypic assays in this chapter.

The isolates from samples with highly diverse resistance profiles were also compared with five seemingly identical isolates of identical resistance patterns to obtain an overview about the baseline phenotypic diversity level. The phenotypic data for isolates from samples which were believed to contain a homogenous population were examined, and no significant differences were apparent. Baseline phenotypic heterogeneity was therefore very low**.**

Chapter Five

Genotypic diversity of Uropathogenic *Escherichia coli* colonies from a single patient

5.1 Introduction

5.1.1 *Escherichia coli* genotypic analysis

E. coli is a genetically heterogeneous group of organisms, and this heterogeneity is thought to be related to acquisition or deletion of genetic elements. Using molecular genetic approaches, genetic diversity can be determined among pathogenic and commensal *E. coli* isolates. It has been demonstrated that using molecular techniques can help in studying the evolutionary process of *E. coli* strains, and can determine the varied genetic content of *E. coli* isolates (Dobrindt et al., 2003). As an illustration, multiplex PCR was used in an epidemiological study to determine widespread variation in presence of virulence factors among 75 ExPEC isolates (Johnson and Stell, 2000).

By applying genome sequencing, it has been shown that UPEC isolates can acquire extra genetic material which may provide specific phenotypic traits for particular strains (Dobrindt, 2005). Most acquired mobile genetic elements are virulence-associated genes that cause structural and functional diversity (Dobrindt, 2005). Brzuszkiewicz et al. (2006) published a study comparing UPEC phenotypes and their genomes to give a better understanding of UPEC diversity by comparing the sequenced isolated genome to the CFT073 genome. They showed that most UPEC virulence genes are located on a pathogenicity island.

Lloyd et al. (2007) applied comparative genomic hybridisation to different UPEC genomes extracted from different disease sets of the urinary tract and from commensal isolates. It was found that UPEC virulence genes were differentially present across examined isolates. These genes were distributed randomly among UPEC bacterial strains (Lloyd et al., 2007). Comparative genomic hybridisation was also carried out on eleven

UPEC of asymptomatic bacteriuria and it was reported that there was diverse genetic content among those isolates (Zdziarski et al., 2008).

Molecular studies have also been carried out on ExPEC to determine within-host diversity using MLST, PFGE and PCR on 226 isolates. Whole-genome sequencing using Illumina technology was used on eight isolates. Together, these showed two types of within-host diversity of ExPEC: polyclonal and monoclonal (Levert et al., 2010). Additionally, 265 of the *E. coli* isolates of inpatients and outpatients were compared phenotypically and using molecular methods. Molecular analysis study was based on using multiplex PCR and MLST. It was shown that, in UTIs caused by the heterogeneous group of *E. coli,* most virulence-associated genes were located on mobile elements such as plasmids (Toval et al., 2014). In 2015, genome sequencing was applied to 19 UPECs to give a better understanding about the degree of diversity among strains and it was found that closely related UPEC isolates belonging to the same sequence type may differ genetically (Lo et al., 2015).

5.1.2 Whole-genome sequencing

Bacterial genotyping is becoming increasingly important, especially in nosocomial infections associated with high morbidity and mortality. Microbiological epidemiological studies are important to understand the distribution and relatedness of bacterial pathogens and to control bacterial infection. Whole-genome sequencing is rapidly replacing traditional genotyping methods as it has become fast and cheaper over time (Bertelli and Greub, 2013). Whole-genome sequencing data of bacterial isolates can provide unparalleled information about virulence and antibiotic resistance genes found within bacterial isolates, and provides a better understanding of evolutionary processes (Mardis, 2011; Bertelli and Greub, 2013). Further, whole-genome sequencing can

provide a rapid and accurate identification of the real potential source of bacteria that causes an outbreak, which enables researchers to inform infection control interventions (Harris et al., 2010; Harris et al., 2013).

5.1.2.1 Using whole-genome sequencing to detect within population variation

Whole-genome sequencing is applied to determine within- and between-host genetic diversity. *De novo* assembly is one method used to provide information about short insertion or deletion between closely related bacterial strains (Maclean et al., 2009). This method was used on isolates from MRSA carriage to detect insertions and deletions (Golubchik et al., 2013). Additionally, aligning of sequenced data against a reference genome is also another method to provide information about SNPs and has importance in investigation of evolutionary process during bacterial infection. Mapping sequenced bacterial isolates to a reference genome provides information about accumulated mutations and was applied on isolates of *B. dolosa* isolated from single patients (Lieberman et al., 2011; Lieberman et al., 2014).

5.1.3 Aim of this chapter

Variation in resistance profile between colonies could be a sign of genotypic diversity, and we observed multiple cases of such variation in colonies from the same patient. In addition, we observed variation in other phenotypic characteristics between colonies obtained from the same patient, which would also suggest significant genotypic variation. In this chapter, we aimed to determine to what extent those phenotypic variation are based on genetic differences. Johnson et al. (2001) reported that vertical inheritance and horizontal transmission within bacterial populations produce divergent patterns of virulence factors. Therefore, in this chapter, determining genetic diversity

within a single patient will create a detailed picture of the level of within-host diversity of UPEC within patients.

We performed whole-genome sequencing on all variants characterised previously in Chapter three from the highly diverse resistance profile samples. The aim was to investigate if there were multiple strain types present in these samples, or if key gene gain events had occurred within the infecting population that led to clinically important shifts in phenotype within the population.

Low resistance profile and identical pattern samples were also evaluated to see if these groups contained identical or low diversity in populations of infecting cells. To do this, 19 random low-diversity samples and the 5 identical samples were selected, and we performed deep population sequencing on complete bacterial growth of those samples. By using a mapping approach, we could then identify the total number of SNPs in each of the infecting populations and quantify the baseline level of genetic diversity across these sample types.

5.1.4 Bacterial strains used in this chapter

Bacterial isolates (variants) were obtained from UPEC samples with varied resistance profiles as described in Chapter 3, were used in this chapter and described in table 5.1.

Table 5.1 Bacterial isolates used in this chapter

Colony-based whole-genome sequencing samples			
High diverse resistance profile samples			
Sample ID	**Sample variant**	**Resistance profile**	**Phenotypic variation**
UTI (F1)	**Colony 2**	Resistance to AP and TM	Varied in biofilm formation and invasion assay
	Colony 6	Resistance to TM	
	Colony 14	Resistance to AP	
	Colony 18W	Resistance to AP	
	Colony 18Y	Intermediate to TM	
UTI (F2)	**Colony 24**	Resistance to AP, GM, CPM, CIP and TM	No phenotypic variation among examined assays
	Colony 30	Resistance to AP, CIP and TM	
	Colony 31	Resistance to GM, AP, CPD, CPM, CIP and TM	
UTI (F3)	**Colony 3**	Resistance to AP and CPM	Varied in biofilm formation
	Colony 7	Sensitive to AP, Intermediate to CPM	
UTI (F4)	**Colony 2**	Resistance to AP	Varied in all examined phenotypic assays
	Colony 20	Sensitive to AP	
UTI (F5)	**Colony 19**	Resistance to CAZ	No phenotypic variation among examined assays
	Colony 28	Sensitive to CAZ	
UTI (F6)	**Colony 2**	Resistance to GM	Varied in invasion assay
	Colony 40	Sensitive to GM	
UTI (M3)	**Colony 5**	Resistance to GM and TM	Varied in invasion assay
	Colony 35	Sensitive to GM and TM	
UTI (M1)	**Colony 2**	Resistance to CIP	No phenotypic variation among examined assays
	Colony 25	Sensitive to CIP	
UTI (M2)	**Colony 18**	Resistance to AP	
	Colony 31	Sensitive to AP	
Low diverse resistance profile sample			
UTI (F17)	**Colony 13**	Varied in CIP	No phenotypic variation among examined assays
	Colony 20		

Deep sequencing samples		
Low diverse resistance profile samples		
Sample ID	**Sample variant**	**Resistance profile**
UTI (F16)	Pooled colonies of the whole bacterial growth	Varied in GM
UTI (F19)		Varied in GM
UTI (F13)		Varied in GM/TM
UTI (F12)		Varied in CAZ/CPM
UTI (F9)		Varied in CIP/TM
UTI (F7)		Varied in GM/CPM/CAZ/MM
UTI (F8)		Varied in GM/CPM/CAZ
UTI (F11)		Varied in GM/CPM/CAZ
UTI (F15)		Varied in GM/CPM
UTI (M11)		Varied in MM
UTI (M12)		Varied in GM
UTI (M16)		Varied in GM
UTI (M5)		Varied in GM/CPM/CAZ/MM
UTI (M4)		Varied in GM/CPM/CAZ/MM/TM
UTI (M7)		Varied in GM/CPM/CAZ
UTI (M9)		Varied in MM
UTI (M6)		Varied in GM/CPM/CAZ
UTI (M10)		Varied in CPM
UTI (M8)		Varied in GM/CPM
Identical resistance profile samples		
UTI (F20)	Pooled colonies of the whole bacterial growth	Sensitive to MM only
UTI (F21)		Resistance to GM, AP and TM
UTI (M21)		Sensitive to MM only
UTI (M20)		Sensitive to all examined antibiotics
UTI (M19)		Sensitive to MM only

5.2 Bioinformatics programs

Raw FASTQ files for bacterial isolates were assembled using SPAdes (version 3.5.0). Detailed information is described in section 2.7.1, **Chapter** 2.

5.2.1 QUAST

QUAST was applied on all assembled genomes providing a range of metric assessments in tables 5.2 and 5.4 (version 2.0; Gurevich et al., 2013) (section 2.7.2, **Chapter** 2).

5.2.2 Multi locus sequence type

Multi locus sequence type (MLST) of whole-genome sequenced isolates was identified using CGE website (http://cge.cbs.dtu.dk/services/MLST/) and are shown in table 5.2 and 5.4 (version 1.8).

5.2.3 Bacterial analysis pipeline

A bacterial analysis pipeline was created for whole genome sequenced samples using CGE website (https://cge.cbs.dtu.dk/services/cge/) (version 2015-04-28). The bacterial analysis pipeline can provide information on carriage of resistance genes, virulence genes, *FimH* adhesion type and plasmids and are shown in table 5.3.

5.2.4 Pan-genome sequence analysis

Pan-genome sequence analysis (https://lfz.corefacility.ca/panseq/page/novel_full.html) was applied to determine the novel regions that varied between sequenced bacterial isolates belonging to a single patient sample and illustrated in tables 5.2 to 5.11 (Laing et al., 2010).

5.2.5 BLASTx

Novel regions identified between bacterial isolates within samples were identified using BLASTx web-base (https://blast.ncbi.nlm.nih.gov/Blast.cgi/).

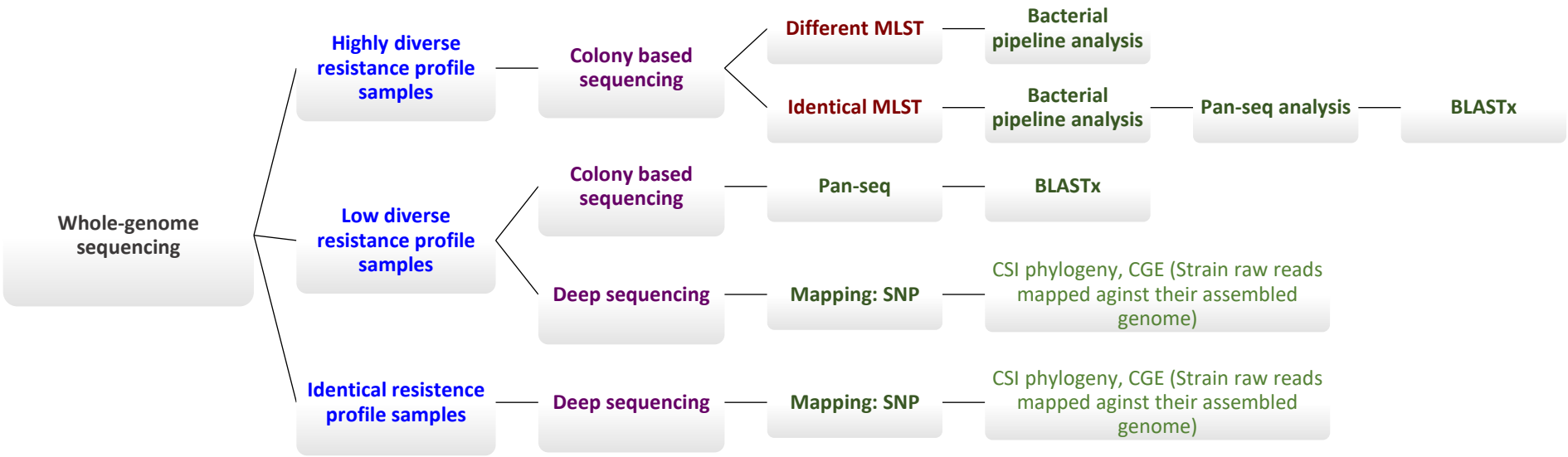


Figure 5.1: Overview of the whole-genome sequencing analysis pipelines.

5.3 Results

5.3.1 Colony-based genome sequencing of highly diverse resistance profile samples

Whole-genome sequencing was performed on twenty-two DNA bacterial isolates extracted from the nine samples of the highly diverse resistance profile group.

5.3.1.1 Quality metric assessment and MLST of highly diverse resistance profile samples

The assembled quality metrics and MLST of all genomes from the highly diverse resistance profile group are shown in table 5.2.

Table 5.2 MLST and quality metrics assessment of colony-based genome sequencing of highly diverse resistance profile samples. MLST sequence type was identified for each sequenced bacterial isolate and quality metric assessments were completed. The number of contigs, the total sequence length and number of contigs that cover half the reference genome (N50) were also determined, with the coverage of sequenced bacterial isolates calculated.

Sample number	Colony number	MLST	Number of contigs	Total length (bp)	N50 (bp)	Coverage
UTI (F1)	2	*M.morgani*	23	3758027	607206	41.58
	6		17	3759560	744281	41.56
	14		19	3760130	744281	41.55
	18W		20	3758396	432235	41.57
	18Y	ST73	117	5091783	275590	30.69
UTI (F2)	24	ST131	77	5285865	276327	44.49
	30		62	5280101	276327	59.02
	31		67	5288069	336308	31.61
UTI (F3)	3	ST88	118	5233663	188663	60.10
	7		105	5225951	208295	48.73
UTI (F4)	2	ST95	109	5280865	268910	29.31
	20	ST10	77	4691891	123294	36.28
UTI (F5)	19	ST131	112	5260716	287246	60.24
	28		112	5282372	285848	56.35
UTI (F6)	2	ST648	141	5322010	215375	44.74
	40		94	5183762	240192	52.27
UTI (M3)	5	ST131	124	5352019	190996	34.43
	35		94	5229679	222525	31.59
UTI (M1)	2	ST69	191	5161261	95321	13.04
	25		164	5162471	106412	13.16
UTI (M2)	18	ST12	134	5244845	191984	16.72
	31		121	5233880	191984	44.69

Sample UTI (F1) had a difference in genome size (total genome length) between all sequenced bacterial isolates (2), (6), (14) and (18W) and (18Y) of about 1 million base pairs. By using MLST, this female patient was shown to have a polymicrobial infection due to the presence of four bacterial isolates of *Morganella morgani* and one bacterial isolate of UPEC.

Patient samples UTI (F4), UTI (F6) and UTI (M3) contain bacterial isolates with a difference in genome size between 100,000 and 1 million base pairs. The largest difference in base pairs was seen in sample UTI (F4). Based on MLST, this female patient had two distinct sequence types of UPEC, ST95 and ST10 while samples UTI (F6) and UTI (M3) contained strains belonging to the same ST, ST648 and ST131 respectively.

Patient sample UTI (F5) and UTI (M2) contained strains with a difference in genome size between 10,000 and 100,000 base pairs. MLST showed that both sequenced bacterial isolates of sample UTI (F5) belong to ST131 while sample UTI (M2) bacterial isolates belong to ST12.

Sequenced bacterial isolates of patient samples UTI (F2), UTI (F3) and UTI (M1) had a difference of less than 10000 base pairs and they belong to ST131, ST88 and ST69 respectively.

5.3.1.2 Resistance and virulence gene analysis of strains from highly diverse resistance profile samples

Using the Centre for Genomic Epidemiology database, a bacterial analysis pipeline was run to compare virulence and resistance genes of sequenced bacterial isolates belonging to a single patient. The bacterial pipeline was run on five female samples and three male samples of the highly diverse resistance profile group. Full details of resistance genes, virulence genes and plasmid replicon type are shown in table 5.3.

Table 5.3 Resistance/ virulence genes of bacterial isolates belong to a single patient of highly diverse resistance profile samples. Sample ID, resistance genes, virulence genes, *FimH* typing and plasmid are included in the table. Resistance genes: Aminoglycoside resistance *aadA5*, *aac*(3)-*IId* ,*StrA*, *StrB,* aph(3')-1; *β-lactam* resistance *bla OXA*-1, *bla TEM 1B, bla TEM 1C;* Sulphonamides resistance *sul1*, *sul2*; Trimethoprim resistance *dFrA17*; Phenicol resistance *CatB3;* Tetracycline resistance *tetB*. Virulence genes: vaculating autotransporter toxin (*vat*), Increase serum survival (*iss*), Cytotoxic necrotizing factor (*cnf1*), Plasmid encoded enteroxin (*senB*), S fimbriae minor subunit (*sfaS*), Enterobactin sidephore receptor protein (*iroN*). Further, the varied resistance gene and plasmid are highlighted.

Sample number	Colony number	MLST	Resistance genes	Virulence genes	*FimH* type	Plasmid
Samples sequenced bacterial isolates have variation among all examined profiles						
UTI (F4)	**2**	**ST95**	*StrA, StrB, bla TEM 1C, Sul2*	*iss, vat, cnf1, senB, sfaS, iroN*	*FimH* 18	IncFII, IncB/O/K/Z, IncFIB Col156, ColRNAI, Col8282
	20	**ST10**	None		*FimH* 34	p0111
Samples sequenced bacterial isolates varied in resistance genes and plasmid types profiles						
UTI (F3)	**3**	**ST88**	*OXA-1, Cat B3*	Identical	Identical	
	7					Col(BS512), Col(8282), Col(MP18), Col(MG828)
UTI (F6)	**2**	**ST648**	*aac(3)IId, StrA, StrB, aph(3')-1°, tetB, Sul2, CaT A1*	Identical	Identical	IncFII, IncFIA, IncFIB, IncQ1
	40					Col(MP18)
UTI (M3)	**5**	**ST131**	*aadA5, aac(3)-IId, Sul1, dFrA17, mph(A)*	Identical	Identical	IncFII, IncFII, IncFIB,IncFIA
	35					
UTI (M2)	**18**	**ST12**	*StrA, StrB, bla TEM 1B, Sul2*	Identical	Identical	Col8282
	31					

Genotypic diversity of Uropathogenic *E. coli* colonies from a single patient

Sample number	Colony number	MLST	Resistance genes	Virulence genes	*FimH* type	Plasmid
Samples sequenced bacterial isolates have variation in plasmid types						
UTI (F2)	**24**	**ST131**	Identical profiles			Col(BS512), Col(MG828)
	30					
	31					
UTI (M1)	**2**	**ST69**	Identical profiles			Col(BS512)
	25					
Samples sequenced bacterial isolates have identical profile						
UTI (F2)	**19**	**ST131**	Identical profiles			
	28					

Table 5.3 displays the differences in resistance and virulence gene content of bacterial strains isolated from single patients. Samples are categorised into groups based on obtained variation among examined profiles (resistance genes, virulence genes, *FimH* type and plasmid type). Firstly, strains that varied among all examined profiles. Female sample UTI (F4) exhibited a diverse genetic content of acquired resistance genes, virulence genes, *FimH* adhesion type and plasmid type. This sample has two different sequence types, and so can be defined as a polyclonal infection.

Secondly, sample bacterial isolates that only varied in resistance genes and plasmid type. Samples UTI (F3), UTI (F6), UTI (M3) and UTI (M2) belong to this category and all sample exhibited monoclonal diversity as these sequenced bacterial isolates of each sample have the same sequence type. Further, samples UTI (F2) and UTI (M1) sequenced bacterial isolates vary only in plasmid type. These two samples also exhibited also monoclonal diversity. Lastly, sample UTI (F5) bacterial isolates have identical genetic content.

Strains of samples UTI (F4), UTI (F6) and UTI (M3) showed variation in Inc plasmid type. Samples UTI (F2), UTI (F3), UTI (F4), UTI (F6), UTI (M3), UTI (M1) and UTI (M2) showed variation in presence of a Col plasmid replicon.

Sample UTI (F4) is the only patient that displays polyclonal diversity, meaning that several patients contain strains having the same sequence type but clear phenotypic variation and potentially differences in gene and plasmid content. To further investigate the levels of within clone diversity, we used Pan-seq analysis on the remaining highly diverse samples to fully characterise differences in gene content between strains from the same patient.

5.3.1.3 Pan-seq genome analysis of strains with identical MLST from highly diverse resistance samples

Pan-seq genome analysis was run on the sequenced bacterial isolates of highly diverse resistance profile samples that have identical sequence type. The analysis was run on four female samples and three male samples. Pan-seq genome analysis can help in defining the unique regions in one bacterial isolate with respect to the other. The putative function of the novel region is then identified using BLASTx (Laing et al., 2010). Novel regions with size less than 500bp has been removed. The identified novel regions obtained between sequenced bacterial isolates of a single patient are shown in tables (5.4 – 5.10).

Table 5.4: Identified novel regions of UTI (F2) bacterial isolates using Pan-seq analysis. Strain specific genes, related to the novel regions, were identified and their putative functions established.

UTI (F2)			
Gene	**Putative Function**	**Gene**	**Putative Function**
Sequenced bacterial isolate (24) with respect to isolate (31)		**Sequenced bacterial isolate (24) with respect to isolate (31)**	
Large Terminase (542bp)			
Replication protein (608bp)	Bacterial replication	Replication protein (1134bp)	Bacterial replication
Hypothetical protein (704bp).		Hypothetical protein (638bp)	
Mobilization protein (1658bp).		Mobilization protein (1077bp)	
Helix-turn-helix domain of transposase IS66 family protein (2690bp)	DNA binding	Helix-turn-helix domain of transposase IS66 family protein (1913bp)	DNA binding
Colibactin non-ribosomal peptide synthetase ClbJ (3833bp)	DNA damage in the host (1)	Colibactin non-ribosomal peptide synthetize ClbJ (3833bp)	DNA damage in the host (1)
Sequenced bacterial isolate (30) with respect to isolate (24)		**Sequenced bacterial isolate (30) with respect to isolate (31)**	
Hypothetical protein (784bp)		Hypothetical protein (935bp), (792bp)	
Acyl-CoA dehydrogenase (773bp)	Fatty acid and amino acid catabolism	Acyl-CoA dehydrogenase (774bp)	Fatty acid and amino acid catabolism
Plasmid replication protein (2946bp)	Bacterial replication	Plasmid replication protein (2946bp)	Bacterial replication
Sequenced bacterial isolate (31) with respect to isolate (24)		**Sequenced bacterial isolate (31) with respect to isolate (30)**	
Hypothetical protein (618bp)		DNA packaging protein (516bp)	Replication
Lipid transfer protein (1142bp)	Lipid transfer activity	Hypothetical protein (734bp), (1570)	
Hypothetical protein (1479bp)			
Integrase core domain-containing protein (9011bp)	DNA integration	Integrase core domain-containing protein (11596bp)	DNA integration

(1): (Nougayrède et al, 2006)

Table 5.5: Identifed novel regions of UTI (F3) bacterial isolates using Pan-seq analysis. Strain specific genes, related to the novel regions, were identified and their putative functions established. Possible varied phenotype related to the gene was also mentioned.

UTI (F3)		
Gene	**Putative Function**	**Linked phenotype to identified gene**
Sequenced bacterial isolate (3) with respect to isolate (7)		
Replication protein (854bp)	Bacterial replication	
Class D β-lactamase OXA-1 (1001bp)	Antibiotic resistance	Resistant to β -lactam
Chloramphenicol acetyltransferase (1490bp)	Antibiotic resistance	Resistant to Phenicol
Colanic acid exporter (3135bp)	Extracellular polysaccharide	Biofilm matrix formation [(1)]
Hypothetical protein (5412bp)		
Mercury (II) reductase (8684bp)	Mercury resistance	
Sequenced bacterial isolate (7) with respect to isolate (3)		
Hypothetical protein (509bp)		
Conjugal transfer protein (610bp)	DNA transfer channel	
Conjugal transfer protein TraG (783bp)	DNA transfer channel	
Hypothetical protein (827bp), (742bp), (959bp)		
Uncharacterised protein (1001bp)		
Replication protein (1435bp), (2565bp)	Bacterial replication	

(1): (Soto et al., 2007)

Table 5.6: Identified novel regions of UTI (F5) bacterial isolates using Pan-seq analysis. Strain specific genes, related to the novel regions, were identified and their putative functions established.

UTI (F5)	
Gene	**Putative Function**
Sequenced bacterial isolate (19) with respect to isolate (28)	
ATP-dependent RNA helicase DbpA (1159bp)	Bacterial enzyme (RNA binding)
Hypothetical protein (2672bp)	
Dipepetidase E (5551bp)	Bacterial enzyme (Dipeptide cleavage)
Structural protein (9446bp)	
Sequenced bacterial isolate (28) with respect to isolate (19)	
Oligopeptide ABC transporter substrate-binding protein OppA (788bp)	Protein and peptide transport
Pitrilysin (1004bp)	Bacterial enzyme
Universal stress protein UspE (1479bp)	Bacterial survival (1)
Amino acid permease (2391bp)	Amino acid transporter
Histidne histamine antiroperter (3607bp)	Antiporter
Phospho-2-dehydro-3-deoxyheptonate aldolase (4075bp)	Amino acid biosynthesis
Hypothetical protein (5650bp)	

(1): (Nachin et al., 2005).

Table 5.7: Identified novel regions of UTI (F6) bacterial isolates using Pan-seq analysis. Strain specific genes, related to the novel regions, were identified and their putative functions established. Possible varied phenotype related to the gene is also mentioned.

UTI (F6)		
Gene	**Putative Function**	**Linked phenotype to identified gene**
Sequenced bacterial isolate (2) with respect to isolate (40)		
Protein ImpB (707bp)	DNA repair	
Conjugal transfer protein TraJ (736bp)	DNA transfer channel	
Amino acid binding protein (833bp)	Amino acid transport	
Aminoglycoside 3'phosphotransferase APH(3') (1038bp)	Antibiotic resistance	Resistant to aminoglycoside
Transposase (1277bp)	Transposition (Mobile genetic element insertion)	
Hypothetical protein (1388bp)		
IcIR family trancriptional regulator (1646bp)	Multi drug resistance and virulence	
IS21 family transposase (1954bp)	Transposition (Mobile genetic element insertion)	
Tetracycline efflux MFS transporter Tet B (2781bp)	Antibiotic resistance	Resistant to tetracycline
Metal ABC transporter substrate binding protein (3401bp)	Metal binding	
Replication protein C (4385bp)	Bacterial replication	
Hypothetical protein (19752bp), (5904bp)		
Tn3 family transposase (5028bp)	Transposition (Mobile genetic element insertion)	
IS1182 family transposase (6974bp)	Transposition (Mobile genetic element insertion)	
ABC transporter substrate-binding protein (7654bp)	Protein binding	
Iron permease (10778bp).	Bacterial pathogenesis	Iron acquisition
Ton B dependent sidephore receptor (11555bp)	Bacterial pathogenesis	Iron acquisition
Sequenced bacterial isolate (40) with respect to isolate (2)		
Replication protein (506bp)	Bacterial replication	
RopB/ MobA like protein (642bp)	Plasmid mobilization	
Plasmid mobilisation protein (714bp)	Plasmid mobilization	
Tail fiber assembly protein (806bp)	Formation tail fiber	
Exonuclease family protein (2170bp)	DNA repair	
DNA-binding protein (3161bp)	Bacterial replication	
RHS Repeat family protein (3398bp)	Polysaccharide synthesis	

Table 5.8: Identified novel regions of UTI (M3) bacterial isolates using Pan-seq analysis. Strain specific genes, related to the novel regions, were identified and their putative functions established. Possible varied phenotype related to the gene is also mentioned.

UTI (M3)		
Gene	**Putative Function**	**Linked phenotype to identified gene**
Sequenced bacterial isolate (5) with respect to isolate (35)		
Hypothetical protein (653bp)		
DNA polymerase V (707bp)	DNA repair	
Hypothetical protein (1006bp)		
Type VI secretion protein IcmF (1157bp)	Bacterial survival and replication	
Phage tail protein (1302bp)	Tail formation	
Enolase (1986bp)	Glycolysis	
Aminoglycoside N(3') acetyltransferase (2598bp)	Antibiotic resistance	Resistant to aminoglycoside
IS66 family transposase (2708bp)	Transposition (Mobile genetic element insertion)	
Ton B dependent receptor (2983bp)	Bacterial pathogenesis	Iron acquisition
IS110 family transposase (3746bp)	Transposition (Mobile genetic element insertion)	
Hypothetical protein (6557bp)		
Hypothetical protein (PapB)	Bacterial virulence	Adhesion[(1)]
Conjugative transfer relaxase/helicase TraI	DNA transfer	
Group II intron reverse transcriptase/ maturase (24619bp)	Mobile genetic element	
Tn3 family transposase: 2 plasmid (7218bp)	Transposition (Mobile genetic element insertion)	
YfcC family protein (8503bp)	Putative amino acid antiporter	
Chromate transporter (10932bp)	Chromate resistance	
Sequenced bacterial isolate (35) with respect to isolate (5)		
Non-ribosomal peptide synthetase (647bp)	Non-ribosomal peptide synthesis	
Efflux ABC transporter permease protein (793bp)	ATP/ nucleotide binding	
Transposase (3429bp)	Transposition (Mobile genetic element insertion)	

(1): (Leffler and Svanbrog-Eden, 1980; Lund et al., 1987)

Table 5.9: Identified novel regions of UTI (M1) bacterial isolates using Pan-seq analysis. Strain specific genes, related to the novel regions, were identified and their putative functions established.

UTI (M1)	
Gene	**Putative Function**
Sequenced bacterial isolate (2) with respect to isolate (25)	
Head completion/ stabilisation protein (2762bp)	Phage replication
Replication protein (3921bp)	Plasmid replication
Sequenced bacterial isolate (25) with respect to isolate (2)	
Filamentous hemagglutinin family domain-containing protein (528bp)	Transport
Type-1 restriction enzyme EcoKI specificity protein (4040bp)	DNA restriction

Table 5.10: Identified novel regions of UTI (M2) bacterial isolates using Pan-seq analysis. Strain specific genes, related to the novel regions, were identified and their putative functions established. Possible varied phenotype related to the gene is also mentioned.

UTI (M2)		
Gene	**Putative Function**	**Linked Phenotype to identified gene**
Sequenced bacterial isolate (18) with respect to isolate (31)		
Toxin YafO (776bp)	Protein synthesis inhibitor	
RepB family plasmid replication initiator protein (1042bp)	Bacterial replication	
Acetyl-CoA acetyltransferase (1309bp)	Fatty acid and lipid catabolism	
TEM family class A β -lactmase (1916bp)	Antibiotic resistance	Resistant to β -lactam
Hypothetical protein (2269bp)		
Str A (3591bp)	Antibiotic resistance	Resistant to aminoglycoside
Sequenced bacterial isolate (31) with respect to isolate (18)		
Flagellar protein (2347bp)	Bacterial virulence [(1)]	Binding /motility
DNA recombinase (2468bp)	DNA exchange	
6-phosphogluconate phosphatase (2846bp)	Bacterial enzyme (Metal binding)	
Cobalt-zinc-cadmium resistance protein (3111bp)	Metal resistance	

(1): (Partridge et al., 2015)

Female patients UTI (F2), UTI (F3), UTI (F5), UTI (F6) and male patient UTI (M3) sequenced bacterial isolates were genetically diverse in term of existed gene contents. Sequenced isolates of each one of these samples contained different genetic fragments of more than 5kbp. Female sample UTI (F2) sequenced bacterial isolates differed in replication protein, plasmid protein and some enzymes. Sample UTI (F3) isolate (3) has mercury reductase and class D β-lactamase, while the other isolate (7) has a unique replication protein and conjugal transfer protein. The third female sample UTI (F5) isolate (19) has a large fragment of structural protein, while isolate (28) has UspE, which has been shown to play a role in phenotypic functions such as motility and adhesion and helps in bacterial survival (Nachin et al., 2005).

In sample UTI (F6), isolate (2) has unique proteins related to conjugal transfer protein, tetracycline efflux transporter, metal binding protein and Ton B dependent siderophores. Isolate (5) in sample UTI (M3) contained a unique type VI secretion protein IcmF, Ton B dependent receptor, *papB* protein and Tn3 transposes which contained two plasmid proteins. These two samples show the largest level of genetic diversity in bacteria studied that belong to the same clone type.

The genetic variation in isolates of samples UTI (M1) and UTI (M2) occurred at similar levels. In male sample UTI (M1), sequenced bacterial isolate (2) has a unique replication protein, while the other sequenced bacterial isolate (25) has a type-1 restriction enzyme Ecor1. In male sample UTI (M2), sequenced bacterial isolate (18) has TEM family class A β-lactam and StrA, while bacterial isolate (31) has a flagellar protein and cobalt-zinc-cadmium resistance protein.

Sequenced bacterial isolates of the same sequence type among highly diverse samples are varied in proteins related to plasmids, such as replication protein, conjugal transfer

protein, IS66, IS110, binding protein, recombinase proteins and IcIR. IS66 and IS110 are responsible for horizontal transmission and acquiring mobile genetic elements which are responsible for bacterial evolution (Han et al., 2001; Tobes and Pareja, 2006). Genetic variation between sequenced bacterial isolates from the same patient also include variation in proteins responsible for drug and metal resistance such mercury reductase, β-lactamase, metal ABC binding protein, tetracycline efflux system, TEM family and metal resistance protein. Also, several iron uptake systems are varied between bacterial isolates such as Ton B protein and iron permease.

5.3.2 Whole-genome sequencing of low diverse resistance profile samples

Whole-genome sequencing was applied on twenty samples of the low diverse resistance profile patterns. Colony-based sequencing was applied on two bacterial isolated from one low diverse sample and deep sequencing was applied on nineteen randomly selected samples.

5.3.2.1 Quality metric assessment and MLST of the low diverse resistance profile samples

The assembly quality metrics and MLST for the whole-genome sequencing of the low diverse resistance profile sample are shown in table 5.11.

Table 5.11 MLST and quality metrics assessment of colony-based and pooled genome sequencing of low diverse resistance profile samples. MLST sequence type was identified for each sequenced bacterial isolate and quality metric assessments were completed. The number of contigs, the total sequence length and number of contigs that cover half the reference genome (N50) were also determined, with the coverage of sequenced bacterial isolates calculated.

Sample number		MLST	Number of contigs	Total length (bp)	N50 (bp)	Coverage
Colony-based sequencing						
UTI (F17)	**13**	**ST10**	97	4772971	178642	32.74
	20		108	4779629	174317	32.69
Deep sequencing						
UTI (F12)		**ST131**	1571	6496694	225002	24.05
UTI (F9)		**ST355**	128	4906722	274347	31.84
UTI (F7)		**ST404**	127	5328251	245659	29.32
UTI (F8)		**ST10**	101	4837046	203697	32.30
UTI (F11)		**ST58**	111	4969202	123969	31.44
UTI (F16)		**ST404**	146	5341348	196359	29.25
UTI (F13)		**ST73**	73	5024121	395954	31.10
UTI (F15)		**ST69**	202	5283128	137122	29.58
UTI (F19)		**ST131**	142	5311854	227152	29.42
UTI (M5)		**ST12**	75	5020826	324524	31.12
UTI (M4)		**ST12**	72	4989540	272338	31.32
UTI (M7)		**ST12**	173	5309323	198138	29.43
UTI (M10)		**ST12**	106	5101415	197068	30.63
UTI (M9)		**ST131**	126	5265576	209516	29.67
UTI (M6)		**ST131**	75	4920096	201241	31.76
UTI (M11)		**ST648**	123	5270316	210219	29.65
UTI (M12)		**ST73**	71	5036192	395957	31.03
UTI (M16)		**ST95**	147	5333419	286696	29.30
UTI (M8)		**ST80**	118	5093302	316363	30.68

5.3.2.2 Colony-based sequencing of bacterial isolates of the low diverse resistance sample

Isolates obtained from female sample UTI (F17) had identical sequence type. Pan-seq and BLASTx analysis was applied on the two sequenced bacterial isolates and revealed some variation in presence of hypothetical proteins and bacterial enzymes. Strain (13) has ribosomal protein, nuclease and iron permease while strain (20) has initiator replication protein, transposase and oxygenase as shown in table 5.12.

Table 5.12: Identified novel regions of UTI (F17) bacterial isolates using Pan-seq analysis. Strain specific genes, related to the novel regions, were identified and their putative functions established.

UTI (F17)	
Gene	**Putative Function**
Sequenced bacterial isolate (13) with respect to isolate (20)	
Nuclease (713bp)	DNA cleavage
Iron ABC transporter permease (1142bp)	Bacterial virulence
30S ribosomal protein S12 methylthiotransferase RimO (1371bp)	Ribosomal proteins transferase
Hypothetical protein (1993bp)	
Sequenced bacterial isolate (20) with respect to isolate (13)	
Transposase family protein (549bp)	Transposition (Mobile genetic element insertion)
β -carotene 15,15 mono-oxygenase (782bp)	Bacterial enzyme
Transposase (1655bp)	Transposition (Mobile genetic element insertion)
YfcC family protein (1714bp)	Amino acid antiporter

5.3.2.3 Deep sequencing of the low diverse resistance profile samples

Deep sequencing was carried out on 19 samples of the low diverse resistance profile. A sweep of bacterial growth was taken and DNA extracted and sequenced to detect the existing SNPs within the bacterial population. To do this, raw reads were mapped against their *de novo* assembled genome using CSI phylogeny on the CGE website (https://cge.cbs.dtu.dk/services/CSIPhylogeny/) (version 1.4); (section 2.7.3, **Chapter** 2). SNPs were determined with base quality > 30 and minimum depth of 10. SNPs selected based on allele frequency calculated from DP4 statistics. DP4 provides information about the forward reference allele, reverse reference allele, forward non-reference allele and reverse non-reference allele reads coverage. Tables 5.13 display the number of mutation/s that found, whether these mutations are synonymous or non-synonymous, and the gene function. Generated variant call format (VCF) files output has been displayed in **Chapter 7**.

Table 5.13 Detected SNPs within low diverse sequenced samples. The number of mutation found in each sample, whether these mutations are synonyms (S) or non-synonyms (NS), and their location.

Sample number	Number of mutations	S/NS	Function
UTI (F12)	2	NS	Hypothetical protein Manganse transport membrane protein MntB
	3	S	Coenzyme A biosynthesis bifunctional protein CoaBC Carboxylate amine ligase YbdK Small toxic polypeptide LdrD
UTI (F9)	2	NS	Hypothetical protein
UTI (F7)	7	NS	Actin cross-linking toxin VgrG1 Hypothetical protein (5) Putative autotransporter
	2	S	Actin cross-linking toxin VgrG1 Antigen 43
UTI (F8)	9	NS	Manganse transport membrane protein MntB (6) Transposase IS1200 like protein tRNA nuclease precursor Phage DNA packaging protein Nu1
UTI (F11)	1	NS	Fructose specific phosphotransferase enzyme lib component
UTI (F16)	6	NS	Putative type II secretion system protein Putative autotransporter precursor Phage related baseplate assembly protein Hypothetical protein (2) Glutamate decarboxylase beta
	2	S	Citrate lyase subunit beta Hypothetical protein
UTI (F13)	3	NS	Phage related baseplate assembly protein Adhesion YadA precursor
UTI (F15)	6	NS	Putative transposase (2) Hypothetical protein (2) gpW Fibronectin type III protein
	3	S	Exodeoxyribonuclease Phage terminase large subunit (GpA) Group II intron encoded protein LtrA

Genotypic diversity of Uropathogenic *E. coli* colonies from a single patient

Sample number	Number of mutations	S/NS	Function
UTI (F19)	5	NS	Antirestriction protein K1cA (2) Putative autotransporter precursor Hypothetical protein IS1 transposase
	3	S	Small toxic polypeptide LdrD (2) Transposase IS66 family protein
UTI (M5)	3	NS	Hypothetical protein (2) DsdX permease
	1	S	Putative autotransporter precursor
UTI (M4)	4	Ns	Electron transport complex protein RnfC (3) Primosomal protein N'
UTI (M7)	4	NS	Hypothetical protein (2) Filamentous hemagglutinin (2)
	2	S	Minor fimbrial protein PrsF precursor Hypothetical protein
UTI (M10)	2	NS	Integrase core domain protein D epimerase tagatose 3
UTI (M9)	6	NS	Terminase like family protein (3) Hypothetical protein (2) Phage tail fibre repeat protein
	4	S	Nicotinate phosphoribosyltransferase Small toxic polypeptide LdrD (2) S transposase 1
UTI (M6)	None		
UTI (M11)	1	NS	IS1 transposase
	3	S	Hypothetical protein (2) Integrase core domain protein
UTI (M12)	1	NS	C1amp binding protein CrfC
UTI (M16)	6	NS	IncFII repA protein family (2) Bifunctional protein Electron transport complex protein RnfC (3)
	1	S	DNA primase TraC
UTI (M8)	2	NS	Integrase core domain protein Filamentous hemagglutinin

Deep sequencing revealed a limited number of mutations. The number of investigated non-synonymous mutations within each sample fewer than 10 mutations. All deep sequenced samples reveal microvariation which reflect low levels of within-host diversity.

5.3.3 Whole-genome sequencing of identical resistance profile samples

Deep sequencing was carried out on the five samples showing identical resistance profile. A sweep of bacterial growth was harvested and DNA extracted and sequenced to detect the existing SNPs within the bacterial population.

5.3.3.1 Quality metric assessment and MLST of the identical resistance profile samples

The assembly quality metrics and MLST for the whole-genome sequencing of the identical resistance profile samples are shown in table 5.14.

Table 5.14 MLST and quality metrics assessment of whole genome sequencing of identical resistance profile samples. MLST sequence type was identified for each sequenced bacterial isolate and quality metric assessments were completed. The number of contigs, the total sequence length and number of contigs that cover half the reference genome (N50) were also determined, with the coverage of sequenced bacterial isolates calculated.

Identical resistance profiles samples					
Sample number	**MLST**	**Number of contigs**	**Total length (bp)**	**N50 (bp)**	**Coverage**
UTI (F20)	**ST131**	121	5142915	180644	30.38
UTI (F21)	**ST69**	139	5213889	175003	29.97
UTI (M21)	**ST131**	366	5634478	188745	27.73
UTI (M19)	**ST131**	143	5404293	202819	28.91
UTI (M20)	**ST681**	94	5168921	196585	30.23

5.3.3.2 Deep sequencing of the identical resistance profile samples

Deep sequencing was carried out on the identical resistance profile samples to detect SNPs within population. Table 5.15 display the number of mutations found in generated VCF files, whether these mutations are synonymous or non-synonymous, and the gene function. Generated variant call format (VCF) files output including all SNPs are displayed in **Chapter 7**.

Table 5.15 Detected SNPs within identical resistance sequenced samples. The number of mutations in each sample, weather these mutations are synonymous (S) or non-synonymous (NS), and their location.

Sample number	Number of mutations	S/NS	Function
UTI (F20)	1	NS	Phage tail fibre repeat protein
UTI (F21)	2	NS	gpW Peptidase S49 family
	1	S	Phage terminase large subunit (GpA)
UTI (M21)	4	NS	Hypothetical protein (3) Dihydrolipollysine acetyl transferase component residue pyruvate dehydrogenase complex
	2	S	Small toxic polypeptide LdrD Ptative autotransporter precursor
UTI (M19)	4	NS	Invasin Small toxic polypeptide LdrD Phage portal protein Integrase core domain protein
	2	S	Small toxic polypeptide LdrD Integrase core domain protein
UTI (M20)	2	NS	Integrase core domain protein Hypothetical protein

Deep sequencing revealed a limited number of mutations. The number of investigated non-synonymous mutations within each sample was fewer than 10 mutations. All identical resistance profile samples reveal microvariation which reflects low levels of within-host diversity.

5.3.4 Summary table linking genotypic with phenotypic variation

Phenotypic and genotypic diversity level can be assessed initially on the variation of antibiotic susceptibility. This is recognised in our highly diverse resistance profile samples. One patient showed that the phenotypic variation is linked with having a polyclonal diversity (two sequence types). The remaining seven samples proved to have a monoclonal diversity linked with further phenotypic and genetic variation. Linked phenotypic characteristics and genotype profile are shown in table 5.16.

To assess the phenotypic and genotypic diversity level of each sample, in comparison to other bacterial isolates which share a susceptibility profile, scoring analysis was performed through chi-square analysis using Graphpad prism software. A chi-squared test was chosen and used in order to assess how likely the observations are assuming that the null hypothesis is true (Campbell, 2007). The diversity levels of four samples within the highly diverse group were predicted as a majority of the examined colonies shared a susceptibility pattern. There were three different resistance profiles noted within the UTI (F3) sample. The phenotypic and genotypic diversity level for two of the bacterial isolates were representative of two different profile patterns. Bacterial isolates (3) shared the resistance profile with 44 other bacterial colonies (91.67 %), assuming that those colonies comprise the same level of phenotypic and genotypic diversity. However, bacterial isolate (7) was the only susceptible isolate among examined well-isolated colonies. UTI (F4) also has three different resistance profiles among the examined well-isolated colonies. The resistant isolate (2) shared a resistance profile with 16 bacterial isolates (72.73 %), assuming that those bacterial isolates belong to ST 95. Comparatively, bacterial isolate (20) was the only susceptible isolate that was found amid well-isolated colonies within the sample. Further research could perform the same

experiment using a higher dose of the varied antibiotic in order to assess whether this has an effect upon the susceptibility to resistance. UTI (F6) showed five different resistance profile patterns. Resistant bacterial isolate (2) shared the resistance profile with 34 other bacterial isolates (85 %), again assuming that those well-isolated colonies have the same diversity level. The other bacterial isolate is the only sensitive isolate found within the sample. Furthermore, UTI (M1) had two different resistance profile patterns. The resistance isolate (5) shared a resistance profile with 38 bacterial isolates (95 %), again assuming an identical diversity level. Likewise, bacterial isolate (35) was the only susceptible isolate that was found within this population. Within all of these mentioned samples, phenotypic and genotypic diversity level can be, and was, predicted, as approximately 80 % of well-isolated colonies can be expected.

The remaining four samples were not associated with a common resistance profile, a feature which could potentially be accompanied by a diverse phenotype and genotype within the population. A diversity level of 15 % was predicted with respect to sample UTI (F2), this was based on the presence of other bacterial isolates which share a resistance pattern with the examined bacterial colonies. Sample UTI (F5) resulted in a diversity level of 35.41 %, this was determined due to the shared resistance pattern amongst examined well-isolated colonies. Penultimate, a diversity level of 22.23 % was observed in sample UTI (M1) as there were 11 different resistance profile patterns noted within this sample. This was comparatively higher than the predicted diversity level of 8.33 % within the final sample, UTI (M2). This was concluded as 6 different resistance patterns were observed. Thus, these samples may have a diverse bacterial population.

Table 5.16 Summary tables of highly diverse resistance sample profiles and their diversity level. This table shows the phenotypic and genotypic variation of bacterial isolates within each high diverse resistance profile samples.

Sample number		Resistance profile		Phenotypic diversity			Genotypic diversity		Diversity level
		Antibiotic	Susceptibility result	Motile ability	Biofilm formation	Association and invasion	MLST	Genetic content	
Highly diverse resistance profile samples									
UTI (F2)	24	AP, GM, CPM	R	No Phenotypic Variation			ST131	Varied plasmid type may affect resistance profile	Monoclonal diversity
	30	AP	R						
	31	GM, AP, CPD, CPM	R						
UTI (F3)	3	AP	R	None	High	None	ST88	Class D *ϐ-lactam* Colanic acid exporter	
	7	AP	S		Low			-	
UTI (F4)	2	AP	R	Motile	High	Highly adherent to T_{24}	ST95	Varied in all resistance, virulence genes and plasmid content	Polyclonal diversity
	20	AP	S	Non motile	Low	Less adherent to T_{24}	ST10	-	
UTI (F5)	19	CAZ	R	No Phenotypic Variation			ST131	Structural protein may affect resistance profile	Monoclonal diversity
	28	CAZ	S					-	
UTI (F6)	2	GM	R	None		Highly invasive to T_{24}	ST648	Aminoglycoside resistance *aadA5*, *aac*(3)-*IId*	
	40	GM	S			Less invasive to T_{24}		-	
UTI (M3)	5	GM and TM	R	None		Highly invasive to T_{24}	ST131	Aminoglycoside resistance *aadA5*, *aac*(3)-*IId* Trimethoprim resistance *dFrA17* *PapB*	
	35	GM and TM	S			Less invasive to T_{24}		-	
UTI (M1)	2	CIP	R	No Phenotypic Variation			ST69	Replication protein may affect resistance profile	
	25	CIP	S					-	
UTI (M2)	18	AP	R	No Phenotypic Variation			ST12	*ϐ-lactam* resistance *bla TEM 1B*	
	31	AP	S					-	

Table 5.17 Scoring analysis for the examined bacterial isolate of each sample. This table shows scoring analysis and represented as a percentage of having phenotypic and genotypic diversity based on the examined bacterial isolate that shared other bacterial isolate shared the susceptibility profile.

Sample number	Number of different resistance profile/s found within sample	Examined bacterial isolate within sample	Resistance profile			
			Varied antibiotic	Susceptibility result	Percentage of the examined bacterial isolate	Percentage of other bacterial isolates having identical resistance profile/s pattern and may have the same phenotypic and genotypic variation level
UTI (F2)	9	24	AP, GM, CPM	R	2.50 %	-
		30	AP	R	2.50 %	7.5 %
		31	GM, AP, CPD, CPM	R	2.50 %	-
UTI (F3)	3	3	AP	R	2.08 %	91.67 %
		7	AP	S	2.08 %	-
UTI (F4)	3	2	AP	R	4.55 %	72.73 %
		20	AP	S	4.55 %	-
UTI (F5)	9	19	CAZ	R	2.08 %	-
		28	CAZ	S	2.08 %	31.25 %
UTI (F6)	5	2	GM	R	2.50 %	85.00 %
		40	GM	S	2.50 %	-
UTI (M3)	2	5	GM and TM	R	2.50 %	95 %
		35	GM and TM	S	2.50 %	-
UTI (M1)	11	2	CIP	R	2.78 %	-
		25	CIP	S	2.78 %	16.67 %
UTI (M2)	6	18	AP	R	2.08 %	4.17 %
		31	AP	S	2.08 %	-

The other two groups of our samples (low and identical resistance profiles) proved to have no phenotypic variation with low level of genetic diversity as shown in table 5.18.

Table 5.18 Summary tables of low and identical resistance sample profiles and their diversity level. This table shows phenotypic and genotypic variation among low diverse and identical resistance profiles patterns.

Low diverse resistance profile samples				
Sample number	**Resistance profile**	**Phenotypic diversity**	**Genotypic diversity**	**Diversity level**
UTI (F12)	Varied in antibiotic susceptibility either from sensitive or resistant to intermediate zone of inhibition	Some of these samples are examined phenotypically and showed no variation	Low SNPs number is detected	Low diversity level
UTI (F9)				
UTI (F7)				
UTI (F8)				
UTI (F11)				
UTI (F16)				
UTI (F13)				
UTI (F15)				
UTI (F19)				
UTI (M5)				
UTI (M4)				
UTI (M7)				
UTI (M10)				
UTI (M9)				
UTI (M6)				
UTI (M11)				
UTI (M12)				
UTI (M16)				
UTI (M8)				
Identical resistance profile samples				
UTI (F20)	Identical resistance profile is detected by all examined well-isolated colonies	No Phenotypic variation is detected	Low SNPs number is detected	Low diversity level
UTI (F21)				
UTI (M21)				
UTI (M19)				
UTI (M20)				

5.4 Discussion

Variation in resistance profiles and phenotypic characteristics within a single UPEC sample (Chapters 3 and 4) may also be related with further diversity of genetic content between bacterial isolates within the sample. Using genomic approaches facilitates detection of varied genetic content within groups of organisms. Acquired genetic elements or genome alterations have a role in pathogen evolution and in within-host adaptation (Dobrindt, 2005). In this chapter, within-host genetic diversity of UPEC was investigated using whole-genome sequencing technology.

Whole-genome sequencing was performed on seventeen bacterial isolates from eight highly diverse resistance patient samples. Sequenced bacterial isolates from the same patient varied in genome size and genetic content. Differences in genome size between bacterial isolates of the same patient may reveal variation of flexible genome size that may be affected by acquisition or loss of genetic material (Dobrindt, 2005).

Presence of polyclonal diversity was shown in one female sample UTI (F4) infected with two different sequence types of UPEC. This explains the variation in resistance profile and phenotypic assays described in chapters 3 and 4, and the difference in genome size, resistance genes, virulence genes and *FimH* adhesion type in this chapter. Having polyclonal diversity (multi sequence types) within a single patient could be due to infection with multiple clones of UPEC or having multiple transmissions over time, as shown in a *B. dolosa* evolution study by Lieberman et al. (2011) and also reported in mother infected with multiple clones of *H. pylori* (Kivi et al., 2007). Further, it was reported that mixed colonization within the host may result from two separate transmission events (Didelot et al., 2016).

Sequenced isolates from a female sample UTI (F6) and from a male sample UTI (M3) had a similar sequence type and are defined as having monoclonal diversity. Isolate (2) of sample UTI (F6) contained unique genetic content compared to the other isolate from this patient and most of the varied genetic content is related to gene acquisition (conjugal transfer and IcIR). In chapter 3, isolate (2) is resistant to aminoglycoside which be linked with presence of aminoglycoside 3'phosphotransferase APH(3') gene in this chapter. Further, isolate (2) characterised with high invasion ability to T24 cell line in chapter 4 and could be linked with presence of varied transposase genes and IcIR family transcriptional regulator which could be responsible for virulence. Furthermore, sequenced isolate (5) of sample UTI (M3) has varied gene regions that are related to antibiotic resistance and bacterial virulence. Presence of Aminoglycoside N(3') acetyltransferase, Aminoglycoside resistance *aadA5*, *aac*(3)-*IId* and Trimethoprim resistance *dFrA17* genes explain resistance ability to GM and TM in chapter 3. In addition, presence of *pap* region in this strain may explain having high invasion ability to T24 cell line. *Pap* may associate with ascension of UTI to kidneys. Having the *pap* operon is important in ascending the UTI and is identified in 80 % of pyelonephritis patients (Kucheria et al., 2004). Such diversity could be explained by long-term infection persistence of a single clone and subsequent reinfection with new bacterial cells of the same clone.

The other five samples of our highly diverse resistance profile exhibit monoclonal diversity due to having the same sequence type of UPEC. Those samples have smaller scale genetic differences between patient sequenced isolates. Horizontal gene transfer may provide the bacterial isolates with a specific trait and play role on pathogen host adaption and have a significant in bacterial evolution (Dobrindt, 2005; Brzuszkiewicz et

al., 2006; Didelot et al., 2016). These samples reveal within host monoclonal diversity and could be related to the long-term persistence of UPEC within the host. Sample UTI (F2) sequenced isolates differ in replication proteins. Sample UTI (F3), bacterial isolate (3) has class D β-lactamase OXA-1 which is responsible for the different antibiotic resistance reported in chapter 3. Also, colanic acid exporter gene is found as a novel region in this strain which may be linked with biofilm formation ability in chapter 4. Lastly, sample UTI (M2) isolate (18) has TEM family class A β -lactamase which is responsible for AP resistance in chapter 3.

Generally, varied genetic content within the highly diverse resistance sample isolates is related to proteins thought to be associated with pathogenicity islands. According to Dobrindt (2005), the pathogenicity island of *E. coli* isolates is thought to carry virulence genes related to iron uptake and adhesion and to varied plasmid effects on drug resistance (Dobrindt, 2005). Variation in these types of genes has been confirmed in this research among our highly diverse resistance samples.

During bacterial infection, pathogens can accumulate mutations that allow adaptation with the host, evading the immune system and resisting antibiotic treatment. Evolution at the gene level was assessed in Lieberman et al. (2011), who found a significant correlation between non-synonymous mutation and gene evolution after initial infection. That study revealed that non-synonymous mutation causes alteration in fluoroquinolone resistance and has a role in pathogenesis (Lieberman et al., 2011). In this research, deep sequencing of low diverse and identical resistance profile samples reveal a limited number of mutations and this indicates limited levels of within-host evolution. Zdziarski et al. (2010) determined a limited number of mutations from isolates of asymptomatic bacteriuria and that prolonged infection may alter genomic content to very limited

levels. These may occur as adaption mechanisms that are required for growth within the urinary tract (Zdziarski et al., 2010).

Taking all considerations together, genomic diversity within a patient infected with UPEC has been demonstrated within our collected samples can be classified into two groups: samples that showed highly diverse genetic content (highly diverse) and samples with very limited genetic diversity. The highly diverse genetic samples are represented by the highly diverse resistance profile samples but examples of polyclonal infection are very rare. Instead phenotypic and genotypic diversity appears to be driven by differential presence of loci associated with either mobile genetic element and horizontal gene transfer (Brzuszkiewicz et al., 2006; Schubert et al., 2009). Polyclonal diversity was found in one female patient and monoclonal diversity was demonstrated in the remaining highly diverse samples. The other two groups of samples (low and identical) are characterised by a limited number of nucleotide substitution.

Chapter Six

General Discussion and Further Directions

6.1 Discussion

Species diversity within a single host has been recently evaluated among various organisms such as *S. aurous, H. pylori, B. dolosa, M. tuberculosis* and *P. aeruginosa* (Kivi et al., 2007; Cespedes et al., 2005; McAdam et al., 2011; Golubchik et al., 2013; Lieberman et al., 2014; O'Neill et al., 2015; Feliziani et al., 2014; Darch et al., 2015). Variation in resistance profiles and phenotypic characters, besides genetic content, were recognised within these species in a single patient. Within patients, diversity was also indicated among patient infected with ExPEC (Levert et al., 2010).

A previous study by our research group established that multiple sequence types of ExPEC were isolated in a single patient with urosepsis when a comparison was performed between blood and urine isolates (McNally et al., 2013). As UPEC is the causative agent of UTIs, this PhD research has been conducted to evaluate the host diversity of a single patient infected with UPEC. The experimental work described here has been based on comparing apparently identical colonies of a single UPEC culture plate using sensitivity testing, various phenotypic assays and genome sequencing technology.

First, 41 samples of patients with UPEC were collected from QMC, Nottingham, UK and classified, based on variation in resistance profile between examined well-isolated colonies, into three groups: highly diverse resistance profile, low diverse resistance profile and identical resistance profile patterns. Afterwards, various phenotypic assays were applied on isolated strains from the highly diverse resistance profile pattern samples and five randomly selected colonies from each apparently identical resistance profile samples. Then, the genetic diversity was also assessed within patients of those three patterns using whole-genome sequencing technology.

Whole-genome sequencing was performed on bacterial isolates that have a varied resistance profile within samples of highly diverse resistance pattern patients, and deep sequencing was applied on 19 samples of the low diverse resistance profile and samples of identical resistance patterns.

This research data can be classified into two groups in term of having within-sample diversity. Samples have a polyclonal or monoclonal diversity, which has clearly been seen in highly diverse resistance profile samples. The other group is recognised within the other two resistance profile patterns (low and identical) and characterised extremely low levels of genomic diversity within the infecting population.

Heteroresistance was recognised in strains isolated from patients whose samples exhibited highly diverse resistance patterns, and this was the first sign of within-host diversity. Heteroresistance is also thought to be related with infection persistence (El-halfawy and Valvano, 2015). Phenotypic diversity was also evaluated via motility using semi solid agar plate, biofilm formation using crystal violet and association and invasion assays to T24 epithelial cell lines. This subsequently proved phenotypic diversity between identical colonies within a single patient of highly diverse resistance profile pattern. Variation in motility, biofilm formation and attachment and invasion abilities are mostly related with having an infection in the upper urinary tract or UTI persistence (Lane et al., 2005; Soto et al., 2007; Blango and Mulvey, 2010). Thus, patients with heteroresistance and phenotypic variation may suffer from UTI ascension or long-term infection. Polyclonal diversity was proved in one female sample due to having two different *E. coli* sequence types. The remaining highly diverse resistance profile samples showed monoclonal diversity. Variation in genetic content within a single host is a sign of UPEC evolution and thought to be

related also to host adaption (Dobrindt, 2005; Brzuszkiewicz et al., 2006). The noticed diversity within these samples could be related to the long-term persistence of UPEC within the host.

Low diversity and identical resistance profiles are the second group identified in this research. Although some change in resistance was recognised within the low diverse samples, phenotypic diversity was not recognised within this group. Similar phenotypic findings were seen in the identical resistance profile samples. Low-levels of genetic variation were determined in low diverse and identical resistance profile samples. Detected SNPs may have occurred due to UPEC adaption within the host. As illustrated by Lieberman et al. (2011), several mutations can accumulate within the pathogen during the infection with the aim of adapting to the host.

Long-term UPEC infection within a patient may drive the obtained phenotypic and genotypic diversity. Recurrent infection may influence the patient's status, and the persistent strain may acquire virulent traits that increase their ability to survive and persist within the host. In a cystic fibrosis patient infected with *S. aureus*, three sequential isolates and clonal genetic diversity were obtained related to variation resistance, growth and virulence (McAdam et al., 2011). Genetic diversity was also determined within four sequential isolates of *H. pylori* and thought to be related to a recombination event due to long term persistence in the host (Kennemann et al., 2011). Feliziani et al. (2014) examined genetic diversity in relation with prolonged existence of *P. aeruginosa*. Two patients were included in the study, and it was revealed that multiple genetic recombinations and hypermutability is related to long-term infection (Feliziani et al., 2014). The long-term existence of a pathogen within the host is thought to be related with pathogen genotypic diversity as

illustrated in *B. dolosa* by Lieberman et al. (2014). Isolates of *S. aureus* from the same individual exhibited microevolution and exchanged mobile elements. This evolution is thought to be related with prolonged carriage and duration of infection (Golubchik et al., 2013). All mentioned studies linked long-term infection persistence with pathogen within host diversity. In this research, highly diverse samples are predicted to be isolated from patients with a long history of UTIs. Further, polyclonal diversity is recognised in female patients, which could be explained by the high prevalence of recurrent infections among female patients (Brumbaugh and Mobley, 2012). Therefore, obtained genetic variations explain the variation in antibiotic resistance and phenotypic assays. As variation in genetic content can also determine the colonisation and pathogenicity levels resulting from having different UTIs (Vejborg et al., 2011), studying patients with a history of UTIs could add a further dimension when investigating UPEC diversity. Overall phenotypic and genotypic diversity is mostly seen in female patients. Our findings show that two-third of our highly diverse sample patterns came from female patient. Assuming that female samples are more susceptible to having a diverse bacterial population, having a larger sample size could prove this hypothesis.

Determined heteroresistance, phenotypic and genotypic diversity, within a single patient, by comparing identical colonies within UPEC infected patients, raises concerns regarding the effectiveness of antibiotic treatment and infection management. Therefore, several considerations arise and this affects the choice of treatment and mode of microbial analysis and diagnosis. Willner et al. (2014) stated that single bacterial isolates of uncomplicated UTIs were representative and effective for patient treatment. Here, within host diversity indicates that single

colony selection is not representative for heterogenous bacterial population. This research has focused on patients with UTIs generally, with no respect to a specific clinical situation. Further, single colony selection may underestimate species diversity, and this was recognised in 16 sequenced isolates of *E. coli* faecal samples. Polyclonal diversity was recognised within patients with presence of multiple types of ESBL. Within clone diversity was also recognised (Stoesser et al., 2015). As UPEC are thought to originate in human intestines (Pitout, 2012a) and extensive within host diversity was shown in faecal *E. coli* samples (Stoesser et al., 2015), this could provide a mechanism for the variation we have observed.

In conclusion, 19.5 % of collected UPEC samples display either highly dynamic within host polyclonal or monoclonal genetic diversity and this may be related with the long-term coexistence of UPEC within the patient's urinary tract. However, it is unclear whether patients with polyclonal and monoclonal diversity are infected with a multi-strain of UPEC at single time point, or because of reinfection with a new strain, or a relapse with the same strain during the current infection. Persistence and recurrent infection may be significantly related to UPEC diversity within a single host. For patients who need treatment, it may be efficient to mix a few colonies together to give the most effective antimicrobial susceptibility result, as was done with *P. aeruginosa* in Darch et al.'s (2015) published study. Finally, our data is the most high-resolution snapshot to date that proved to have phenotypic and genotypic diversity within a single patient and demonstrate homogenous bacterial community may reveal phenotypic and genotypic variations.

6.2 Future work

The presented research evaluated the diversity level within a single patient. Based on our findings, knowledge of patient history might be a crucial determinant of the diversity level within a single patient. Our findings also determined the presence of varied proteins between bacterial isolates that belongs to a single patient. We hypothesized that replication protein, conjugal transfer protein, IS66, IS110, binding protein and recombinase proteins are responsible for the obtained phenotypic and genotypic variation and might be responsible for within-host diversity of a single patient. Therefore, it would be worthwhile in future work to evaluate the effect of these proteins on within-host diversity. Such a finding could provide a better understanding and sufficient knowledge of single host diversity.

Chapter Seven

Appendix

Table 7.1 VCF files of deep mapping sequence samples (low and identical resistance profiles): this table displays position, altered amino acid, quality and mapping information: REF: the reference base, ALT: alternative base, Quality: quality score, DP: depth, AF1: allele frequency and DP4: number of forward reference allele; reverse reference allele; forward non reference allele and reverse non reference allele. Positions highlighted in red is the non-synonyms mutation, the yellowish positions are the synonyms mutation.

UTI (F9)													
Position		REF	ALT	Quality	Information								
2566275		C	T	125	DP	75	AF1	0.5	DP4	14	14	22	25
4857708	Hypothetical protein	T	C	160	DP	109	AF1	0.5	DP4	13	26	36	33
2566315		C	T	225	DP	77	AF1	0.5	DP4	12	12	24	28
47462		C	T	135	DP	45	AF1	1	DP4	0	1	18	20
182801		C	G	186	DP	57	AF1	1	DP4	0	1	19	34
261785		A	G	140	DP	35	AF1	1	DP4	0	4	23	8
416322		C	A	164	DP	32	AF1	1	DP4	4	0	9	18
426579		C	A	131	DP	40	AF1	1	DP4	0	4	25	8
433011		G	T	181	DP	49	AF1	1	DP4	1	0	13	23
586455		C	G	222	DP	71	AF1	1	DP4	0	0	29	39
653170		C	A	144	DP	81	AF1	1	DP4	0	4	42	8
724271		C	T	161	DP	62	AF1	1	DP4	0	1	34	9
1065711		G	T	162	DP	57	AF1	1	DP4	2	0	16	30
1191053		T	G	209	DP	44	AF1	1	DP4	1	0	16	27
1380217		C	A	196	DP	50	AF1	1	DP4	6	0	23	21
1586918		A	T	222	DP	70	AF1	1	DP4	0	1	33	36
1711518		T	G	222	DP	45	AF1	1	DP4	1	0	19	23
2033773		T	C	141	DP	37	AF1	1	DP4	3	0	10	22
2037751		G	C	139	DP	37	AF1	1	DP4	0	5	25	6
2058289		C	G	222	DP	66	AF1	1	DP4	0	0	34	25
2594640		C	A	120	DP	45	AF1	1	DP4	0	4	21	13
2698830		T	C	222	DP	90	AF1	1	DP4	4	7	27	42
2698914		C	T	224	DP	98	AF1	1	DP4	15	8	35	39
2698966		G	A	222	DP	92	AF1	1	DP4	10	8	29	41
2698977		C	T	225	DP	88	AF1	1	DP4	8	7	25	37

2699766		T	C	222	DP	72	AF1	1	DP4	6	4	24	33
2699883		C	T	222	DP	79	AF1	1	DP4	3	0	35	35
2825903		C	T	163	DP	68	AF1	1	DP4	1	3	35	23
3076738		C	A	172	DP	74	AF1	1	DP4	0	11	35	26
3294541		C	A	163	DP	58	AF1	1	DP4	0	1	33	24
3308848		C	A	134	DP	50	AF1	1	DP4	0	3	32	8
3308870		C	A	176	DP	55	AF1	1	DP4	0	1	37	6
3499664		C	A	220	DP	54	AF1	1	DP4	1	0	28	25
3513464		G	T	151	DP	55	AF1	1	DP4	0	1	31	21
3553547		G	T	134	DP	32	AF1	1	DP4	2	0	14	14
3573038		C	A	128	DP	66	AF1	1	DP4	0	0	34	8
3970886		G	T	125	DP	58	AF1	1	DP4	1	0	18	24
3982836		G	T	166	DP	42	AF1	1	DP4	1	0	7	21
3982881		G	T	148	DP	43	AF1	1	DP4	0	1	9	23
4051247		C	T	222	DP	37	AF1	1	DP4	1	0	13	19
4078420		C	A	123	DP	45	AF1	1	DP4	0	2	29	8
4078431		G	T	150	DP	43	AF1	1	DP4	0	5	25	11
4078459		C	T	169	DP	41	AF1	1	DP4	0	3	24	9
4281615		T	G	222	DP	72	AF1	1	DP4	0	0	31	39
4375154		A	C	146	DP	59	AF1	1	DP4	9	0	16	28
4375171		G	T	125	DP	58	AF1	1	DP4	7	0	11	30
4558824		C	A	125	DP	39	AF1	1	DP4	0	5	15	17
4565452		T	A	222	DP	62	AF1	1	DP4	0	0	33	22
4748127		C	G	154	DP	40	AF1	1	DP4	0	0	20	10
UTI (M19)													
4599529		C	T	225	DP	150	AF1	0.5	DP4	27	33	42	45
5368396	Integrase core domain protein	A	T	225	DP	148	AF1	0.5	DP4	27	30	46	36
5375822		T	C	225	DP	160	AF1	0.5	DP4	28	38	41	50
5368473	Integrase core domain protein	A	G	191	DP	150	AF1	0.5	DP4	23	33	44	40
5368396	Integrase core domain protein	G	T	179	DP	143	AF1	0.5	DP4	26	29	47	33
4599579		G	T	164	DP	140	AF1	0.5	DP4	24	33	43	37
4599233	Small toxic polypeptide LdrD	G	A	135	DP	154	AF1	0.5	DP4	29	17	54	51

2838182	.	A	G	123	DP	173	AF1	0.5	DP4	36	31	52	49
4599716	Small toxic polypeptide LdrD	G	A	122	DP	120	AF1	0.5	DP4	12	15	47	44
74769		A	C	133	DP	42	AF1	1	DP4	4	0	8	18
74879		G	T	170	DP	57	AF1	1	DP4	0	0	27	21
107510		C	A	222	DP	58	AF1	1	DP4	0	1	12	33
237470		C	T	124	DP	59	AF1	1	DP4	5	1	10	35
264901		A	T	222	DP	55	AF1	1	DP4	1	0	14	40
298201		C	G	179	DP	100	AF1	1	DP4	1	0	45	49
453441		G	T	179	DP	49	AF1	1	DP4	5	0	14	27
453477		G	T	138	DP	58	AF1	1	DP4	9	0	17	32
453496		G	A	133	DP	61	AF1	1	DP4	0	0	12	31
871038		G	A	122	DP	54	AF1	1	DP4	0	0	7	35
871049		G	A	153	DP	52	AF1	1	DP4	0	0	11	34
898945		G	A	127	DP	51	AF1	1	DP4	1	0	5	34
1460423		C	A	144	DP	44	AF1	1	DP4	0	0	24	10
1460448		C	T	131	DP	46	AF1	1	DP4	0	2	28	10
1460481		C	A	155	DP	52	AF1	1	DP4	0	3	27	15
1650551		G	C	198	DP	100	AF1	1	DP4	0	1	57	33
1946071		C	A	158	DP	66	AF1	1	DP4	0	0	38	17
2204233		T	G	124	DP	69	AF1	1	DP4	1	0	25	38
2570287		G	T	148	DP	59	AF1	1	DP4	3	0	8	39
2570330		T	A	131	DP	57	AF1	1	DP4	7	0	14	33
2570367		G	A	204	DP	54	AF1	1	DP4	0	0	13	31
2570383		G	A	134	DP	56	AF1	1	DP4	0	0	16	30
2717371		A	G	222	DP	63	AF1	1	DP4	1	1	27	32
2950155		G	T	163	DP	53	AF1	1	DP4	1	0	10	31
2953381		T	C	157	DP	88	AF1	1	DP4	2	7	32	41
2968290		T	A	222	DP	67	AF1	1	DP4	1	0	26	34
3052591		G	T	222	DP	93	AF1	1	DP4	1	0	56	36
3060191		A	C	183	DP	106	AF1	1	DP4	0	1	42	60
3191904		A	G	128	DP	44	AF1	1	DP4	0	4	19	19
3191938		C	A	172	DP	46	AF1	1	DP4	0	1	20	15

3196168		C	A	164	DP	60	AF1	1	DP4	0	1	32	19
3196196		C	A	171	DP	64	AF1	1	DP4	0	7	36	20
3347998		G	T	124	DP	41	AF1	1	DP4	2	0	8	14
3393087		G	T	145	DP	73	AF1	1	DP4	9	1	34	26
3393180		A	G	132	DP	77	AF1	1	DP4	7	1	33	32
3686171		T	A	222	DP	77	AF1	1	DP4	0	1	41	25
3960927		C	A	135	DP	86	AF1	1	DP4	0	8	44	18
4572570		A	T	222	DP	85	AF1	1	DP4	0	1	40	43
4872136		C	T	175	DP	53	AF1	1	DP4	5	0	21	24
4872200		G	T	120	DP	48	AF1	1	DP4	3	0	12	20
4941485		C	A	152	DP	67	AF1	1	DP4	0	0	37	7
5021394		C	G	222	DP	67	AF1	1	DP4	0	1	34	27
5114795		C	A	153	DP	47	AF1	1	DP4	0	2	25	12
5348608		G	T	131	DP	83	AF1	1	DP4	2	2	22	43
5423389		T	G	134	DP	24	AF1	1	DP4	1	1	9	11
UTI (F7)													
5299267	Antigen 43	C	A	122	DP	221	AF1	0.5	DP4	54	34	79	53
5296801		C	T	129	DP	284	AF1	0.5	DP4	48	37	111	72
5322388	Actin cross-linking toxin VgrG1	T	C	133	DP	308	AF1	0.5	DP4	50	31	133	79
3636345	Actin cross-linking toxin VgrG1	A	T	144	DP	340	AF1	0.5	DP4	53	70	89	127
5299319		A	T	160	DP	211	AF1	0.5	DP4	45	33	74	57
5296988	Hypothetical protein	A	G	164	DP	304	AF1	0.5	DP4	50	48	106	94
5296310		A	G	188	DP	269	AF1	0.5	DP4	47	59	69	93
53910	Hypothetical protein	C	T	225	DP	209	AF1	0.5	DP4	13	33	88	74
3918536	Putative autotransporter	G	A	225	DP	98	AF1	0.5	DP4	12	17	36	32
5281901	Hypothetical protein	T	C	225	DP	134	AF1	0.5	DP4	30	15	51	35
5282012	Hypothetical protein	A	G	225	DP	179	AF1	0.5	DP4	39	23	67	47
5295629	Hypothetical protein	A	G	225	DP	321	AF1	0.5	DP4	4	61	104	100
5296451		T	C	225	DP	83	AF1	0.5	DP4	9	15	22	37
5296867		C	T	225	DP	294	AF1	0.5	DP4	64	34	111	79
3620060		G	A	125	DP	87	AF1	1	DP4	2	0	18	42
5169068		G	T	126	DP	72	AF1	1	DP4	1	10	30	30

2947500		C	T	137	DP	104	AF1	1	DP4	0	5	49	19
4656086		C	A	141	DP	92	AF1	1	DP4	0	4	49	10
1160396		G	T	142	DP	110	AF1	1	DP4	0	4	54	37
3250416		C	A	145	DP	112	AF1	1	DP4	1	9	50	30
191240		C	A	148	DP	70	AF1	1	DP4	0	9	35	22
3475826		G	T	165	DP	139	AF1	1	DP4	7	0	54	67
190701		G	T	177	DP	166	AF1	1	DP4	1	0	72	73
2829970		T	C	188	DP	89	AF1	1	DP4	0	0	33	55
1407412		C	A	209	DP	108	AF1	1	DP4	0	4	64	36
470216		A	T	222	DP	120	AF1	1	DP4	1	0	52	62
641516		A	C	222	DP	132	AF1	1	DP4	0	0	63	65
1760388		C	A	222	DP	196	AF1	1	DP4	1	1	72	102
2260808		T	C	222	DP	147	AF1	1	DP4	0	1	79	54
2265808		G	T	222	DP	67	AF1	1	DP4	1	0	22	36
3473826		C	T	222	DP	58	AF1	1	DP4	0	0	23	33
3918941		T	A	222	DP	99	AF1	1	DP4	0	0	35	54
3918965		T	G	222	DP	100	AF1	1	DP4	1	0	42	49
UTI (M5)													
2226609		C	A	148	DP	58	AF1	0.5	DP4	1	8	22	12
4993266	Hypothetical protein	C	T	128	DP	329	AF1	0.5	DP4	66	58	88	101
4946892		G	A	139	DP	23	AF1	0.5	DP4	3	2	8	8
831267		T	G	157	DP	122	AF1	0.5	DP4	16	19	31	46
831255		C	T	161	DP	119	AF1	0.5	DP4	17	19	39	33
4994873	DsdX permease	G	A	167	DP	275	AF1	0.5	DP4	51	65	70	79
4999469		T	C	177	DP	233	AF1	0.5	DP4	33	66	49	81
4993368	Hypothetical protein	G	A	225	DP	301	AF1	0.5	DP4	54	70	76	91
4999142		C	T	225	DP	225	AF1	0.5	DP4	41	51	54	69
317001		G	A	222	DP	115	AF1	1	DP4	0	0	56	55
1202383		C	T	140	DP	67	AF1	1	DP4	0	6	47	10
1365636		A	G	222	DP	117	AF1	1	DP4	1	0	59	50
3383983		T	C	222	DP	90	AF1	1	DP4	0	0	59	17
3564534		C	T	186	DP	113	AF1	1	DP4	0	0	60	41

3700768		A	G	157	DP	64	AF1	1	DP4	2	0	15	40
3748176		C	A	207	DP	124	AF1	1	DP4	0	6	72	34
3769829		G	T	160	DP	97	AF1	1	DP4	4	0	19	47
3850354		G	T	222	DP	206	AF1	1	DP4	0	1	99	94
3879854		G	T	222	DP	69	AF1	1	DP4	0	0	34	27
3905954		A	T	222	DP	107	AF1	1	DP4	1	0	46	60
4761309		C	A	222	DP	99	AF1	1	DP4	1	0	40	56
UTI (F8)													
4799131	Manganese transport system membrane protein MntB	T	G	225	DP	86	AF1	0.5	DP4	15	15	32	24
4799410	Manganese transport system membrane protein MntB	T	C	225	DP	110	AF1	0.5	DP4	21	17	32	34
4808498	Transposase IS200 like protein	G	A	225	DP	45	AF1	0.5	DP4	2	11	9	23
4816913		C	T	225	DP	55	AF1	0.5	DP4	8	11	14	21
4817783		T	C	225	DP	98	AF1	0.5	DP4	18	20	28	32
4811088	tRNA nuclease WapA precursor	A	G	223	DP	79	AF1	0.5	DP4	23	6	47	3
4816574		C	T	177	DP	49	AF1	0.5	DP4	10	4	20	14
4816671	Phage DNA packaging protein Nu1	G	A	170	DP	49	AF1	0.5	DP4	12	4	19	14
4799230	Manganese transport system membrane protein MntB	T	C	166	DP	92	AF1	0.5	DP4	16	18	28	29
4799254	Manganese transport system membrane protein MntB	T	C	163	DP	86	AF1	0.5	DP4	17	15	25	28
4799857	Manganese transport system membrane protein MntB	T	C	161	DP	67	AF1	0.5	DP4	7	11	17	25
4808461		T	C	158	DP	43	AF1	0.5	DP4	3	10	10	20
4816445		A	G	152	DP	32	AF1	0.5	DP4	9	1	16	6
4816868		T	C	126	DP	57	AF1	0.5	DP4	9	11	18	19
4799113	Manganese transport system membrane protein MntB	T	C	125	DP	88	AF1	0.5	DP4	16	15	33	23
3252096		G	T	122	DP	33	AF1	0.51	DP4	10	0	11	11
17101		T	G	148	DP	44	AF1	1	DP4	0	1	20	23
48301		C	T	162	DP	38	AF1	1	DP4	2	0	8	26
115624		C	A	152	DP	31	AF1	1	DP4	0	1	18	7
127401		G	T	154	DP	47	AF1	1	DP4	2	0	26	19
170501		G	C	202	DP	39	AF1	1	DP4	0	0	16	18
181424		G	A	146	DP	38	AF1	1	DP4	0	3	21	9
233982		G	T	129	DP	36	AF1	1	DP4	8	0	14	14
334337		G	T	122	DP	28	AF1	1	DP4	4	0	10	14

438442		C	A	151	DP	35	AF1	1	DP4	0	4	21	9
438452		C	A	126	DP	35	AF1	1	DP4	0	4	21	10
443017		C	T	205	DP	55	AF1	1	DP4	1	1	26	25
445636		G	A	120	DP	31	AF1	1	DP4	4	0	8	13
445769		G	T	175	DP	43	AF1	1	DP4	0	0	9	25
536405		C	A	159	DP	34	AF1	1	DP4	0	2	15	13
686073		G	T	121	DP	50	AF1	1	DP4	1	0	11	27
719542		A	G	135	DP	38	AF1	1	DP4	0	2	21	9
758679		G	T	148	DP	46	AF1	1	DP4	9	0	17	19
779117		T	C	222	DP	38	AF1	1	DP4	1	0	24	13
788861		C	T	138	DP	26	AF1	1	DP4	0	0	3	13
788897		G	T	122	DP	27	AF1	1	DP4	2	0	6	15
810821		G	T	126	DP	45	AF1	1	DP4	1	0	5	24
843462		A	G	170	DP	39	AF1	1	DP4	0	1	20	17
917032		G	C	158	DP	23	AF1	1	DP4	0	2	13	7
985462		A	T	202	DP	35	AF1	1	DP4	0	1	16	17
1059362		T	G	194	DP	28	AF1	1	DP4	1	0	16	9
1075756		T	A	222	DP	37	AF1	1	DP4	1	0	15	10
1392089		G	T	126	DP	30	AF1	1	DP4	1	1	4	22
1434534		G	T	222	DP	52	AF1	1	DP4	0	1	19	30
1555856		C	A	167	DP	42	AF1	1	DP4	3	0	14	19
1562441		G	A	222	DP	22	AF1	1	DP4	0	1	15	6
1563041		G	A	200	DP	39	AF1	1	DP4	0	1	19	16
1629004		C	A	167	DP	41	AF1	1	DP4	0	1	26	5
1638668		G	T	133	DP	58	AF1	1	DP4	4	0	10	35
1655719		G	T	193	DP	50	AF1	1	DP4	0	0	20	13
1841297		C	A	157	DP	35	AF1	1	DP4	0	0	16	9
1852536		A	T	125	DP	27	AF1	1	DP4	1	0	3	21
2001467		T	A	222	DP	36	AF1	1	DP4	0	1	20	8
2043696		T	G	172	DP	43	AF1	1	DP4	1	0	15	26
2062053		G	T	169	DP	32	AF1	1	DP4	1	0	6	19
2181725		G	T	221	DP	38	AF1	1	DP4	3	0	14	19

2293668		C	A	152	DP	33	AF1	1	DP4	0	3	15	12
2400369		T	G	150	DP	39	AF1	1	DP4	4	0	16	19
2400380		G	T	143	DP	36	AF1	1	DP4	3	0	11	18
2400407		C	G	131	DP	33	AF1	1	DP4	3	0	6	16
2400611		C	A	178	DP	36	AF1	1	DP4	1	0	12	16
2401517		C	T	120	DP	27	AF1	1	DP4	0	0	17	6
2414184		G	T	130	DP	25	AF1	1	DP4	0	3	18	4
2501117		G	T	222	DP	49	AF1	1	DP4	0	1	27	19
2560715		G	T	132	DP	27	AF1	1	DP4	4	0	11	11
2560727		G	T	140	DP	27	AF1	1	DP4	3	0	14	10
2568691		C	A	163	DP	29	AF1	1	DP4	0	0	12	10
2617724		C	A	120	DP	25	AF1	1	DP4	0	3	13	5
2638669		C	A	138	DP	33	AF1	1	DP4	5	0	10	17
2638688		G	T	146	DP	40	AF1	1	DP4	1	0	7	18
2638715		G	T	197	DP	46	AF1	1	DP4	2	0	17	19
2706919		G	T	135	DP	27	AF1	1	DP4	0	0	5	13
2886762		G	C	171	DP	30	AF1	1	DP4	0	0	9	15
2986960		T	G	134	DP	36	AF1	1	DP4	5	0	17	14
3033473		C	G	176	DP	47	AF1	1	DP4	1	0	16	26
3037154		G	T	144	DP	36	AF1	1	DP4	1	0	7	18
3059511		C	A	154	DP	48	AF1	1	DP4	0	0	25	12
3059620		G	T	173	DP	37	AF1	1	DP4	0	4	16	15
3059646		C	A	166	DP	38	AF1	1	DP4	0	1	16	14
3317688		T	A	181	DP	45	AF1	1	DP4	1	0	18	26
3349929		G	T	178	DP	48	AF1	1	DP4	0	2	22	20
3361945		T	G	165	DP	35	AF1	1	DP4	6	0	7	21
3433943		G	T	177	DP	52	AF1	1	DP4	5	0	18	25
3575500		G	T	157	DP	44	AF1	1	DP4	2	0	12	27
3672504		C	A	141	DP	43	AF1	1	DP4	1	0	24	16
3681405		C	A	125	DP	41	AF1	1	DP4	5	0	6	29
3735104		C	A	211	DP	52	AF1	1	DP4	1	0	27	23
3817554		G	T	122	DP	44	AF1	1	DP4	3	0	2	25

3866743		A	T	127	DP	38	AF1	1	DP4	6	0	14	12
3967433		T	C	187	DP	51	AF1	1	DP4	0	0	25	17
4016759		C	A	129	DP	31	AF1	1	DP4	0	0	21	5
4253602		C	A	222	DP	41	AF1	1	DP4	0	0	20	17
4254702		G	T	148	DP	19	AF1	1	DP4	0	1	4	14
4254786		C	T	212	DP	22	AF1	1	DP4	1	2	5	13
4300631		C	T	127	DP	34	AF1	1	DP4	1	0	12	20
4300841		A	T	138	DP	32	AF1	1	DP4	2	0	8	19
4323326		G	T	164	DP	27	AF1	1	DP4	2	0	12	11
4468927		C	A	123	DP	39	AF1	1	DP4	0	4	12	19
4468940		C	A	191	DP	40	AF1	1	DP4	0	0	7	20
4643744		G	T	120	DP	41	AF1	1	DP4	0	4	22	6
4666734		C	A	148	DP	25	AF1	1	DP4	2	0	7	16
UTI (F11)													
1767247	Fructose-specific phosphotransferase enzyme IIb component	G	A	141	DP	58	AF1	0.5	DP4	3	11	20	24
4553709		C	T	159	DP	91	AF1	0.5	DP4	19	16	13	41
4712621		A	G	201	DP	488	AF1	0.5	DP4	91	104	128	149
4694980		C	T	225	DP	86	AF1	0.5	DP4	13	16	28	29
143754		C	T	173	DP	49	AF1	1	DP4	0	5	28	16
143764		G	T	222	DP	50	AF1	1	DP4	0	2	28	19
143774		C	A	170	DP	52	AF1	1	DP4	0	0	30	9
143790		C	T	134	DP	52	AF1	1	DP4	0	5	30	10
288520		G	T	147	DP	36	AF1	1	DP4	1	0	4	25
288538		G	A	131	DP	36	AF1	1	DP4	2	0	5	24
345851		C	T	144	DP	40	AF1	1	DP4	5	0	4	30
397355		C	A	124	DP	48	AF1	1	DP4	0	2	29	11
658269		G	A	222	DP	50	AF1	1	DP4	0	1	23	25
666962		C	A	173	DP	42	AF1	1	DP4	0	2	23	8
963101		C	A	178	DP	62	AF1	1	DP4	7	0	22	33
1088454		C	A	191	DP	52	AF1	1	DP4	0	1	28	18
1088487		C	A	189	DP	53	AF1	1	DP4	0	1	28	13
1207612		G	T	150	DP	60	AF1	1	DP4	7	0	21	30

1371373		C	T	175	DP	56	AF1	1	DP4	0	2	21	21
1464847		C	A	139	DP	60	AF1	1	DP4	0	2	32	10
1530662		G	T	155	DP	50	AF1	1	DP4	8	0	12	30
1707315		G	T	131	DP	70	AF1	1	DP4	1	0	32	27
1813711		G	T	178	DP	41	AF1	1	DP4	1	0	13	19
1813723		G	T	133	DP	42	AF1	1	DP4	7	0	13	19
1893382		A	T	222	DP	24	AF1	1	DP4	1	0	10	12
2004351		G	T	178	DP	58	AF1	1	DP4	0	0	30	26
2030951		G	C	167	DP	56	AF1	1	DP4	0	1	26	27
2210852		G	A	181	DP	62	AF1	1	DP4	0	1	39	20
2499913		T	C	222	DP	42	AF1	1	DP4	1	0	21	20
2500902		G	T	149	DP	35	AF1	1	DP4	4	0	9	21
2543199		C	T	149	DP	54	AF1	1	DP4	0	4	26	13
2625178		G	T	128	DP	51	AF1	1	DP4	1	0	7	33
2680004		G	T	204	DP	46	AF1	1	DP4	0	3	26	9
2905781		G	T	149	DP	82	AF1	1	DP4	0	0	35	45
3025129		C	A	161	DP	30	AF1	1	DP4	0	0	14	8
3159881		C	A	128	DP	66	AF1	1	DP4	0	5	30	24
3298166		G	T	125	DP	39	AF1	1	DP4	8	0	9	22
3510537		C	G	180	DP	60	AF1	1	DP4	0	1	31	25
3515838		G	T	149	DP	48	AF1	1	DP4	6	0	22	18
3603027		G	A	122	DP	68	AF1	1	DP4	0	10	29	26
3614238		G	T	163	DP	64	AF1	1	DP4	3	0	14	34
3678716		G	T	160	DP	41	AF1	1	DP4	3	1	16	14
3713326		T	A	222	DP	43	AF1	1	DP4	1	0	21	18
3713515		C	A	151	DP	41	AF1	1	DP4	0	3	7	17
3713684		G	T	151	DP	38	AF1	1	DP4	3	0	14	19
3832019		G	C	222	DP	54	AF1	1	DP4	0	1	27	18
4293566		C	A	147	DP	54	AF1	1	DP4	0	7	20	26
4311350		G	T	166	DP	41	AF1	1	DP4	1	0	4	26
4353935		G	T	156	DP	60	AF1	1	DP4	6	0	18	33
4369326		G	T	166	DP	58	AF1	1	DP4	9	0	20	28

4369341		T	C	128	DP	58	AF1	1	DP4	7	0	23	25
4369388		G	T	184	DP	56	AF1	1	DP4	9	0	16	30
4369417		G	T	172	DP	66	AF1	1	DP4	8	0	21	30
4369428		C	T	179	DP	65	AF1	1	DP4	7	0	23	33
4391683		T	C	146	DP	54	AF1	1	DP4	0	8	29	16
4401561		G	T	152	DP	50	AF1	1	DP4	2	1	14	26
4401629		G	T	142	DP	44	AF1	1	DP4	5	0	15	22
4401656		G	T	143	DP	41	AF1	1	DP4	1	0	14	21
4497607		C	A	135	DP	65	AF1	1	DP4	0	4	33	18
4499992		T	G	222	DP	81	AF1	1	DP4	1	0	35	45
4695535		C	A	141	DP	28	AF1	1	DP4	0	0	18	7
UTI (M4)													
375768	Electron transport complex protein RnfC	G	A	176	DP	82	AF1	0.5	DP4	11	15	27	22
375756	Electron transport complex protein RnfC	A	C	156	DP	82	AF1	0.5	DP4	11	13	35	22
375705	Electron transport complex protein RnfC	C	T	121	DP	82	AF1	0.5	DP4	5	10	37	28
3908592	Primosomal protein N'	C	A	120	DP	50	AF1	0.5	DP4	0	5	7	15
192701		C	A	207	DP	52	AF1	1	DP4	0	1	25	24
277501		T	G	156	DP	49	AF1	1	DP4	0	1	25	20
375301		C	A	138	DP	37	AF1	1	DP4	2	0	21	13
568634		T	G	150	DP	27	AF1	1	DP4	0	0	12	6
741298		C	A	155	DP	64	AF1	1	DP4	0	7	40	15
891356		A	T	150	DP	54	AF1	1	DP4	2	0	17	26
891369		G	T	122	DP	56	AF1	1	DP4	9	0	13	31
891451		G	A	216	DP	59	AF1	1	DP4	0	0	20	32
891716		C	A	190	DP	54	AF1	1	DP4	0	6	30	17
914150		C	T	179	DP	43	AF1	1	DP4	0	1	21	19
954350		T	A	222	DP	53	AF1	1	DP4	0	1	25	27
958752		G	T	133	DP	52	AF1	1	DP4	3	0	7	28
958823		G	T	168	DP	51	AF1	1	DP4	0	0	18	26
1016464		G	T	145	DP	50	AF1	1	DP4	1	0	18	23
1095150		A	T	222	DP	58	AF1	1	DP4	0	0	22	33
1112708		G	A	159	DP	49	AF1	1	DP4	5	0	11	31

1112738		G	T	151	DP	52	AF1	1	DP4	2	1	13	28
1124279		G	T	124	DP	42	AF1	1	DP4	2	0	6	20
1220308		C	A	148	DP	66	AF1	1	DP4	0	5	27	28
1224066		C	T	122	DP	41	AF1	1	DP4	0	0	16	12
1224084		C	T	138	DP	39	AF1	1	DP4	0	4	19	12
1247881		C	A	143	DP	64	AF1	1	DP4	1	7	42	13
1321050		T	G	222	DP	48	AF1	1	DP4	0	0	23	22
1597326		G	C	222	DP	51	AF1	1	DP4	0	0	31	15
1597424		C	A	192	DP	54	AF1	1	DP4	0	1	17	16
1614626		T	A	170	DP	56	AF1	1	DP4	1	0	28	25
1769811		C	A	154	DP	62	AF1	1	DP4	0	4	28	12
1770473		G	T	156	DP	60	AF1	1	DP4	0	8	33	17
1957026		G	A	222	DP	52	AF1	1	DP4	1	0	22	29
2441175		G	T	137	DP	54	AF1	1	DP4	1	0	11	25
2542630		G	A	156	DP	42	AF1	1	DP4	1	0	10	21
2542739		A	T	191	DP	43	AF1	1	DP4	5	0	18	16
2542930		A	T	123	DP	61	AF1	1	DP4	7	0	20	25
2836690		G	A	222	DP	46	AF1	1	DP4	0	2	31	13
2841290		A	G	222	DP	54	AF1	1	DP4	0	1	25	25
3031643		G	T	180	DP	55	AF1	1	DP4	2	0	19	22
3076416		G	T	130	DP	34	AF1	1	DP4	6	0	12	11
3210570		C	A	166	DP	36	AF1	1	DP4	1	2	13	17
3210682		C	A	133	DP	36	AF1	1	DP4	0	4	18	12
3283634		G	A	160	DP	72	AF1	1	DP4	0	5	37	20
3292642		T	C	142	DP	35	AF1	1	DP4	5	0	17	12
3417542		T	C	155	DP	39	AF1	1	DP4	0	7	22	8
3549316		G	T	131	DP	74	AF1	1	DP4	1	0	5	48
3785750		C	A	171	DP	74	AF1	1	DP4	0	1	31	22
3799235		G	T	174	DP	34	AF1	1	DP4	0	1	12	16
3922136		C	T	213	DP	68	AF1	1	DP4	0	1	39	28
3996560		C	T	121	DP	45	AF1	1	DP4	0	0	34	4
4027213		C	A	128	DP	63	AF1	1	DP4	0	4	31	11

4447096		C	T	141	DP	49	AF1	1	DP4	6	0	8	33
UTI (M7)													
5253754	Hypothetical protein	A	G	123	DP	130	AF1	0.5	DP4	9	22	39	58
5280458		T	C	142	DP	92	AF1	0.5	DP4	23	13	36	20
5192453	Minor fimbrial protein PrsF precursor	A	G	151	DP	74	AF1	0.5	DP4	6	21	28	19
5254581		T	C	151	DP	96	AF1	0.5	DP4	2	28	6	60
5237347	Hypothetical protein	A	G	158	DP	69	AF1	0.5	DP4	15	7	32	15
5257122	Filamentous hemagglutinin	T	G	162	DP	100	AF1	0.5	DP4	18	22	27	33
5254064		A	C	164	DP	154	AF1	0.5	DP4	29	29	44	51
5253461	Hypothetical protein	T	C	225	DP	115	AF1	0.5	DP4	19	10	55	29
5254539		C	T	225	DP	124	AF1	0.5	DP4	9	33	11	71
5257161	Filamentous hemagglutinin	G	A	225	DP	98	AF1	0.5	DP4	17	20	23	34
5280383		A	G	225	DP	82	AF1	0.5	DP4	18	13	27	22
116501		C	T	194	DP	50	AF1	1	DP4	2	1	21	22
116534		G	T	144	DP	47	AF1	1	DP4	2	1	16	24
127171		C	A	126	DP	50	AF1	1	DP4	2	5	25	14
130967		G	T	120	DP	21	AF1	1	DP4	0	3	9	9
150284		A	T	222	DP	26	AF1	1	DP4	2	1	17	5
252900		T	C	222	DP	66	AF1	1	DP4	1	0	23	41
432367		G	T	167	DP	34	AF1	1	DP4	3	0	6	19
432400		A	C	222	DP	35	AF1	1	DP4	1	0	14	17
586870		C	T	168	DP	54	AF1	1	DP4	0	0	10	37
613850		G	T	143	DP	35	AF1	1	DP4	4	0	24	7
700187		G	T	122	DP	30	AF1	1	DP4	3	1	6	20
901360		T	C	134	DP	51	AF1	1	DP4	8	0	11	27
901413		T	C	135	DP	54	AF1	1	DP4	1	0	12	33
968779		G	T	128	DP	42	AF1	1	DP4	3	1	9	23
978379		T	C	190	DP	43	AF1	1	DP4	1	0	17	24
1208749		G	T	148	DP	48	AF1	1	DP4	1	0	15	24
1249259		C	A	132	DP	24	AF1	1	DP4	3	0	4	17
1320685		A	T	220	DP	48	AF1	1	DP4	0	1	24	21
1356285		C	A	222	DP	32	AF1	1	DP4	1	0	6	24

1363907		G	T	199	DP	48	AF1	1	DP4	1	0	17	17
1363920		G	T	150	DP	47	AF1	1	DP4	2	0	27	16
1594777		C	T	222	DP	50	AF1	1	DP4	1	0	12	36
1852082		C	G	222	DP	64	AF1	1	DP4	1	0	37	23
1852449		G	C	145	DP	39	AF1	1	DP4	2	2	6	28
2021443		C	A	153	DP	60	AF1	1	DP4	0	5	29	11
2057006		G	T	152	DP	46	AF1	1	DP4	1	3	24	16
2063318		C	A	144	DP	41	AF1	1	DP4	0	1	23	7
2063346		C	A	138	DP	39	AF1	1	DP4	2	0	17	8
2063970		G	A	222	DP	39	AF1	1	DP4	0	2	16	21
2091236		G	T	207	DP	53	AF1	1	DP4	0	1	29	17
2107170		C	A	200	DP	34	AF1	1	DP4	1	0	21	12
2160538		G	C	141	DP	56	AF1	1	DP4	1	0	7	32
2182788		C	A	149	DP	44	AF1	1	DP4	0	5	24	12
2304843		C	T	132	DP	24	AF1	1	DP4	2	0	4	16
2305178		C	A	129	DP	48	AF1	1	DP4	4	0	28	11
2305213		T	G	178	DP	45	AF1	1	DP4	1	0	27	11
2633384		G	T	130	DP	29	AF1	1	DP4	3	0	10	12
2918256		T	C	126	DP	44	AF1	1	DP4	0	8	25	10
3165659		A	C	142	DP	18	AF1	1	DP4	0	1	12	5
3184913		C	A	203	DP	70	AF1	1	DP4	0	2	32	18
3276659		G	A	130	DP	23	AF1	1	DP4	0	1	14	8
3377416		T	C	198	DP	36	AF1	1	DP4	1	1	15	18
3599555		G	T	121	DP	35	AF1	1	DP4	6	0	9	14
3655562		A	G	222	DP	37	AF1	1	DP4	1	0	23	13
3898922		G	T	175	DP	45	AF1	1	DP4	1	0	9	25
4067967		A	C	130	DP	38	AF1	1	DP4	4	0	10	23
4190262		C	A	172	DP	43	AF1	1	DP4	0	2	19	16
4219037		C	T	222	DP	46	AF1	1	DP4	1	0	8	23
4283494		G	T	166	DP	53	AF1	1	DP4	0	0	24	21
4458805		A	C	222	DP	38	AF1	1	DP4	0	5	21	11
4458845		A	G	122	DP	41	AF1	1	DP4	0	6	23	11

4458877		C	G	140	DP	49	AF1	1	DP4	0	2	29	7
4469504		A	T	137	DP	36	AF1	1	DP4	6	0	10	19
4956296		C	A	222	DP	36	AF1	1	DP4	1	1	15	19
5179672		G	T	124	DP	21	AF1	1	DP4	1	0	16	2
5209653		C	A	150	DP	67	AF1	1	DP4	0	1	36	13
5233310		C	T	184	DP	45	AF1	1	DP4	2	0	18	21
UTI (M10)													
5013035		T	G	225	DP	331	AF1	0.5	DP4	32	56	87	140
5013003		T	C	202	DP	343	AF1	0.5	DP4	64	76	80	118
5012980		A	G	177	DP	351	AF1	0.5	DP4	70	77	84	119
5012938		C	T	173	DP	380	AF1	0.5	DP4	70	79	92	121
4259528	Integrase core domain protein	G	A	163	DP	147	AF1	0.5	DP4	19	36	44	46
5094004		G	A	162	DP	51	AF1	0.5	DP4	5	12	13	21
5042126		A	G	161	DP	50	AF1	0.5	DP4	5	7	21	16
2644044		G	A	155	DP	48	AF1	0.5	DP4	4	7	21	15
3649523	D-tagatose 3-epimerase	C	A	120	DP	35	AF1	0.572	DP4	0	8	16	9
83896		G	T	137	DP	25	AF1	1	DP4	1	0	6	14
91151		C	A	132	DP	44	AF1	1	DP4	0	5	24	13
104101		T	G	129	DP	41	AF1	1	DP4	1	0	18	22
152501		A	G	190	DP	35	AF1	1	DP4	1	0	19	13
251405		G	T	126	DP	41	AF1	1	DP4	1	0	8	26
311701		C	T	181	DP	36	AF1	1	DP4	2	0	17	15
349801		A	C	128	DP	39	AF1	1	DP4	1	0	16	22
392180		T	G	155	DP	32	AF1	1	DP4	4	0	7	20
430501		C	A	222	DP	31	AF1	1	DP4	0	0	17	9
430689		C	A	135	DP	33	AF1	1	DP4	0	2	18	8
441601		C	A	159	DP	41	AF1	1	DP4	0	2	18	19
484860		G	T	156	DP	22	AF1	1	DP4	0	0	15	3
510364		G	T	162	DP	50	AF1	1	DP4	1	0	9	31
707600		G	T	131	DP	21	AF1	1	DP4	3	0	9	9
757215		G	T	150	DP	43	AF1	1	DP4	2	0	11	23
775730		G	T	181	DP	21	AF1	1	DP4	1	0	7	13

807009		G	T	181	DP	38	AF1	1	DP4	0	0	20	13
845630		G	T	133	DP	39	AF1	1	DP4	4	0	9	24
975784		C	A	128	DP	49	AF1	1	DP4	0	4	18	9
981425		C	G	165	DP	46	AF1	1	DP4	0	0	17	23
985225		G	A	193	DP	42	AF1	1	DP4	0	0	23	18
1008629		G	T	124	DP	27	AF1	1	DP4	2	0	8	15
1057825		C	T	158	DP	34	AF1	1	DP4	2	0	14	15
1057840		G	T	176	DP	39	AF1	1	DP4	0	0	11	16
1104025		A	T	222	DP	51	AF1	1	DP4	1	0	14	35
1107603		C	A	142	DP	24	AF1	1	DP4	4	0	9	11
1116745		C	A	131	DP	32	AF1	1	DP4	0	2	18	6
1125825		G	A	150	DP	31	AF1	1	DP4	1	0	10	17
1203874		C	A	174	DP	52	AF1	1	DP4	1	4	26	17
1228521		G	T	145	DP	37	AF1	1	DP4	1	0	4	24
1236747		C	A	162	DP	26	AF1	1	DP4	2	0	10	10
1236775		G	T	179	DP	23	AF1	1	DP4	0	1	9	11
1236793		C	A	159	DP	25	AF1	1	DP4	1	0	9	12
1237325		A	G	204	DP	30	AF1	1	DP4	0	1	14	14
1299733		T	C	127	DP	48	AF1	1	DP4	5	0	14	26
1315670		G	A	222	DP	37	AF1	1	DP4	0	1	21	15
1378071		A	C	133	DP	39	AF1	1	DP4	8	0	12	17
1406802		G	T	161	DP	47	AF1	1	DP4	2	0	7	27
1581870		G	A	222	DP	24	AF1	1	DP4	0	1	13	7
1641892		G	A	131	DP	42	AF1	1	DP4	2	0	5	28
1645484		G	T	163	DP	28	AF1	1	DP4	3	0	6	14
1645506		G	T	130	DP	31	AF1	1	DP4	3	0	6	18
1645610		G	T	155	DP	27	AF1	1	DP4	0	0	8	15
1657971		C	A	182	DP	31	AF1	1	DP4	1	0	6	15
1662896		T	C	132	DP	33	AF1	1	DP4	4	0	9	20
1662917		G	A	163	DP	41	AF1	1	DP4	1	0	9	19
1663277		G	A	207	DP	38	AF1	1	DP4	0	0	19	10
1697593		G	T	142	DP	25	AF1	1	DP4	0	0	7	12

1813340		C	A	134	DP	26	AF1	1	DP4	2	0	2	18
1876742		T	G	124	DP	38	AF1	1	DP4	5	0	6	24
1894681		A	G	188	DP	27	AF1	1	DP4	0	1	8	16
1943681		C	A	163	DP	32	AF1	1	DP4	1	0	16	11
2158793		G	T	171	DP	39	AF1	1	DP4	0	0	12	23
2260293		G	C	211	DP	37	AF1	1	DP4	0	0	23	12
2281893		A	G	213	DP	19	AF1	1	DP4	0	1	11	5
2383260		A	G	126	DP	35	AF1	1	DP4	0	4	14	15
2383328		G	T	134	DP	37	AF1	1	DP4	0	7	16	12
2447067		A	T	222	DP	41	AF1	1	DP4	0	0	20	17
2483689		G	T	160	DP	38	AF1	1	DP4	3	0	19	13
2632004		C	A	123	DP	32	AF1	1	DP4	0	0	21	5
2636442		G	T	134	DP	31	AF1	1	DP4	3	0	14	12
2636511		G	T	149	DP	29	AF1	1	DP4	2	1	13	11
2638916		A	G	191	DP	33	AF1	1	DP4	0	3	13	14
2692572		G	T	139	DP	35	AF1	1	DP4	6	0	4	23
2762683		G	A	201	DP	47	AF1	1	DP4	0	1	24	16
2908313		C	A	153	DP	44	AF1	1	DP4	0	1	22	9
2914035		G	T	120	DP	28	AF1	1	DP4	2	0	3	17
2988528		A	C	222	DP	28	AF1	1	DP4	0	1	16	11
3082141		C	A	146	DP	30	AF1	1	DP4	0	0	17	8
3178238		T	A	135	DP	41	AF1	1	DP4	5	0	19	17
3220534		C	A	158	DP	44	AF1	1	DP4	0	0	20	23
3225840		C	A	205	DP	44	AF1	1	DP4	0	0	26	8
3225870		G	A	222	DP	44	AF1	1	DP4	0	0	21	9
3225902		T	A	166	DP	44	AF1	1	DP4	0	4	16	20
3235134		C	A	216	DP	33	AF1	1	DP4	0	0	14	18
3309970		T	A	141	DP	57	AF1	1	DP4	8	0	16	27
3309986		G	A	160	DP	58	AF1	1	DP4	4	0	19	27
3333690		C	A	157	DP	44	AF1	1	DP4	0	5	25	11
3438281		G	T	161	DP	27	AF1	1	DP4	1	0	7	18
3438876		C	A	137	DP	29	AF1	1	DP4	0	2	13	7

3544802		T	A	222	DP	31	AF1	1	DP4	1	0	8	21
3544840		A	T	222	DP	22	AF1	1	DP4	1	0	6	13
3600789		T	A	181	DP	35	AF1	1	DP4	4	0	11	19
3689685		C	A	138	DP	24	AF1	1	DP4	0	1	8	11
3696089		G	T	143	DP	29	AF1	1	DP4	4	0	9	14
3719862		C	A	124	DP	44	AF1	1	DP4	0	5	19	18
3734397		C	A	128	DP	31	AF1	1	DP4	0	5	13	10
3759027		G	T	161	DP	29	AF1	1	DP4	0	3	17	9
3771505		T	G	145	DP	39	AF1	1	DP4	0	3	21	13
3787692		G	A	222	DP	29	AF1	1	DP4	0	2	9	17
3859255		G	A	123	DP	27	AF1	1	DP4	0	0	11	11
3959932		C	A	180	DP	43	AF1	1	DP4	1	0	21	19
3980094		G	C	156	DP	39	AF1	1	DP4	1	2	12	16
3980571		C	A	190	DP	47	AF1	1	DP4	0	1	19	11
4078451		A	C	160	DP	69	AF1	1	DP4	0	6	32	27
4120387		C	A	122	DP	28	AF1	1	DP4	0	0	14	9
4154019		C	T	125	DP	40	AF1	1	DP4	0	0	28	4
4154046		C	A	132	DP	42	AF1	1	DP4	0	2	31	5
4180646		C	A	133	DP	38	AF1	1	DP4	0	2	17	13
4180678		G	A	222	DP	41	AF1	1	DP4	0	1	23	15
4248693		A	C	130	DP	22	AF1	1	DP4	1	1	15	3
4288349		G	T	144	DP	35	AF1	1	DP4	2	0	5	24
4322812		C	T	222	DP	45	AF1	1	DP4	2	0	19	19
4323011		C	T	154	DP	46	AF1	1	DP4	0	5	27	12
4444836		G	T	125	DP	60	AF1	1	DP4	3	0	9	32
4511096		G	A	141	DP	28	AF1	1	DP4	2	0	2	18
4539485		A	G	128	DP	48	AF1	1	DP4	0	6	26	16
4696873		C	A	139	DP	39	AF1	1	DP4	0	3	17	13
4697862		C	A	133	DP	36	AF1	1	DP4	0	6	19	11
4718680		T	A	222	DP	50	AF1	1	DP4	0	0	18	28
4739328		C	A	187	DP	43	AF1	1	DP4	0	1	16	15
4745971		T	C	126	DP	19	AF1	1	DP4	3	0	5	11

4956447		T	A	157	DP	48	AF1	1	DP4	4	0	13	30
4986109		G	C	193	DP	40	AF1	1	DP4	0	2	17	16
5042463		G	A	222	DP	44	AF1	1	DP4	0	2	19	21
5075111		C	A	127	DP	49	AF1	1	DP4	0	2	7	29
5084139		G	T	147	DP	16	AF1	1	DP4	1	0	4	11
UTI (F16)													
5319547		G	C	123	DP	99	AF1	0.5	DP4	15	15	35	32
2000311	Citrate lyase subunit beta	C	A	125	DP	34	AF1	0.6243	DP4	6	0	4	23
1327153		T	C	139	DP	27	AF1	0.501	DP4	6	0	5	13
342709	Putative type II secretion system protein K	G	T	143	DP	38	AF1	0.508	DP4	8	0	14	11
5307367		A	G	145	DP	90	AF1	0.5	DP4	9	16	28	37
5337913	Hypothetical protein	T	C	152	DP	40	AF1	0.5	DP4	4	6	14	16
5322246		T	G	153	DP	123	AF1	0.5	DP4	14	13	32	53
5306775		T	C	166	DP	109	AF1	0.5	DP4	22	13	36	35
5306797		A	G	166	DP	111	AF1	0.5	DP4	20	21	39	30
5294582	Hypothetical protein	A	G	168	DP	113	AF1	0.5	DP4	20	18	33	39
5306871		G	A	174	DP	109	AF1	0.5	DP4	15	24	35	31
5294106	Hypothetical protein	G	A	225	DP	112	AF1	0.5	DP4	16	22	40	31
5294194		T	C	225	DP	113	AF1	0.5	DP4	19	16	47	29
5313754	Glutamate decarboxylase beta	C	T	225	DP	63	AF1	0.5	DP4	14	5	31	9
5322064		T	A	225	DP	119	AF1	0.5	DP4	16	8	67	22
3336195	Phage-related baseplate assembly protein	A	G	225	DP	47	AF1	0.5	DP4	7	6	18	14
3749157	Putative autotransporter precursor	A	G	225	DP	42	AF1	0.5	DP4	4	2	16	16
59567		A	T	222	DP	39	AF1	1	DP4	1	0	11	19
60533		C	T	222	DP	30	AF1	1	DP4	1	0	13	12
60601		T	A	198	DP	28	AF1	1	DP4	1	0	13	13
73401		C	A	219	DP	49	AF1	1	DP4	2	0	26	20
139401		G	T	216	DP	31	AF1	1	DP4	0	0	12	16
189301		C	T	201	DP	31	AF1	1	DP4	0	0	12	18
462220		C	A	143	DP	33	AF1	1	DP4	0	2	19	9
528305		G	T	135	DP	35	AF1	1	DP4	2	0	9	14
546600		G	T	139	DP	41	AF1	1	DP4	0	1	23	15

581322		G	A	126	DP	38	AF1	1	DP4	1	0	9	21
674270		G	T	156	DP	43	AF1	1	DP4	0	0	8	27
674300		T	G	168	DP	42	AF1	1	DP4	2	0	14	26
738752		C	A	120	DP	29	AF1	1	DP4	0	1	20	4
741196		G	A	130	DP	23	AF1	1	DP4	2	1	11	8
744351		G	T	129	DP	34	AF1	1	DP4	1	0	14	13
744381		G	T	196	DP	33	AF1	1	DP4	4	0	13	15
793248		G	T	127	DP	25	AF1	1	DP4	3	0	13	6
860802		C	A	162	DP	47	AF1	1	DP4	0	0	28	15
862432		G	T	145	DP	51	AF1	1	DP4	1	0	14	26
864861		C	A	146	DP	43	AF1	1	DP4	0	4	25	10
864902		C	A	132	DP	42	AF1	1	DP4	0	1	26	9
878247		T	A	188	DP	21	AF1	1	DP4	1	1	10	8
896815		C	A	152	DP	27	AF1	1	DP4	0	0	15	8
962502		T	A	197	DP	31	AF1	1	DP4	0	0	15	14
965517		C	T	137	DP	38	AF1	1	DP4	2	0	7	14
965547		C	G	136	DP	36	AF1	1	DP4	4	0	11	16
980347		G	A	121	DP	26	AF1	1	DP4	0	2	16	6
1069320		G	T	169	DP	28	AF1	1	DP4	0	0	9	11
1237550		C	T	208	DP	25	AF1	1	DP4	1	0	13	11
1335112		C	A	120	DP	41	AF1	1	DP4	6	0	6	25
1335164		G	A	148	DP	47	AF1	1	DP4	4	0	10	28
1335180		T	G	125	DP	45	AF1	1	DP4	5	0	10	27
1335671		G	T	172	DP	45	AF1	1	DP4	1	0	21	17
1397024		T	C	165	DP	39	AF1	1	DP4	0	1	14	24
1397045		C	A	185	DP	37	AF1	1	DP4	0	1	16	14
1475574		C	A	154	DP	27	AF1	1	DP4	5	0	8	14
1571924		C	G	134	DP	20	AF1	1	DP4	0	0	6	13
1572282		A	T	134	DP	25	AF1	1	DP4	3	0	9	12
1573470		C	T	174	DP	33	AF1	1	DP4	1	0	7	18
1703116		T	A	132	DP	27	AF1	1	DP4	0	4	11	9
1758224		C	A	146	DP	25	AF1	1	DP4	0	1	10	9

1799369		G	T	184	DP	38	AF1	1	DP4	0	0	13	20
1809620		C	A	171	DP	31	AF1	1	DP4	0	1	16	14
1811616		A	C	125	DP	32	AF1	1	DP4	0	6	16	9
2057138		G	A	207	DP	49	AF1	1	DP4	0	1	20	28
2066438		G	T	209	DP	31	AF1	1	DP4	1	0	15	15
2097571		C	A	171	DP	32	AF1	1	DP4	0	0	13	11
2097597		C	A	121	DP	35	AF1	1	DP4	0	1	15	4
2189746		C	A	182	DP	43	AF1	1	DP4	0	1	18	13
2189786		C	T	148	DP	51	AF1	1	DP4	0	6	24	18
2189822		C	A	121	DP	51	AF1	1	DP4	0	6	24	17
2223724		G	A	132	DP	18	AF1	1	DP4	0	2	12	4
2236405		C	T	222	DP	34	AF1	1	DP4	2	0	15	15
2237224		A	G	223	DP	17	AF1	1	DP4	0	1	6	8
2354398		C	G	127	DP	29	AF1	1	DP4	4	0	11	13
2425459		G	A	163	DP	30	AF1	1	DP4	0	0	7	18
2444726		G	T	222	DP	42	AF1	1	DP4	0	1	23	17
2479804		G	T	166	DP	36	AF1	1	DP4	0	0	13	19
2540626		T	A	189	DP	46	AF1	1	DP4	0	1	17	28
2550855		G	T	189	DP	32	AF1	1	DP4	0	0	10	18
2672922		A	T	133	DP	38	AF1	1	DP4	5	0	15	12
2672973		G	T	208	DP	32	AF1	1	DP4	0	0	9	13
2802767		G	T	157	DP	41	AF1	1	DP4	3	0	8	28
2815961		C	T	222	DP	39	AF1	1	DP4	2	0	14	17
2816041		G	T	203	DP	46	AF1	1	DP4	1	0	17	23
2938650		G	T	146	DP	48	AF1	1	DP4	3	0	20	18
3045816		A	T	139	DP	28	AF1	1	DP4	0	3	18	7
3071946		C	A	222	DP	44	AF1	1	DP4	0	0	17	19
3072011		C	A	137	DP	39	AF1	1	DP4	0	3	16	13
3134983		C	A	145	DP	28	AF1	1	DP4	0	0	13	7
3135616		G	A	174	DP	26	AF1	1	DP4	0	1	15	9
3140485		G	T	148	DP	39	AF1	1	DP4	1	0	12	17
3140536		G	T	203	DP	46	AF1	1	DP4	0	0	18	16

3140580		G	T	174	DP	47	AF1	1	DP4	0	0	19	17
3140590		G	T	162	DP	43	AF1	1	DP4	2	0	20	14
3140622		G	A	162	DP	47	AF1	1	DP4	0	0	17	15
3207901		G	A	172	DP	28	AF1	1	DP4	0	3	15	8
3207927		G	A	146	DP	25	AF1	1	DP4	0	0	14	6
3211599		A	C	222	DP	29	AF1	1	DP4	0	1	8	18
3214464		G	T	125	DP	33	AF1	1	DP4	2	0	6	19
3214655		G	T	125	DP	37	AF1	1	DP4	1	0	5	22
3221765		G	T	162	DP	58	AF1	1	DP4	0	0	9	31
3221835		C	A	186	DP	48	AF1	1	DP4	2	0	7	29
3259370		T	A	132	DP	28	AF1	1	DP4	2	1	14	11
3294999		C	T	222	DP	25	AF1	1	DP4	1	0	9	15
3336165		T	G	141	DP	45	AF1	1	DP4	4	4	23	13
3336525		C	T	142	DP	15	AF1	1	DP4	0	0	2	12
3352082		G	T	222	DP	38	AF1	1	DP4	0	1	22	15
3413710		C	A	123	DP	28	AF1	1	DP4	0	1	17	5
3414982		G	C	164	DP	28	AF1	1	DP4	0	0	15	12
3479283		G	A	132	DP	23	AF1	1	DP4	1	0	12	10
3488769		G	T	128	DP	27	AF1	1	DP4	3	0	9	10
3570583		T	C	140	DP	33	AF1	1	DP4	1	0	13	19
3748362		A	G	222	DP	21	AF1	1	DP4	0	0	11	9
3748416		A	G	222	DP	20	AF1	1	DP4	0	1	10	9
3748449		A	G	222	DP	20	AF1	1	DP4	0	0	11	6
3748626		G	A	222	DP	42	AF1	1	DP4	1	5	17	19
3748665		T	G	222	DP	40	AF1	1	DP4	2	4	16	18
3748707		A	G	222	DP	40	AF1	1	DP4	1	4	16	14
3748749		C	A	147	DP	43	AF1	1	DP4	1	5	20	11
3748926		A	G	222	DP	52	AF1	1	DP4	4	1	28	15
3756214		C	T	131	DP	31	AF1	1	DP4	0	5	18	8
3788980		C	A	124	DP	26	AF1	1	DP4	0	4	13	9
3821351		C	T	222	DP	38	AF1	1	DP4	1	0	16	21
3842391		C	A	129	DP	29	AF1	1	DP4	0	4	17	7

3929823		G	A	133	DP	28	AF1	1	DP4	0	0	8	13
4043865		G	A	126	DP	37	AF1	1	DP4	6	0	9	18
4216645		T	A	222	DP	26	AF1	1	DP4	0	1	17	6
4251083		C	A	133	DP	39	AF1	1	DP4	3	0	9	20
4267192		C	A	179	DP	47	AF1	1	DP4	0	0	18	23
4298264		G	T	131	DP	42	AF1	1	DP4	3	0	10	23
4320362		G	T	128	DP	34	AF1	1	DP4	2	0	5	21
4330494		C	A	222	DP	39	AF1	1	DP4	0	1	20	18
4333367		C	A	186	DP	25	AF1	1	DP4	0	0	4	17
4406815		G	T	167	DP	38	AF1	1	DP4	4	0	15	17
4433119		G	A	164	DP	33	AF1	1	DP4	0	1	15	12
4566191		G	T	167	DP	28	AF1	1	DP4	0	0	11	14
4570591		G	T	160	DP	30	AF1	1	DP4	3	0	13	11
4671423		C	A	167	DP	49	AF1	1	DP4	0	0	24	14
4749357		G	T	147	DP	38	AF1	1	DP4	0	3	21	9
4752396		G	T	179	DP	37	AF1	1	DP4	1	0	14	22
4831670		G	T	189	DP	28	AF1	1	DP4	0	1	12	15
4875819		C	A	164	DP	23	AF1	1	DP4	0	0	13	7
4876029		C	A	158	DP	27	AF1	1	DP4	0	0	7	18
4876050		C	A	179	DP	26	AF1	1	DP4	1	1	6	18
4876136		G	A	144	DP	18	AF1	1	DP4	0	1	0	15
4961429		C	A	168	DP	25	AF1	1	DP4	1	0	5	16
4961509		G	T	201	DP	21	AF1	1	DP4	0	0	6	11
5041004		C	A	222	DP	38	AF1	1	DP4	2	0	19	16
5041772		C	A	136	DP	33	AF1	1	DP4	0	2	18	9
5049581		C	A	158	DP	31	AF1	1	DP4	1	0	9	17
5049950		G	T	130	DP	33	AF1	1	DP4	3	0	12	14
5061527		C	T	159	DP	35	AF1	1	DP4	1	0	19	8
5326531		G	T	123	DP	22	AF1	1	DP4	0	1	8	10
5332930		T	C	222	DP	20	AF1	1	DP4	0	1	5	13
5337431		T	C	222	DP	20	AF1	1	DP4	2	1	7	10
UTI (F13)													

3171736	Phage-related baseplate assembly protein	T	C	147	DP	47	AF1	0.5	DP4	1	8	21	15
3623009	Adhesion YadA precursor	A	G	225	DP	71	AF1	0.5	DP4	11	9	18	22
5008430		G	A	158	DP	76	AF1	0.5	DP4	17	12	25	21
29601		C	G	167	DP	48	AF1	1	DP4	1	1	22	22
126863		C	T	144	DP	42	AF1	1	DP4	0	0	7	26
149809		C	A	167	DP	55	AF1	1	DP4	0	5	24	17
160900		G	A	222	DP	40	AF1	1	DP4	0	1	15	24
195800		T	G	166	DP	33	AF1	1	DP4	1	0	12	20
214200		C	G	222	DP	56	AF1	1	DP4	0	0	26	23
297768		G	T	148	DP	40	AF1	1	DP4	7	0	20	13
848147		C	A	139	DP	23	AF1	1	DP4	0	1	6	12
948184		C	A	133	DP	55	AF1	1	DP4	0	10	28	17
960542		A	G	120	DP	34	AF1	1	DP4	5	0	10	16
960567		G	T	188	DP	33	AF1	1	DP4	0	0	9	15
1271820		G	A	179	DP	26	AF1	1	DP4	1	1	9	10
1431443		G	T	222	DP	53	AF1	1	DP4	0	1	27	11
1471370		T	G	212	DP	38	AF1	1	DP4	0	1	24	13
1732570		C	T	155	DP	67	AF1	1	DP4	1	0	36	29
1795102		A	T	128	DP	47	AF1	1	DP4	5	0	15	26
1825787		G	T	165	DP	36	AF1	1	DP4	2	0	10	20
1938163		C	A	121	DP	33	AF1	1	DP4	0	3	19	10
1941012		C	A	120	DP	31	AF1	1	DP4	0	1	17	5
1990181		C	A	122	DP	33	AF1	1	DP4	0	3	11	11
1997012		C	A	132	DP	39	AF1	1	DP4	1	1	8	24
2014390		C	T	222	DP	32	AF1	1	DP4	0	0	7	17
2445898		A	C	212	DP	37	AF1	1	DP4	0	1	16	17
2460498		C	A	211	DP	41	AF1	1	DP4	1	0	23	17
2472255		G	A	222	DP	30	AF1	1	DP4	0	0	13	12
2753738		A	C	140	DP	25	AF1	1	DP4	0	3	14	8
2753748		G	A	222	DP	25	AF1	1	DP4	0	0	15	5
2791098		T	A	210	DP	52	AF1	1	DP4	0	1	17	34
2797582		A	T	149	DP	24	AF1	1	DP4	1	2	13	8

2929452		A	C	191	DP	34	AF1	1	DP4	1	1	20	11
3247567		G	T	129	DP	57	AF1	1	DP4	4	0	9	35
3276191		T	G	152	DP	31	AF1	1	DP4	0	1	22	4
3440219		G	T	177	DP	42	AF1	1	DP4	0	0	12	22
3444629		G	T	121	DP	42	AF1	1	DP4	0	0	3	25
3483425		G	T	130	DP	31	AF1	1	DP4	4	0	6	18
3547577		G	A	131	DP	38	AF1	1	DP4	0	1	14	16
3623273		G	A	200	DP	56	AF1	1	DP4	5	1	25	23
3623309		G	A	222	DP	53	AF1	1	DP4	1	0	21	23
3634733		G	C	150	DP	42	AF1	1	DP4	0	3	16	17
3688112		G	T	124	DP	46	AF1	1	DP4	0	6	31	8
3688140		A	C	125	DP	50	AF1	1	DP4	0	5	34	6
3753796		C	G	164	DP	31	AF1	1	DP4	1	0	6	15
3770920		C	A	170	DP	23	AF1	1	DP4	0	0	6	12
3856396		A	T	125	DP	47	AF1	1	DP4	0	5	20	18
3897296		A	T	160	DP	48	AF1	1	DP4	1	0	22	24
4009147		G	T	128	DP	37	AF1	1	DP4	5	0	8	21
4097532		G	T	173	DP	60	AF1	1	DP4	1	0	32	23
4097948		C	G	170	DP	37	AF1	1	DP4	0	4	17	12
4176212		C	A	143	DP	34	AF1	1	DP4	1	3	16	6
4231057		C	T	214	DP	38	AF1	1	DP4	1	0	17	20
4303943		C	A	140	DP	28	AF1	1	DP4	3	0	7	12
4310431		C	A	135	DP	36	AF1	1	DP4	0	4	15	9
4472331		G	T	147	DP	43	AF1	1	DP4	5	0	20	16
4521001		T	G	161	DP	28	AF1	1	DP4	0	3	11	14
4521042		T	C	148	DP	26	AF1	1	DP4	0	1	6	13
4567815		C	T	222	DP	27	AF1	1	DP4	1	0	16	10
4662894		C	A	140	DP	66	AF1	1	DP4	0	1	34	23
UTI (M9)													
4829840		T	G	225	DP	53	AF1	0.5	DP4	2	16	20	15
5137026		A	C	225	DP	233	AF1	0.5	DP4	49	22	117	40
5202463	Terminase-like family protein	T	A	225	DP	129	AF1	0.5	DP4	20	30	39	35

5202520	Terminase-like family protein	A	G	225	DP	133	AF1	0.5	DP4	26	23	35	41
5202547	Terminase-like family protein	C	T	225	DP	134	AF1	0.5	DP4	29	20	34	46
5265729		T	C	225	DP	54	AF1	0.5	DP4	13	5	26	10
5265885		T	C	225	DP	61	AF1	0.5	DP4	5	14	17	25
5004916		A	G	199	DP	66	AF1	0.5	DP4	9	7	15	19
5254296		T	C	187	DP	84	AF1	0.5	DP4	4	20	17	27
5233107	Hypothetical protein	C	A	178	DP	115	AF1	0.5	DP4	20	23	28	43
2266421	Small toxic polypeptide LdrD	G	A	173	DP	107	AF1	0.5	DP4	26	13	36	31
2266904	Small toxic polypeptide LdrD	G	A	167	DP	85	AF1	0.5	DP4	12	14	34	25
3723912		T	C	167	DP	68	AF1	0.5	DP4	8	10	29	20
5232987	Hypothetical protein	T	C	165	DP	110	AF1	0.5	DP4	26	17	39	26
3723751		G	A	164	DP	67	AF1	0.5	DP4	8	7	20	30
3723899		C	T	159	DP	71	AF1	0.5	DP4	10	10	29	21
5279414	Phage tail fibre repeat protein	C	T	152	DP	55	AF1	0.5	DP4	16	4	11	24
3723727		C	T	147	DP	71	AF1	0.5	DP4	7	8	25	30
5250759	Is1 transposase	C	T	145	DP	258	AF1	0.5	DP4	53	48	80	67
4829822		G	A	139	DP	58	AF1	0.5	DP4	3	18	18	18
286865		G	T	177	DP	45	AF1	1	DP4	6	0	17	22
287501		G	T	193	DP	46	AF1	1	DP4	0	2	20	23
387844		C	A	144	DP	36	AF1	1	DP4	0	4	14	17
458111		T	G	222	DP	90	AF1	1	DP4	0	1	25	44
613572		C	A	179	DP	44	AF1	1	DP4	3	0	19	20
613626		G	T	123	DP	40	AF1	1	DP4	1	0	10	23
779772		A	G	208	DP	62	AF1	1	DP4	1	1	30	28
792472		C	T	150	DP	56	AF1	1	DP4	0	1	30	23
812766		G	A	222	DP	70	AF1	1	DP4	0	1	48	20
816993		A	G	125	DP	28	AF1	1	DP4	0	1	7	13
875917		T	A	122	DP	48	AF1	1	DP4	0	5	34	7
975729		C	G	127	DP	40	AF1	1	DP4	0	3	29	3
1079488		A	T	133	DP	47	AF1	1	DP4	0	1	29	16
1101798		G	T	188	DP	48	AF1	1	DP4	1	0	8	26
1101819		G	T	156	DP	45	AF1	1	DP4	0	0	11	23

1107536		C	A	127	DP	35	AF1	1	DP4	1	4	6	19
1202989		C	A	147	DP	37	AF1	1	DP4	1	1	20	7
1369518		G	A	122	DP	23	AF1	1	DP4	0	0	8	6
1369544		C	A	180	DP	27	AF1	1	DP4	0	0	14	8
1384808		C	A	124	DP	55	AF1	1	DP4	0	7	29	16
1459031		A	T	157	DP	38	AF1	1	DP4	1	0	15	22
1542288		C	A	128	DP	48	AF1	1	DP4	0	2	18	19
1542310		C	T	139	DP	48	AF1	1	DP4	0	3	24	15
1731936		C	A	121	DP	37	AF1	1	DP4	0	2	16	4
1731949		C	A	130	DP	40	AF1	1	DP4	0	2	19	7
1831796		G	T	175	DP	44	AF1	1	DP4	2	0	18	17
1891118		C	T	222	DP	48	AF1	1	DP4	1	0	22	25
2007189		G	A	139	DP	28	AF1	1	DP4	1	0	5	20
2053015		A	C	222	DP	55	AF1	1	DP4	0	2	19	34
2134651		G	T	149	DP	44	AF1	1	DP4	7	0	13	18
2228893		G	T	124	DP	31	AF1	1	DP4	4	0	6	19
2228928		A	T	129	DP	35	AF1	1	DP4	5	0	9	18
2236253		T	A	129	DP	38	AF1	1	DP4	8	0	14	13
2271681		G	A	193	DP	44	AF1	1	DP4	0	1	19	16
2329392		C	A	126	DP	59	AF1	1	DP4	0	2	25	18
2356127		C	A	135	DP	35	AF1	1	DP4	0	0	22	6
2434005		A	G	217	DP	61	AF1	1	DP4	0	1	25	33
2463527		C	T	158	DP	39	AF1	1	DP4	0	0	20	5
2463560		G	A	126	DP	40	AF1	1	DP4	0	2	16	4
2475405		G	C	222	DP	57	AF1	1	DP4	1	0	24	28
2504405		C	T	222	DP	44	AF1	1	DP4	2	0	17	24
2595657		C	A	165	DP	36	AF1	1	DP4	0	0	21	8
2669779		A	C	122	DP	43	AF1	1	DP4	0	7	24	12
2723652		G	A	222	DP	83	AF1	1	DP4	0	4	47	26
2748998		C	A	144	DP	49	AF1	1	DP4	0	6	23	19
2770346		G	T	136	DP	48	AF1	1	DP4	4	0	14	19
2825032		G	C	222	DP	46	AF1	1	DP4	0	1	20	24

3003084		C	A	161	DP	70	AF1	1	DP4	0	0	37	26
3039231		A	C	131	DP	52	AF1	1	DP4	0	8	24	17
3078191		A	G	152	DP	40	AF1	1	DP4	6	0	6	27
3188453		C	T	149	DP	35	AF1	1	DP4	0	1	24	3
3211733		G	T	154	DP	38	AF1	1	DP4	2	0	11	21
3211744		G	A	196	DP	40	AF1	1	DP4	0	0	14	19
3294744		C	A	174	DP	53	AF1	1	DP4	1	0	19	31
3321468		G	A	175	DP	33	AF1	1	DP4	4	0	11	18
3614249		T	G	135	DP	49	AF1	1	DP4	2	0	8	30
3810438		G	T	142	DP	38	AF1	1	DP4	0	0	21	5
3884238		A	C	209	DP	54	AF1	1	DP4	0	1	36	17
4190647		C	T	222	DP	64	AF1	1	DP4	1	0	16	35
4190663		G	T	128	DP	62	AF1	1	DP4	3	0	11	33
4190684		G	T	152	DP	62	AF1	1	DP4	0	1	12	33
4206647		C	A	211	DP	50	AF1	1	DP4	1	0	23	26
4211712		C	A	163	DP	53	AF1	1	DP4	0	1	26	20
4302704		A	C	146	DP	51	AF1	1	DP4	3	0	8	23
4302739		C	A	222	DP	45	AF1	1	DP4	0	0	12	21
4302814		G	T	161	DP	53	AF1	1	DP4	1	0	14	28
4370303		A	C	179	DP	38	AF1	1	DP4	0	4	18	16
4397112		G	T	150	DP	35	AF1	1	DP4	2	0	9	17
4429617		C	A	133	DP	44	AF1	1	DP4	0	3	25	7
4470685		T	A	222	DP	49	AF1	1	DP4	0	3	21	20
4576030		A	T	193	DP	66	AF1	1	DP4	0	1	26	39
4659608		G	A	222	DP	41	AF1	1	DP4	0	2	22	9
4659632		C	A	124	DP	46	AF1	1	DP4	0	3	27	8
4710806		C	A	168	DP	56	AF1	1	DP4	0	4	28	16
4756148		G	T	130	DP	55	AF1	1	DP4	9	0	15	27
4778886		T	G	156	DP	44	AF1	1	DP4	1	0	19	24
4996530		G	T	140	DP	105	AF1	1	DP4	1	0	8	74
5265408		C	T	139	DP	58	AF1	1	DP4	0	0	14	14
5265431		C	A	136	DP	56	AF1	1	DP4	0	17	14	24

5265457		T	A	152	DP	48	AF1	1	DP4	0	8	12	26
UTI (M6)													
2124178		G	T	128	DP	30	AF1	0.5016	DP4	0	7	16	7
55781		G	T	123	DP	42	AF1	1	DP4	3	0	7	22
307920		G	T	147	DP	44	AF1	1	DP4	5	0	11	26
358482		C	T	136	DP	44	AF1	1	DP4	0	0	24	3
529095		C	T	176	DP	34	AF1	1	DP4	0	0	14	10
552871		C	T	176	DP	34	AF1	1	DP4	5	0	11	17
552902		G	T	132	DP	30	AF1	1	DP4	1	0	4	16
780451		G	T	121	DP	23	AF1	1	DP4	2	1	7	11
843789		C	T	156	DP	48	AF1	1	DP4	0	4	24	14
909423		G	T	143	DP	30	AF1	1	DP4	0	2	17	10
909494		G	T	127	DP	30	AF1	1	DP4	0	4	15	7
909529		C	G	127	DP	29	AF1	1	DP4	0	2	13	6
959143		G	T	144	DP	35	AF1	1	DP4	3	0	7	16
959253		G	T	124	DP	32	AF1	1	DP4	2	1	7	16
959323		G	T	201	DP	30	AF1	1	DP4	0	0	9	19
1109223		T	G	166	DP	45	AF1	1	DP4	2	0	19	21
1186770		G	T	140	DP	26	AF1	1	DP4	0	3	11	10
1190753		G	T	138	DP	33	AF1	1	DP4	5	0	8	20
1208837		G	T	170	DP	48	AF1	1	DP4	6	0	18	24
1209024		C	A	155	DP	63	AF1	1	DP4	0	8	28	26
1321029		G	A	120	DP	23	AF1	1	DP4	0	3	17	3
1392107		G	T	144	DP	32	AF1	1	DP4	3	0	9	12
1392127		G	T	138	DP	35	AF1	1	DP4	5	0	9	17
1668780		G	T	137	DP	27	AF1	1	DP4	1	0	10	14
1768559		C	A	137	DP	37	AF1	1	DP4	0	4	21	9
1843115		G	A	169	DP	29	AF1	1	DP4	0	1	17	7
1936581		G	A	124	DP	22	AF1	1	DP4	0	0	11	7
1980299		C	A	121	DP	32	AF1	1	DP4	0	1	20	3
2008000		G	T	147	DP	49	AF1	1	DP4	7	0	16	26
2152268		G	T	128	DP	28	AF1	1	DP4	0	3	19	6

2216464		C	G	120	DP	31	AF1	1	DP4	1	1	13	5
2436474		A	C	126	DP	21	AF1	1	DP4	0	2	11	7
2484492		G	A	192	DP	48	AF1	1	DP4	0	1	28	17
2514978		C	A	120	DP	63	AF1	1	DP4	0	8	34	19
2585635		G	T	141	DP	32	AF1	1	DP4	5	0	9	16
2614727		G	T	167	DP	35	AF1	1	DP4	0	0	16	11
2634289		C	T	222	DP	48	AF1	1	DP4	0	0	15	25
2678433		G	C	217	DP	52	AF1	1	DP4	0	1	24	26
2721741		A	T	178	DP	46	AF1	1	DP4	1	0	7	23
2721852		G	T	156	DP	34	AF1	1	DP4	3	0	9	21
2854407		A	C	142	DP	33	AF1	1	DP4	5	0	12	14
2854476		G	A	203	DP	37	AF1	1	DP4	0	0	12	14
2929539		T	G	131	DP	35	AF1	1	DP4	6	0	13	15
2929593		G	T	163	DP	26	AF1	1	DP4	0	0	9	13
3018212		C	A	206	DP	47	AF1	1	DP4	0	0	16	29
3046012		C	A	157	DP	32	AF1	1	DP4	0	2	15	13
3046051		G	A	222	DP	28	AF1	1	DP4	0	0	13	5
3223185		T	C	178	DP	39	AF1	1	DP4	0	0	23	16
3249208		C	A	212	DP	48	AF1	1	DP4	4	0	22	20
3258735		C	A	161	DP	23	AF1	1	DP4	0	0	7	9
3258834		C	A	155	DP	24	AF1	1	DP4	2	0	9	11
3334449		G	T	134	DP	39	AF1	1	DP4	10	0	12	17
3395432		G	T	133	DP	40	AF1	1	DP4	5	0	15	15
3455042		G	T	123	DP	34	AF1	1	DP4	1	0	4	21
3459783		G	A	173	DP	30	AF1	1	DP4	0	1	20	3
3497165		G	A	139	DP	29	AF1	1	DP4	2	0	5	17
3505667		G	T	164	DP	27	AF1	1	DP4	2	1	9	15
3650538		A	G	153	DP	38	AF1	1	DP4	0	3	24	8
3911951		G	T	139	DP	37	AF1	1	DP4	2	0	8	19
4024126		C	A	138	DP	48	AF1	1	DP4	1	0	21	23
4152074		C	T	197	DP	50	AF1	1	DP4	1	0	15	32
4270553		A	G	121	DP	39	AF1	1	DP4	0	7	24	8

4418886		G	T	174	DP	37	AF1	1	DP4	1	0	12	21
4419293		G	T	125	DP	35	AF1	1	DP4	4	0	19	12
4435494		C	A	122	DP	42	AF1	1	DP4	0	2	22	8
4557952		G	T	136	DP	22	AF1	1	DP4	0	1	15	3
4561740		G	T	143	DP	30	AF1	1	DP4	3	0	11	10
4663428		G	C	147	DP	44	AF1	1	DP4	0	0	21	22
4708516		C	A	134	DP	37	AF1	1	DP4	0	0	22	10
4736151		T	C	171	DP	27	AF1	1	DP4	1	0	11	15
4753666		G	T	143	DP	28	AF1	1	DP4	1	0	16	4
4753693		C	A	175	DP	32	AF1	1	DP4	0	5	15	11
4825763		T	C	122	DP	28	AF1	1	DP4	2	1	14	9
UTI (M11)													
3973036	Hypothetical protein	A	T	225	DP	144	AF1	0.5	DP4	21	14	51	56
5217992		G	A	225	DP	589	AF1	0.5	DP4	123	133	138	167
5237315	Integrase core domain protein	C	A	225	DP	218	AF1	0.5	DP4	48	41	60	64
5258646		G	C	225	DP	138	AF1	0.5	DP4	31	24	47	32
5277619		G	T	225	DP	199	AF1	0.5	DP4	46	35	68	50
5277696	Hypothetical protein	G	A	225	DP	214	AF1	0.5	DP4	29	43	68	72
4634033	Is1 transposase	C	A	145	DP	150	AF1	0.5	DP4	17	31	39	58
5244417		C	A	132	DP	229	AF1	0.5	DP4	37	61	51	78
210069		C	G	125	DP	36	AF1	1	DP4	4	0	4	25
392889		C	T	157	DP	70	AF1	1	DP4	1	7	46	10
498197		C	A	166	DP	76	AF1	1	DP4	0	4	41	16
749353		T	C	222	DP	64	AF1	1	DP4	1	0	27	35
958681		C	A	156	DP	85	AF1	1	DP4	0	0	47	23
958792		C	A	136	DP	87	AF1	1	DP4	0	7	39	18
1553651		C	A	168	DP	53	AF1	1	DP4	0	7	31	14
1654597		C	A	147	DP	78	AF1	1	DP4	1	2	43	13
1785108		C	T	127	DP	86	AF1	1	DP4	0	1	35	49
1896567		C	T	136	DP	83	AF1	1	DP4	0	1	41	7
1962825		C	A	139	DP	90	AF1	1	DP4	0	6	41	20
1974638		C	A	161	DP	87	AF1	1	DP4	0	1	41	21

1986247		T	G	157	DP	49	AF1	1	DP4	0	8	27	12
2138042		C	A	151	DP	75	AF1	1	DP4	0	9	39	20
2313844		C	T	222	DP	97	AF1	1	DP4	1	0	42	52
2479926		C	T	135	DP	56	AF1	1	DP4	0	1	37	6
3319347		G	A	143	DP	74	AF1	1	DP4	0	4	46	12
3607209		G	C	201	DP	91	AF1	1	DP4	0	1	47	37
3726034		C	A	132	DP	68	AF1	1	DP4	0	11	27	26
3875488		T	A	222	DP	33	AF1	1	DP4	0	0	7	24
3929180		G	T	180	DP	46	AF1	1	DP4	1	0	10	21
4304387		G	T	187	DP	59	AF1	1	DP4	0	1	37	17
4304408		C	T	222	DP	57	AF1	1	DP4	0	1	38	18
4390262		A	C	206	DP	55	AF1	1	DP4	0	1	28	25
5159138		C	A	169	DP	102	AF1	1	DP4	0	5	54	29
UTI (F20)													
2783629		A	C	142	DP	26	AF1	0.502	DP4	3	1	11	11
4392589	Phage tail fibre repeat protein	T	A	145	DP	24	AF1	0.6243	DP4	2	1	12	8
2783570		C	T	146	DP	35	AF1	0.5	DP4	5	5	10	14
2783583		T	C	157	DP	34	AF1	0.5	DP4	5	4	10	15
2783398		C	T	159	DP	32	AF1	0.5	DP4	0	5	12	15
2783422		G	A	170	DP	31	AF1	0.5	DP4	2	6	11	11
212101		A	C	184	DP	19	AF1	1	DP4	0	1	9	9
212801		A	T	222	DP	15	AF1	1	DP4	1	0	6	7
393082		C	T	121	DP	17	AF1	1	DP4	0	0	7	7
561459		A	C	141	DP	21	AF1	1	DP4	0	3	7	11
779189		C	A	159	DP	23	AF1	1	DP4	1	0	4	13
923153		C	G	143	DP	21	AF1	1	DP4	0	2	12	5
1376749		C	T	222	DP	22	AF1	1	DP4	2	0	5	14
1621327		A	T	126	DP	26	AF1	1	DP4	0	0	10	11
1736124		C	A	121	DP	22	AF1	1	DP4	0	4	13	5
1831968		C	A	195	DP	20	AF1	1	DP4	0	1	9	10
1931625		G	T	140	DP	24	AF1	1	DP4	0	0	4	13
1990015		G	A	172	DP	21	AF1	1	DP4	0	1	9	7

2613458		C	A	125	DP	21	AF1	1	DP4	0	1	11	6
2989307		C	A	151	DP	18	AF1	1	DP4	1	0	5	12
3100700		C	A	133	DP	19	AF1	1	DP4	2	0	7	8
3396836		C	T	132	DP	25	AF1	1	DP4	2	0	5	15
3396863		T	G	137	DP	22	AF1	1	DP4	3	0	7	11
3447111		G	T	206	DP	22	AF1	1	DP4	0	0	8	11
3649843		G	T	135	DP	22	AF1	1	DP4	0	0	7	11
3857621		C	T	195	DP	24	AF1	1	DP4	1	0	12	11
3954814		T	A	133	DP	19	AF1	1	DP4	0	1	11	6
4138491		T	C	150	DP	19	AF1	1	DP4	1	0	9	9
4259943		G	T	144	DP	26	AF1	1	DP4	1	0	7	7
4274423		A	G	196	DP	23	AF1	1	DP4	0	1	10	12
4663744		G	A	121	DP	13	AF1	1	DP4	0	1	9	3
4670344		A	G	222	DP	27	AF1	1	DP4	0	1	18	8
4674188		G	T	120	DP	26	AF1	1	DP4	0	3	18	5
4674289		C	A	124	DP	25	AF1	1	DP4	0	0	18	3
4674444		T	G	222	DP	20	AF1	1	DP4	1	0	4	15
4691800		T	C	142	DP	21	AF1	1	DP4	3	0	8	10
4860162		C	A	124	DP	22	AF1	1	DP4	0	1	11	7
4887512		C	A	151	DP	23	AF1	1	DP4	0	3	11	8
4988187		C	A	166	DP	19	AF1	1	DP4	0	0	10	5
4988284		T	A	144	DP	16	AF1	1	DP4	1	0	6	9
UTI (M8)													
4978588		G	T	225	DP	135	AF1	0.5	DP4	26	24	40	38
4978628		A	G	225	DP	123	AF1	0.5	DP4	29	13	37	34
4973249	Filamentous hemagglutinin	G	A	176	DP	99	AF1	0.5	DP4	15	22	29	27
4126341		T	C	160	DP	68	AF1	0.5	DP4	14	10	21	22
5052692	Integrase core domain protein	A	G	146	DP	109	AF1	0.5	DP4	18	23	37	24
4126505		C	T	131	DP	53	AF1	0.5	DP4	5	5	24	15
37992		C	A	172	DP	33	AF1	1	DP4	2	0	8	17
38199		G	T	131	DP	38	AF1	1	DP4	6	0	6	26
100839		G	A	131	DP	34	AF1	1	DP4	0	5	15	14

200736		G	T	126	DP	29	AF1	1	DP4	1	0	4	16
251179		T	G	160	DP	38	AF1	1	DP4	1	0	8	22
338901		C	G	221	DP	28	AF1	1	DP4	1	0	2	23
425565		G	T	151	DP	24	AF1	1	DP4	0	2	18	1
428201		C	G	194	DP	44	AF1	1	DP4	1	0	18	19
445673		C	A	137	DP	47	AF1	1	DP4	0	1	30	7
466454		G	T	146	DP	24	AF1	1	DP4	1	0	9	6
632301		G	T	222	DP	44	AF1	1	DP4	0	1	23	18
746646		C	T	221	DP	28	AF1	1	DP4	0	0	13	14
813146		C	G	154	DP	27	AF1	1	DP4	1	0	14	11
1107838		G	T	131	DP	34	AF1	1	DP4	1	0	10	20
1136246		T	A	200	DP	32	AF1	1	DP4	1	0	14	10
1169696		G	T	202	DP	42	AF1	1	DP4	0	0	16	17
1170610		C	A	124	DP	40	AF1	1	DP4	0	5	20	13
1362316		G	T	222	DP	45	AF1	1	DP4	0	1	22	20
1414265		A	G	140	DP	47	AF1	1	DP4	8	0	15	23
1548616		C	A	165	DP	29	AF1	1	DP4	0	2	14	11
1611438		G	T	142	DP	41	AF1	1	DP4	2	0	12	17
1768145		G	T	222	DP	38	AF1	1	DP4	0	1	11	25
1880620		G	A	122	DP	49	AF1	1	DP4	1	1	8	29
1880634		C	A	149	DP	47	AF1	1	DP4	6	0	11	29
1881996		C	A	142	DP	49	AF1	1	DP4	7	0	21	21
2025045		C	A	133	DP	40	AF1	1	DP4	0	2	17	10
2083120		A	C	222	DP	27	AF1	1	DP4	0	1	12	12
2132115		C	T	222	DP	28	AF1	1	DP4	1	0	12	13
2158246		A	C	163	DP	29	AF1	1	DP4	0	0	6	19
2160197		C	T	127	DP	28	AF1	1	DP4	0	0	4	15
2228625		C	A	176	DP	44	AF1	1	DP4	0	3	24	14
2325304		C	A	127	DP	23	AF1	1	DP4	0	1	13	6
2607822		C	A	205	DP	34	AF1	1	DP4	0	2	20	9
2651910		G	T	140	DP	29	AF1	1	DP4	3	0	13	13
2833461		C	G	122	DP	37	AF1	1	DP4	3	0	16	9

2833510		C	G	122	DP	46	AF1	1	DP4	0	0	3	23
2833755		C	A	129	DP	48	AF1	1	DP4	0	2	25	13
3138908		G	A	183	DP	39	AF1	1	DP4	1	0	16	21
3331815		G	T	138	DP	45	AF1	1	DP4	0	5	29	11
3331848		C	A	127	DP	53	AF1	1	DP4	0	5	30	8
3506271		A	G	209	DP	41	AF1	1	DP4	2	0	25	14
3542071		A	G	222	DP	75	AF1	1	DP4	1	0	28	43
3593144		G	T	146	DP	41	AF1	1	DP4	3	0	9	24
3593188		G	C	122	DP	26	AF1	1	DP4	1	1	11	11
3593210		A	C	126	DP	25	AF1	1	DP4	3	0	9	12
3661117		C	T	222	DP	60	AF1	1	DP4	1	0	25	33
3748117		A	T	211	DP	41	AF1	1	DP4	0	1	15	23
3779900		C	A	123	DP	43	AF1	1	DP4	4	0	15	18
3779959		C	G	168	DP	44	AF1	1	DP4	2	0	13	18
3780759		A	T	209	DP	44	AF1	1	DP4	0	1	26	17
3844124		C	T	177	DP	25	AF1	1	DP4	0	0	5	16
3848345		A	T	203	DP	24	AF1	1	DP4	5	0	7	12
3888111		C	T	123	DP	36	AF1	1	DP4	0	1	20	11
3888140		C	A	124	DP	32	AF1	1	DP4	0	2	20	6
3888155		C	T	140	DP	33	AF1	1	DP4	0	3	20	8
3958832		G	T	133	DP	35	AF1	1	DP4	1	0	4	25
4126591		C	T	128	DP	56	AF1	1	DP4	1	5	24	20
4126645		C	A	158	DP	61	AF1	1	DP4	1	1	34	8
4203376		C	A	158	DP	37	AF1	1	DP4	0	1	21	8
4204712		A	T	222	DP	65	AF1	1	DP4	2	0	26	33
4281146		G	T	150	DP	31	AF1	1	DP4	2	0	7	15
4356784		G	T	138	DP	39	AF1	1	DP4	4	0	10	20
4414750		A	C	131	DP	36	AF1	1	DP4	0	5	21	10
4539349		G	T	145	DP	34	AF1	1	DP4	5	0	10	18
4556519		T	C	129	DP	33	AF1	1	DP4	8	0	8	16
4599104		G	A	222	DP	33	AF1	1	DP4	0	2	14	13
4756399		C	A	164	DP	35	AF1	1	DP4	1	0	16	16

4756899		C	T	222	DP	36	AF1	1	DP4	1	0	12	23
4764199		C	A	154	DP	38	AF1	1	DP4	1	0	22	15
UTI (F21)													
1556195		G	A	143	DP	49	AF1	0.8294	DP4	1	3	18	20
1555946		G	A	133	DP	40	AF1	0.504	DP4	0	5	11	21
222		A	G	197	DP	51	AF1	0.5	DP4	3	7	26	15
424173		T	A	225	DP	210	AF1	0.5	DP4	45	34	84	41
1556088		C	A	166	DP	60	AF1	0.5	DP4	8	7	15	23
4825834		G	A	225	DP	35	AF1	0.5	DP4	5	5	20	5
5122322		A	G	225	DP	339	AF1	0.5	DP4	61	48	105	119
5134146		A	C	122	DP	142	AF1	0.5	DP4	29	15	56	32
5134281		T	C	146	DP	141	AF1	0.5	DP4	27	17	47	47
5134827		G	A	176	DP	109	AF1	0.5	DP4	18	22	35	31
5135301		A	G	168	DP	108	AF1	0.5	DP4	14	14	34	33
5136859		T	C	225	DP	153	AF1	0.5	DP4	32	24	44	48
5135859	gpW	A	C	225	DP	120	AF1	0.5	DP4	23	25	35	35
5198830	Peptidase family S49	A	G	155	DP	106	AF1	0.5	DP4	15	16	43	31
5135240	Phage terminase large subunit (GpA)	T	C	172	DP	110	AF1	0.5	DP4	22	18	32	36
259848		C	A	131	DP	41	AF1	1	DP4	1	1	26	7
516446		G	C	210	DP	50	AF1	1	DP4	1	0	22	27
788973		G	C	126	DP	37	AF1	1	DP4	4	0	4	24
988085		C	T	136	DP	34	AF1	1	DP4	0	0	22	2
988101		T	A	120	DP	40	AF1	1	DP4	0	7	21	9
988151		G	T	222	DP	40	AF1	1	DP4	0	3	18	11
992612		G	T	139	DP	37	AF1	1	DP4	4	0	8	18
1083191		G	T	170	DP	52	AF1	1	DP4	3	0	16	22
1083377		G	T	151	DP	46	AF1	1	DP4	3	0	17	21
1083400		G	T	158	DP	48	AF1	1	DP4	1	0	17	23
1083623		G	A	222	DP	38	AF1	1	DP4	0	1	14	15
1400608		C	A	144	DP	44	AF1	1	DP4	0	6	25	11
1453483		C	A	137	DP	37	AF1	1	DP4	1	0	14	12
1458548		A	T	148	DP	70	AF1	1	DP4	0	0	35	33

1786317		G	T	222	DP	43	AF1	1	DP4	0	1	27	11
1844717		T	A	216	DP	27	AF1	1	DP4	1	0	13	13
1908978		C	A	135	DP	29	AF1	1	DP4	0	4	17	6
1942109		C	A	121	DP	31	AF1	1	DP4	0	6	18	6
1958785		T	A	222	DP	42	AF1	1	DP4	0	0	9	32
1963613		T	G	159	DP	48	AF1	1	DP4	0	5	31	10
2002484		C	T	172	DP	50	AF1	1	DP4	1	0	27	22
2035284		C	A	166	DP	37	AF1	1	DP4	0	5	19	10
2092198		G	A	165	DP	32	AF1	1	DP4	1	0	13	16
2343592		C	T	162	DP	38	AF1	1	DP4	0	1	19	9
2380624		C	A	180	DP	43	AF1	1	DP4	0	1	18	24
2515283		G	C	179	DP	39	AF1	1	DP4	3	1	14	20
2624310		G	T	139	DP	37	AF1	1	DP4	1	0	9	22
2641953		G	T	136	DP	29	AF1	1	DP4	2	0	7	16
2641964		G	T	128	DP	29	AF1	1	DP4	4	0	8	16
2969698		G	T	149	DP	43	AF1	1	DP4	0	0	31	3
3009452		C	A	139	DP	35	AF1	1	DP4	0	3	10	9
3123162		C	G	210	DP	43	AF1	1	DP4	1	0	12	27
3136465		G	T	122	DP	54	AF1	1	DP4	2	0	8	18
3177365		G	T	157	DP	31	AF1	1	DP4	2	3	15	11
3189987		G	T	151	DP	39	AF1	1	DP4	0	1	26	4
3267198		C	A	150	DP	41	AF1	1	DP4	0	8	18	14
3270345		A	T	124	DP	46	AF1	1	DP4	8	0	9	26
3318053		C	A	175	DP	63	AF1	1	DP4	0	1	29	18
3333954		A	C	155	DP	42	AF1	1	DP4	0	2	28	5
3493271		C	A	136	DP	35	AF1	1	DP4	0	6	16	12
3497829		G	T	128	DP	29	AF1	1	DP4	0	5	13	11
3568183		G	T	123	DP	33	AF1	1	DP4	5	0	13	11
3652959		C	T	120	DP	42	AF1	1	DP4	0	3	29	4
3719608		C	A	144	DP	25	AF1	1	DP4	3	0	4	18
3719635		C	G	125	DP	24	AF1	1	DP4	3	0	6	15
3719778		A	C	182	DP	25	AF1	1	DP4	0	1	15	8

3977567		T	G	196	DP	51	AF1	1	DP4	0	1	19	28
4121066		T	G	152	DP	30	AF1	1	DP4	5	0	7	18
4121581		C	G	217	DP	61	AF1	1	DP4	0	0	37	21
4129037		T	A	222	DP	57	AF1	1	DP4	0	1	27	25
4129048		C	A	140	DP	57	AF1	1	DP4	0	2	31	14
4299171		C	A	144	DP	25	AF1	1	DP4	0	1	19	5
4339461		G	A	127	DP	45	AF1	1	DP4	0	1	24	20
4778111		G	T	143	DP	55	AF1	1	DP4	7	0	14	34
4778255		G	T	134	DP	59	AF1	1	DP4	5	0	24	26
4940652		C	T	133	DP	45	AF1	1	DP4	1	0	7	26
UTI (M21)													
895266	Dihydrolipoyllysine-residue acetyltransferase component of pyruvate dehydrogenase complex	G	T	225	DP	27	AF1	0.53	DP4	1	4	7	11
2675063	Hypothetical protein	A	G	225	DP	43	AF1	0.5	DP4	3	2	20	18
2873142	Putative autotransporter precursor	C	T	225	DP	109	AF1	0.5	DP4	15	26	30	32
5462324		C	T	225	DP	92	AF1	0.5	DP4	12	18	27	34
5588730	Hypothetical protein	A	G	176	DP	116	AF1	0.5	DP4	26	7	36	21
2981086		G	A	169	DP	73	AF1	0.5	DP4	12	5	22	31
5421648		T	C	165	DP	110	AF1	0.5	DP4	18	25	29	37
5454617	Hypothetical protein	G	A	158	DP	102	AF1	0.5	DP4	27	11	47	15
2981234		C	T	152	DP	65	AF1	0.5	DP4	11	7	25	22
4626896	Small toxic polypeptide LdrD	G	A	149	DP	120	AF1	0.5	DP4	20	16	45	36
2981247		T	C	148	DP	68	AF1	0.5	DP4	9	6	27	25
5355859		A	G	136	DP	156	AF1	0.5	DP4	29	31	42	50
5605170		T	G	136	DP	74	AF1	0.5	DP4	18	9	26	21
2981062		C	T	135	DP	81	AF1	0.5	DP4	11	6	27	37
259848		C	A	131	DP	41	AF1	1	DP4	1	1	26	7
516446		G	C	210	DP	50	AF1	1	DP4	1	0	22	27
788973		G	C	126	DP	37	AF1	1	DP4	4	0	4	24
988085		C	T	136	DP	34	AF1	1	DP4	0	0	22	2
988101		T	A	120	DP	40	AF1	1	DP4	0	7	21	9
988151		G	T	222	DP	40	AF1	1	DP4	0	3	18	11
992612		G	T	139	DP	37	AF1	1	DP4	4	0	8	18

1083191		G	T	170	DP	52	AF1	1	DP4	3	0	16	22
1083377		G	T	151	DP	46	AF1	1	DP4	3	0	17	21
1083400		G	T	158	DP	48	AF1	1	DP4	1	0	17	23
1083623		G	A	222	DP	38	AF1	1	DP4	0	1	14	15
1400608		C	A	144	DP	44	AF1	1	DP4	0	6	25	11
1453483		C	A	137	DP	37	AF1	1	DP4	1	0	14	12
1458548		A	T	148	DP	70	AF1	1	DP4	0	0	35	33
1786317		G	T	222	DP	43	AF1	1	DP4	0	1	27	11
1844717		T	A	216	DP	27	AF1	1	DP4	1	0	13	13
1908978		C	A	135	DP	29	AF1	1	DP4	0	4	17	6
1942109		C	A	121	DP	31	AF1	1	DP4	0	6	18	6
1958785		T	A	222	DP	42	AF1	1	DP4	0	0	9	32
1963613		T	G	159	DP	48	AF1	1	DP4	0	5	31	10
2002484		C	T	172	DP	50	AF1	1	DP4	1	0	27	22
2035284		C	A	166	DP	37	AF1	1	DP4	0	5	19	10
2092198		G	A	165	DP	32	AF1	1	DP4	1	0	13	16
2343592		C	T	162	DP	38	AF1	1	DP4	0	1	19	9
2380624		C	A	180	DP	43	AF1	1	DP4	0	1	18	24
2515283		G	C	179	DP	39	AF1	1	DP4	3	1	14	20
2624310		G	T	139	DP	37	AF1	1	DP4	1	0	9	22
2641953		G	T	136	DP	29	AF1	1	DP4	2	0	7	16
2641964		G	T	128	DP	29	AF1	1	DP4	4	0	8	16
2969698		G	T	149	DP	43	AF1	1	DP4	0	0	31	3
3009452		C	A	139	DP	35	AF1	1	DP4	0	3	10	9
3123162		C	G	210	DP	43	AF1	1	DP4	1	0	12	27
3136465		G	T	122	DP	54	AF1	1	DP4	2	0	8	18
3177365		G	T	157	DP	31	AF1	1	DP4	2	3	15	11
3189987		G	T	151	DP	39	AF1	1	DP4	0	1	26	4
3267198		C	A	150	DP	41	AF1	1	DP4	0	8	18	14
3270345		A	T	124	DP	46	AF1	1	DP4	8	0	9	26
3318053		C	A	175	DP	63	AF1	1	DP4	0	1	29	18
3333954		A	C	155	DP	42	AF1	1	DP4	0	2	28	5

3493271		C	A	136	DP	35	AF1	1	DP4	0	6	16	12
3497829		G	T	128	DP	29	AF1	1	DP4	0	5	13	11
3568183		G	T	123	DP	33	AF1	1	DP4	5	0	13	11
3652959		C	T	120	DP	42	AF1	1	DP4	0	3	29	4
3719608		C	A	144	DP	25	AF1	1	DP4	3	0	4	18
3719635		C	G	125	DP	24	AF1	1	DP4	3	0	6	15
3719778		A	C	182	DP	25	AF1	1	DP4	0	1	15	8
3977567		T	G	196	DP	51	AF1	1	DP4	0	1	19	28
4121066		T	G	152	DP	30	AF1	1	DP4	5	0	7	18
4121581		C	G	217	DP	61	AF1	1	DP4	0	0	37	21
4129037		T	A	222	DP	57	AF1	1	DP4	0	1	27	25
4129048		C	A	140	DP	57	AF1	1	DP4	0	2	31	14
4299171		C	A	144	DP	25	AF1	1	DP4	0	1	19	5
4339461		G	A	127	DP	45	AF1	1	DP4	0	1	24	20
4778111		G	T	143	DP	55	AF1	1	DP4	7	0	14	34
4778255		G	T	134	DP	59	AF1	1	DP4	5	0	24	26
4940652		C	T	133	DP	45	AF1	1	DP4	1	0	7	26
UTI (M12)													
4930263		A	C	170	DP	143	AF1	0.5	DP4	30	28	41	41
4933235	C1amp-binding protein CrfC	T	G	173	DP	153	AF1	0.5	DP4	24	36	45	43
4934934		T	C	225	DP	179	AF1	0.5	DP4	42	32	49	51
284418		T	C	152	DP	86	AF1	1	DP4	10	0	26	46
285247		C	T	170	DP	61	AF1	1	DP4	0	8	32	13
285268		G	T	124	DP	54	AF1	1	DP4	0	10	27	16
600227		C	A	140	DP	83	AF1	1	DP4	0	4	42	23
624159		C	A	133	DP	54	AF1	1	DP4	0	7	23	14
705810		G	T	157	DP	68	AF1	1	DP4	2	0	18	38
937985		C	A	162	DP	75	AF1	1	DP4	0	3	34	21
940744		G	T	168	DP	59	AF1	1	DP4	6	0	19	29
1105139		G	T	147	DP	63	AF1	1	DP4	0	3	33	13
1105152		C	A	134	DP	64	AF1	1	DP4	0	2	33	21
1454171		C	A	157	DP	79	AF1	1	DP4	10	0	25	40

1454248		A	T	165	DP	94	AF1	1	DP4	4	0	38	49
1810872		G	T	182	DP	88	AF1	1	DP4	11	0	32	42
1966992		C	A	161	DP	63	AF1	1	DP4	0	9	29	15
1990180		T	C	222	DP	70	AF1	1	DP4	0	1	38	29
2210963		T	A	207	DP	52	AF1	1	DP4	1	0	26	16
2405697		C	A	180	DP	89	AF1	1	DP4	0	8	50	22
2478362		C	A	177	DP	65	AF1	1	DP4	0	4	41	14
2519606		G	T	221	DP	64	AF1	1	DP4	2	0	19	31
2519725		G	T	124	DP	90	AF1	1	DP4	0	0	30	50
2524328		C	T	163	DP	78	AF1	1	DP4	0	3	42	16
2670904		G	A	222	DP	111	AF1	1	DP4	0	0	47	58
3064461		A	T	201	DP	68	AF1	1	DP4	1	0	33	33
3215959		G	T	184	DP	54	AF1	1	DP4	6	0	24	24
3377403		G	T	166	DP	70	AF1	1	DP4	8	0	21	40
3377431		G	T	151	DP	76	AF1	1	DP4	6	0	24	36
3395737		C	A	164	DP	73	AF1	1	DP4	0	9	41	21
3554920		G	A	222	DP	98	AF1	1	DP4	0	1	42	49
3616620		A	C	220	DP	41	AF1	1	DP4	0	1	20	16
3651252		G	A	177	DP	70	AF1	1	DP4	0	0	36	20
3796843		T	G	153	DP	74	AF1	1	DP4	10	0	22	40
3900140		C	A	147	DP	76	AF1	1	DP4	0	6	36	26
4043641		G	T	147	DP	96	AF1	1	DP4	8	0	29	51
4043690		A	T	222	DP	85	AF1	1	DP4	1	0	37	47
4300961		T	A	222	DP	91	AF1	1	DP4	0	1	51	37
4420685		G	A	222	DP	101	AF1	1	DP4	0	1	49	47
4587187		G	T	144	DP	33	AF1	1	DP4	3	0	14	16
4587214		G	T	143	DP	27	AF1	1	DP4	3	0	12	12
4663977		C	A	124	DP	66	AF1	1	DP4	0	2	27	9
4877373		G	T	147	DP	81	AF1	1	DP4	3	0	26	42
UTI (F15)													
5179020		G	A	225	DP	111	AF1	0.5	DP4	17	25	24	38
2214404	Fibronectin type III protein	C	T	225	DP	34	AF1	0.5	DP4	7	1	6	17

5179203	gpW	A	C	225	DP	94	AF1	0.5	DP4	19	14	34	24
5190670	Exodeoxyribonuclease 8	C	T	225	DP	123	AF1	0.5	DP4	26	23	43	30
5209070	Putative transposase	A	C	225	DP	126	AF1	0.5	DP4	13	32	31	41
5209303		A	G	225	DP	103	AF1	0.5	DP4	22	15	34	32
5228578		T	A	225	DP	225	AF1	0.5	DP4	23	42	52	106
5230869		A	G	225	DP	94	AF1	0.5	DP4	14	18	29	33
5261372	Hypothetical protein	A	G	225	DP	151	AF1	0.5	DP4	25	27	41	36
5261519	Hypothetical protein	A	G	225	DP	181	AF1	0.5	DP4	33	38	53	45
5269447		C	T	225	DP	92	AF1	0.5	DP4	25	10	32	24
5228143	Group II intron-encoded protein LtrA	A	G	195	DP	264	AF1	0.5	DP4	69	42	78	68
5179220		C	T	188	DP	91	AF1	0.5	DP4	18	14	32	24
5179621		A	C	177	DP	115	AF1	0.5	DP4	23	19	35	37
5178584	Phage terminase large subunit (GpA)	C	T	174	DP	91	AF1	0.5	DP4	22	12	34	21
5209348		A	G	172	DP	110	AF1	0.5	DP4	25	15	33	35
5209361		T	C	167	DP	110	AF1	0.5	DP4	27	17	31	35
5209016	Putative transposase	A	G	147	DP	137	AF1	0.5	DP4	23	31	37	40
5190505	Exodeoxyribonuclease 8	C	A	144	DP	101	AF1	0.5	DP4	18	21	37	25
5209319		T	G	137	DP	105	AF1	0.5	DP4	23	14	31	31
2401		A	G	222	DP	66	AF1	1	DP4	0	1	38	27
166101		C	A	222	DP	71	AF1	1	DP4	1	0	36	34
1127475		A	C	129	DP	32	AF1	1	DP4	4	0	8	19
1146726		C	A	185	DP	47	AF1	1	DP4	0	1	22	14
1163692		T	G	138	DP	48	AF1	1	DP4	8	0	13	25
1163709		C	A	153	DP	47	AF1	1	DP4	5	0	11	26
1163724		A	T	141	DP	46	AF1	1	DP4	2	0	11	28
1243044		C	A	152	DP	36	AF1	1	DP4	0	5	20	10
1319917		G	T	158	DP	53	AF1	1	DP4	2	0	24	21
1321254		C	T	120	DP	35	AF1	1	DP4	0	3	16	9
1335221		C	T	212	DP	33	AF1	1	DP4	1	0	10	17
1398205		C	G	166	DP	46	AF1	1	DP4	1	0	18	25
1398221		G	T	160	DP	47	AF1	1	DP4	4	0	17	22
1399332		G	A	143	DP	32	AF1	1	DP4	0	3	19	7

1594705		A	C	222	DP	74	AF1	1	DP4	0	1	32	36
1637150		C	A	149	DP	49	AF1	1	DP4	0	5	25	18
1708815		G	T	180	DP	64	AF1	1	DP4	2	0	17	37
1746772		G	T	127	DP	36	AF1	1	DP4	3	0	5	22
1805361		T	C	222	DP	65	AF1	1	DP4	1	0	24	38
1890043		C	A	126	DP	28	AF1	1	DP4	0	4	12	10
2123253		C	A	197	DP	48	AF1	1	DP4	0	4	26	14
2131013		G	T	145	DP	42	AF1	1	DP4	0	0	17	21
2149313		C	T	205	DP	50	AF1	1	DP4	0	0	23	24
2168413		G	T	222	DP	38	AF1	1	DP4	0	0	17	20
2200067		C	A	141	DP	29	AF1	1	DP4	0	0	15	8
2237236		C	T	185	DP	50	AF1	1	DP4	1	0	20	28
2244632		G	T	121	DP	49	AF1	1	DP4	5	0	12	32
2341336		A	T	222	DP	55	AF1	1	DP4	0	0	30	21
2364636		G	T	191	DP	60	AF1	1	DP4	1	0	18	41
2659279		C	A	217	DP	45	AF1	1	DP4	0	1	22	21
2762030		G	T	134	DP	94	AF1	1	DP4	4	0	11	42
2898987		G	T	147	DP	62	AF1	1	DP4	6	0	10	37
2906731		T	C	201	DP	40	AF1	1	DP4	0	0	15	23
3093913		C	A	133	DP	62	AF1	1	DP4	0	6	32	22
3108870		T	G	179	DP	60	AF1	1	DP4	0	0	26	31
3263254		C	A	164	DP	51	AF1	1	DP4	0	1	28	14
3316800		T	C	125	DP	18	AF1	1	DP4	0	0	4	14
3475780		C	A	207	DP	65	AF1	1	DP4	0	1	30	19
3540274		A	T	222	DP	42	AF1	1	DP4	1	0	13	25
3658719		G	T	153	DP	74	AF1	1	DP4	5	0	22	37
3862054		G	T	161	DP	56	AF1	1	DP4	3	0	19	28
3862229		G	A	134	DP	58	AF1	1	DP4	0	3	31	14
3862373		G	A	126	DP	62	AF1	1	DP4	2	1	5	43
3886959		A	T	222	DP	55	AF1	1	DP4	0	1	24	28
4229041		T	A	175	DP	43	AF1	1	DP4	0	0	20	20
4292034		C	A	160	DP	48	AF1	1	DP4	0	3	28	8

4322426		T	G	173	DP	36	AF1	1	DP4	4	0	13	17
4690388		C	A	130	DP	62	AF1	1	DP4	0	5	30	7
4738282		G	T	148	DP	44	AF1	1	DP4	6	0	18	12
4792156		G	A	182	DP	49	AF1	1	DP4	0	0	14	29
4864165		C	T	153	DP	44	AF1	1	DP4	2	0	25	16
4864197		A	G	222	DP	44	AF1	1	DP4	0	0	23	16
4864209		T	C	207	DP	42	AF1	1	DP4	0	0	20	18
5172250		C	T	175	DP	42	AF1	1	DP4	0	1	18	16
5183360		G	T	167	DP	34	AF1	1	DP4	0	1	13	8
5248624		G	T	163	DP	56	AF1	1	DP4	6	13	18	19
UTI (M16)													
5169385		C	T	131	DP	137	AF1	0.5	DP4	22	28	21	66
5310182		T	C	132	DP	86	AF1	0.5	DP4	16	16	25	29
5299182	IncFII repA protein family	T	C	136	DP	57	AF1	0.5	DP4	4	12	17	18
306799	Bifunctional protein Fo1C	T	C	137	DP	39	AF1	0.5005	DP4	2	2	21	14
5213081		T	G	144	DP	90	AF1	0.5	DP4	13	19	22	32
5358449		C	T	145	DP	21	AF1	0.5025	DP4	3	0	11	7
1639567	Electron transport complex protein RnfC	A	C	151	DP	52	AF1	0.5	DP4	4	4	19	24
5205886		C	T	155	DP	44	AF1	0.5	DP4	10	5	17	12
5276307		C	A	177	DP	76	AF1	0.5	DP4	12	10	27	20
5358463		A	G	225	DP	21	AF1	0.5003	DP4	3	0	11	7
1639543	Electron transport complex protein RnfC	G	A	225	DP	58	AF1	0.5	DP4	7	7	18	26
1639555	Electron transport complex protein RnfC	A	G	225	DP	56	AF1	0.5	DP4	5	6	17	25
4860839	DNA primase TraC	T	A	225	DP	31	AF1	0.5	DP4	5	3	9	11
4864309		C	T	225	DP	99	AF1	0.5	DP4	18	21	28	31
5164966		A	G	225	DP	21	AF1	0.5	DP4	2	3	8	8
5299152	IncFII repA protein family	A	G	225	DP	59	AF1	0.5	DP4	8	13	17	21
5299217		G	T	225	DP	48	AF1	0.5	DP4	10	3	15	18
56301		C	A	215	DP	32	AF1	1	DP4	1	0	14	15
102855		C	A	137	DP	39	AF1	1	DP4	0	4	15	16
149869		C	A	121	DP	21	AF1	1	DP4	0	0	11	7
292232		C	A	174	DP	42	AF1	1	DP4	0	2	18	14

292252		C	A	128	DP	41	AF1	1	DP4	0	5	17	10
329509		G	T	143	DP	37	AF1	1	DP4	2	0	9	16
329520		C	A	168	DP	35	AF1	1	DP4	4	0	16	13
329556		G	T	155	DP	48	AF1	1	DP4	0	0	19	20
574633		A	T	222	DP	44	AF1	1	DP4	4	0	16	23
737884		T	C	142	DP	24	AF1	1	DP4	0	1	13	5
774704		G	T	182	DP	45	AF1	1	DP4	2	0	14	23
827063		C	T	174	DP	51	AF1	1	DP4	0	1	19	29
855138		T	C	136	DP	43	AF1	1	DP4	5	0	12	24
878552		C	A	141	DP	58	AF1	1	DP4	1	3	35	8
880582		C	A	158	DP	53	AF1	1	DP4	0	0	34	8
880599		C	T	143	DP	55	AF1	1	DP4	0	1	33	9
910863		A	G	148	DP	18	AF1	1	DP4	1	0	3	14
1105063		G	T	222	DP	67	AF1	1	DP4	0	0	37	26
1144391		C	A	121	DP	36	AF1	1	DP4	0	0	5	27
1162204		G	T	125	DP	33	AF1	1	DP4	4	0	11	16
1163712		C	A	182	DP	40	AF1	1	DP4	0	2	19	14
1183248		C	T	127	DP	26	AF1	1	DP4	0	2	8	13
1183392		C	A	166	DP	28	AF1	1	DP4	0	0	9	16
1205869		G	T	122	DP	39	AF1	1	DP4	1	0	6	19
1390439		C	T	122	DP	37	AF1	1	DP4	4	0	5	17
1390504		A	C	146	DP	42	AF1	1	DP4	8	0	11	21
1390514		A	G	125	DP	43	AF1	1	DP4	9	0	9	22
1524317		G	A	167	DP	37	AF1	1	DP4	0	2	18	13
1524331		G	T	222	DP	35	AF1	1	DP4	0	0	17	6
1624796		A	G	169	DP	49	AF1	1	DP4	1	0	20	18
1902422		C	G	197	DP	30	AF1	1	DP4	1	0	9	20
1908063		G	T	153	DP	45	AF1	1	DP4	4	0	9	16
2071822		G	T	153	DP	57	AF1	1	DP4	1	1	36	19
2074222		C	T	222	DP	54	AF1	1	DP4	1	0	19	33
2176839		G	T	133	DP	61	AF1	1	DP4	7	1	19	31
2437021		C	A	120	DP	38	AF1	1	DP4	0	4	21	11

2609198		C	A	178	DP	41	AF1	1	DP4	0	1	20	16
2609216		T	C	152	DP	42	AF1	1	DP4	0	7	19	16
2683045		C	T	166	DP	31	AF1	1	DP4	0	0	15	14
2683281		C	A	156	DP	26	AF1	1	DP4	0	0	13	9
2683411		G	T	120	DP	32	AF1	1	DP4	6	1	11	12
2730545		C	T	222	DP	46	AF1	1	DP4	0	0	16	28
2931725		C	G	126	DP	32	AF1	1	DP4	5	0	7	16
2931764		C	A	128	DP	29	AF1	1	DP4	3	0	8	16
3272804		C	A	145	DP	34	AF1	1	DP4	0	1	17	2
3339750		C	A	120	DP	40	AF1	1	DP4	5	0	4	25
3339771		T	C	136	DP	41	AF1	1	DP4	5	0	9	26
3466977		G	A	215	DP	44	AF1	1	DP4	1	0	10	18
3615777		A	T	222	DP	44	AF1	1	DP4	1	0	22	20
3671501		C	A	120	DP	36	AF1	1	DP4	0	0	22	10
3725299		G	C	153	DP	32	AF1	1	DP4	0	0	3	19
3725835		G	T	151	DP	52	AF1	1	DP4	0	0	13	25
4006737		G	A	138	DP	43	AF1	1	DP4	4	0	13	20
4053151		G	T	122	DP	36	AF1	1	DP4	5	0	6	16
4063886		T	G	220	DP	45	AF1	1	DP4	1	0	15	27
4071444		T	G	150	DP	43	AF1	1	DP4	0	1	30	11
4101403		C	A	161	DP	59	AF1	1	DP4	0	2	32	12
4515909		C	T	222	DP	41	AF1	1	DP4	1	0	13	27
4576092		C	A	141	DP	59	AF1	1	DP4	1	0	26	30
4698429		C	A	126	DP	34	AF1	1	DP4	0	3	17	8
4761126		C	A	222	DP	41	AF1	1	DP4	0	1	24	15
4774326		A	T	222	DP	43	AF1	1	DP4	1	0	17	24
4831146		G	T	157	DP	53	AF1	1	DP4	1	0	25	26
4860498		A	C	146	DP	49	AF1	1	DP4	1	6	21	20
4860591		C	A	132	DP	42	AF1	1	DP4	0	3	17	11
4860605		C	T	144	DP	42	AF1	1	DP4	0	3	16	17
5046530		T	G	138	DP	47	AF1	1	DP4	5	0	12	25
5046561		T	C	138	DP	45	AF1	1	DP4	5	0	10	26

5046610		C	A	167	DP	42	AF1	1	DP4	4	1	10	27
5047570		G	C	167	DP	30	AF1	1	DP4	0	3	17	7
5102850		A	T	166	DP	34	AF1	1	DP4	3	1	11	16
5201353		C	A	177	DP	27	AF1	1	DP4	0	1	9	16
5221464		C	A	162	DP	25	AF1	1	DP4	0	1	7	14
5279588		G	A	132	DP	25	AF1	1	DP4	0	2	6	12
5294174		T	C	222	DP	22	AF1	1	DP4	2	3	11	6
UTI (M20)													
5153575	Hypothetical protein	G	T	222	DP	56	AF1	1	DP4	4	2	24	23
3975549		G	A	225	DP	49	AF1	0.5	DP4	11	4	11	23
5137817	Integrase core domain protein	T	G	225	DP	628	AF1	0.5	DP4	142	142	172	163
5141778		G	A	225	DP	234	AF1	0.5	DP4	49	43	69	58
5147501		C	T	195	DP	187	AF1	0.5	DP4	36	41	44	63
2977259		C	T	130	DP	106	AF1	0.5	DP4	10	23	30	43
5141709		A	G	130	DP	224	AF1	0.5	DP4	46	45	69	60
208601		A	C	213	DP	113	AF1	1	DP4	0	1	51	58
536409		C	A	192	DP	107	AF1	1	DP4	0	4	50	32
549772		G	T	120	DP	75	AF1	1	DP4	4	0	23	31
695461		A	G	222	DP	83	AF1	1	DP4	4	0	21	56
736361		C	T	150	DP	94	AF1	1	DP4	1	0	40	52
802496		T	A	222	DP	85	AF1	1	DP4	1	1	39	44
950710		C	A	150	DP	85	AF1	1	DP4	8	0	32	38
1591422		G	C	120	DP	47	AF1	1	DP4	0	3	20	8
1591432		C	A	206	DP	46	AF1	1	DP4	0	2	19	17
1981672		G	A	222	DP	78	AF1	1	DP4	7	1	35	27
1981690		G	A	203	DP	74	AF1	1	DP4	1	2	33	29
1981720		G	T	222	DP	70	AF1	1	DP4	0	0	30	28
2011562		A	T	214	DP	95	AF1	1	DP4	2	0	43	46
2202802		G	T	218	DP	90	AF1	1	DP4	0	1	47	42
2521737		G	T	222	DP	78	AF1	1	DP4	0	0	39	38
2656397		G	T	129	DP	92	AF1	1	DP4	3	0	6	55
2729487		C	T	147	DP	78	AF1	1	DP4	0	4	37	12

2814242		G	A	147	DP	55	AF1	1	DP4	1	0	22	25
2928590		C	G	209	DP	47	AF1	1	DP4	0	0	19	24
3070500		T	C	222	DP	116	AF1	1	DP4	0	1	50	64
3103445		C	A	135	DP	74	AF1	1	DP4	2	6	30	20
3232009		G	T	123	DP	78	AF1	1	DP4	4	0	24	42
3232033		G	T	169	DP	79	AF1	1	DP4	1	0	25	42
3403781		A	G	174	DP	78	AF1	1	DP4	0	1	34	37
3722379		G	T	151	DP	84	AF1	1	DP4	6	0	23	37
3722450		G	T	155	DP	88	AF1	1	DP4	2	0	30	42
3944877		G	T	184	DP	41	AF1	1	DP4	1	0	10	24
4104601		G	C	215	DP	112	AF1	1	DP4	0	1	53	58
4262242		G	T	126	DP	44	AF1	1	DP4	1	0	3	31
4562485		G	T	214	DP	71	AF1	1	DP4	0	1	34	34
4570129		G	T	126	DP	67	AF1	1	DP4	3	0	10	35
4571402		G	T	122	DP	50	AF1	1	DP4	1	0	9	23
4728532		C	T	139	DP	45	AF1	1	DP4	0	0	26	7
4913927		G	T	144	DP	78	AF1	1	DP4	9	0	21	44
5150943		T	C	222	DP	66	AF1	1	DP4	3	2	30	24
5153362		A	G	133	DP	29	AF1	1	DP4	1	0	17	8
5153509		A	G	201	DP	56	AF1	1	DP4	0	0	31	25
5153536		T	C	222	DP	53	AF1	1	DP4	2	0	22	24
5153575		G	T	222	DP	56	AF1	1	DP4	4	2	24	23
UTI (F19)													
5193123		A	C	121	DP	145	AF1	0.5	DP4	41	18	54	30
764134		G	T	126	DP	46	AF1	0.5	DP4	14	0	12	16
4643865	Anti restriction protein K1cA	A	G	129	DP	92	AF1	0.50	DP4	13	0	32	40
5192970		C	A	132	DP	96	AF1	0.5	DP4	19	14	42	19
4644344		A	G	137	DP	154	AF1	0.5	DP4	29	32	37	52
3788197		C	T	153	DP	77	AF1	0.5	DP4	15	14	23	25
1990634	Small toxic polypeptide LdrD	C	T	156	DP	111	AF1	0.5	DP4	12	20	24	54
1990151	Small toxic polypeptide LdrD	C	T	157	DP	86	AF1	0.5	DP4	14	7	32	30
3677065		T	C	158	DP	67	AF1	0.5	DP4	9	13	25	19

5299014		A	G	163	DP	55	AF1	0.5	DP4	11	6	26	10
3677147	Putative autotransporter precursor	A	G	168	DP	72	AF1	0.5	DP4	5	13	19	27
3788210		T	C	168	DP	81	AF1	0.5	DP4	12	13	31	25
5032962	Hypothetical protein	G	A	178	DP	135	AF1	0.5	DP4	26	27	49	27
4644272		A	G	184	DP	149	AF1	0.5	DP4	31	27	42	42
3676966		A	G	225	DP	106	AF1	0.5	DP4	11	31	33	29
4643880	Antirestriction protein K1cA	A	G	225	DP	92	AF1	0.5	DP4	15	1	35	40
5208312		C	T	225	DP	141	AF1	0.5	DP4	35	21	46	36
5209699		C	T	225	DP	139	AF1	0.5	DP4	24	32	33	46
5252249		C	T	225	DP	87	AF1	0.5	DP4	17	16	28	26
5275233		C	A	225	DP	74	AF1	0.5	DP4	22	5	29	18
5291706	Is1 transposase	C	A	225	DP	203	AF1	0.5	DP4	28	18	93	57
5302106	Transposase IS66 family protein	C	A	225	DP	84	AF1	0.5	DP4	20	7	37	8
10901		G	A	158	DP	40	AF1	1	DP4	1	0	21	18
387967		C	T	120	DP	38	AF1	1	DP4	0	2	23	4
545937		A	C	222	DP	68	AF1	1	DP4	0	1	25	41
656518		C	A	178	DP	60	AF1	1	DP4	0	0	27	13
656563		C	A	170	DP	69	AF1	1	DP4	0	8	37	13
688607		G	A	122	DP	51	AF1	1	DP4	4	0	8	25
760837		G	C	162	DP	52	AF1	1	DP4	0	1	38	13
981051		A	C	143	DP	61	AF1	1	DP4	0	7	28	24
981132		C	A	167	DP	57	AF1	1	DP4	0	5	28	16
1453004		C	T	175	DP	55	AF1	1	DP4	0	4	27	13
1497550		G	T	212	DP	50	AF1	1	DP4	1	0	25	15
1497573		G	T	169	DP	37	AF1	1	DP4	1	0	20	8
1750960		C	G	122	DP	34	AF1	1	DP4	0	7	20	7
1808580		T	A	139	DP	49	AF1	1	DP4	0	8	28	10
1855114		G	T	134	DP	43	AF1	1	DP4	3	0	8	25
1910019		T	G	184	DP	45	AF1	1	DP4	5	0	19	21
1976649		C	A	177	DP	50	AF1	1	DP4	1	0	18	25
2394133		G	T	154	DP	44	AF1	1	DP4	2	0	5	21
2514162		T	A	146	DP	35	AF1	1	DP4	0	4	17	13

2514187		G	C	131	DP	34	AF1	1	DP4	0	5	15	11
2514198		C	A	136	DP	37	AF1	1	DP4	0	2	18	8
2636723		G	A	222	DP	58	AF1	1	DP4	0	0	28	14
2649562		C	A	145	DP	67	AF1	1	DP4	0	5	28	16
2905967		G	T	182	DP	36	AF1	1	DP4	1	0	5	24
2909962		C	A	170	DP	42	AF1	1	DP4	0	0	29	3
3056375		T	A	126	DP	60	AF1	1	DP4	1	0	15	44
3168854		G	T	217	DP	41	AF1	1	DP4	1	0	15	23
3169638		C	A	153	DP	65	AF1	1	DP4	8	0	18	39
3433420		T	C	202	DP	33	AF1	1	DP4	1	0	11	21
3482020		C	G	154	DP	53	AF1	1	DP4	0	1	24	27
3548720		G	A	165	DP	79	AF1	1	DP4	0	1	46	32
3553124		G	A	125	DP	47	AF1	1	DP4	6	0	20	18
3635182		T	A	189	DP	64	AF1	1	DP4	0	0	34	27
3677463		C	T	222	DP	75	AF1	1	DP4	2	2	27	37
3688282		G	T	195	DP	69	AF1	1	DP4	0	0	33	33
3739146		G	T	130	DP	44	AF1	1	DP4	5	0	19	18
3786374		A	C	222	DP	58	AF1	1	DP4	0	1	29	26
4051834		A	T	146	DP	65	AF1	1	DP4	1	0	28	35
4132465		C	A	155	DP	37	AF1	1	DP4	0	0	25	6
4167131		C	A	166	DP	55	AF1	1	DP4	0	0	22	12
4202117		T	A	186	DP	60	AF1	1	DP4	0	7	34	18
4397400		C	A	185	DP	57	AF1	1	DP4	0	1	29	25
4411970		G	A	222	DP	40	AF1	1	DP4	0	0	14	23
4435580		A	G	137	DP	47	AF1	1	DP4	0	7	17	23
4565513		G	A	124	DP	31	AF1	1	DP4	5	1	8	15
4583393		C	A	128	DP	35	AF1	1	DP4	2	0	5	23
4618479		G	T	153	DP	45	AF1	1	DP4	0	1	30	6
4682121		C	T	130	DP	33	AF1	1	DP4	0	1	20	5
4682387		A	C	129	DP	34	AF1	1	DP4	0	2	25	7
4682787		A	G	204	DP	44	AF1	1	DP4	0	1	23	19
4703403		G	T	135	DP	66	AF1	1	DP4	3	0	14	39

4906586		G	T	160	DP	46	AF1	1	DP4	3	0	7	30
5003758		A	T	125	DP	50	AF1	1	DP4	0	9	25	15
5264633		A	C	159	DP	31	AF1	1	DP4	1	0	14	16
UTI (F12)													
5440371	ECF RNA plymerase sigma factor SigW	G	A	169	DP	119	AF1	0.5	DP4	34	11	46	24
5500129	Manganese transport system membrane protein MntB	T	C	165	DP	175	AF1	0.5	DP4	30	33	54	58
4242957	Small toxic polypeptide LdrD	G	A	162	DP	118	AF1	0.5	DP4	29	11	44	34
5500474	Manganese transport system membrane protein MntB	A	G	160	DP	152	AF1	0.5	DP4	27	29	41	53
7034249		C	T	160	DP	69	AF1	0.5	DP4	12	10	26	21
6134236		A	G	142	DP	81	AF1	0.5	DP4	15	13	30	18
1169066	Carboxylate-amine ligase YbdK	C	A	135	DP	51	AF1	0.6243	DP4	0	6	22	13
939450	Hypothetical protein	C	A	133	DP	44	AF1	0.5002	DP4	3	9	12	19
3995602	Coenzyme A biosynthesis bifunctional protein CoaBC	G	T	132	DP	31	AF1	0.5013	DP4	7	2	9	12
5499772		A	G	121	DP	131	AF1	0.5	DP4	32	14	61	23
44074		C	A	159	DP	56	AF1	1	DP4	0	5	28	13
47401		G	A	222	DP	53	AF1	1	DP4	0	1	25	25
227470		C	A	220	DP	47	AF1	1	DP4	0	1	28	12
261700		A	C	133	DP	42	AF1	1	DP4	2	0	13	23
261840		C	T	129	DP	37	AF1	1	DP4	1	0	15	16
286901		G	A	222	DP	54	AF1	1	DP4	0	1	34	18
289201		A	G	222	DP	76	AF1	1	DP4	0	1	32	33
341216		C	A	125	DP	75	AF1	1	DP4	0	1	34	22
519501		G	T	212	DP	63	AF1	1	DP4	0	1	37	25
541902		T	C	169	DP	65	AF1	1	DP4	8	0	17	40
546360		C	A	144	DP	36	AF1	1	DP4	0	0	21	5
623139		G	T	144	DP	69	AF1	1	DP4	3	0	9	34
636786		G	T	154	DP	68	AF1	1	DP4	9	0	17	35
740001		G	A	159	DP	87	AF1	1	DP4	0	0	55	29
1044723		G	T	138	DP	53	AF1	1	DP4	5	0	15	24
1097707		G	A	155	DP	91	AF1	1	DP4	0	1	46	43
1212795		A	C	146	DP	55	AF1	1	DP4	2	0	12	34
1228382		C	A	140	DP	30	AF1	1	DP4	0	4	16	10

1335307		C	G	140	DP	25	AF1	1	DP4	1	0	16	4
1539188		C	A	172	DP	67	AF1	1	DP4	0	6	31	26
1772067		C	A	121	DP	47	AF1	1	DP4	1	3	20	19
1866971		A	C	222	DP	44	AF1	1	DP4	1	1	12	30
1894550		G	T	133	DP	38	AF1	1	DP4	0	5	14	13
1917634		C	A	170	DP	40	AF1	1	DP4	0	2	24	8
1917707		G	T	222	DP	42	AF1	1	DP4	0	0	20	11
2067265		G	T	167	DP	42	AF1	1	DP4	4	0	7	22
2075971		G	A	222	DP	56	AF1	1	DP4	0	1	25	26
2076071		G	A	164	DP	50	AF1	1	DP4	0	2	17	29
2077405		G	T	160	DP	58	AF1	1	DP4	6	1	21	27
2379119		G	T	164	DP	38	AF1	1	DP4	0	0	11	21
2391064		G	T	197	DP	56	AF1	1	DP4	2	0	19	27
2391171		G	T	150	DP	59	AF1	1	DP4	4	0	18	33
2511499		G	T	171	DP	76	AF1	1	DP4	1	0	47	27
2653801		C	A	168	DP	53	AF1	1	DP4	1	0	24	14
2918431		T	G	219	DP	58	AF1	1	DP4	1	0	25	30
2932918		C	A	144	DP	57	AF1	1	DP4	0	1	30	17
2932931		T	G	222	DP	58	AF1	1	DP4	0	1	31	25
3132001		G	T	155	DP	38	AF1	1	DP4	2	0	11	21
3132323		G	T	222	DP	57	AF1	1	DP4	0	1	28	26
3263301		G	T	126	DP	30	AF1	1	DP4	0	0	3	23
3263504		C	A	222	DP	53	AF1	1	DP4	0	1	25	22
3352646		G	C	120	DP	25	AF1	1	DP4	2	0	7	14
3675308		C	A	145	DP	55	AF1	1	DP4	0	5	27	15
3743067		G	T	151	DP	84	AF1	1	DP4	2	1	27	41
3808171		C	T	127	DP	34	AF1	1	DP4	0	3	9	16
3817212		C	A	132	DP	58	AF1	1	DP4	0	3	35	11
3855317		T	G	222	DP	71	AF1	1	DP4	1	0	17	52
4033605		G	T	198	DP	61	AF1	1	DP4	0	1	29	17
4161756		G	T	196	DP	63	AF1	1	DP4	0	0	18	36
4183047		G	A	154	DP	54	AF1	1	DP4	0	7	22	21

4516543		**G**	**A**	**164**	**DP**	**49**	**AF1**	**1**	**DP4**	**7**	**0**	**15**	**26**
4718041		**C**	**A**	**139**	**DP**	**85**	**AF1**	**1**	**DP4**	**0**	**2**	**46**	**21**
4983183		**C**	**A**	**222**	**DP**	**63**	**AF1**	**1**	**DP4**	**0**	**1**	**35**	**27**
5182399		**A**	**C**	**143**	**DP**	**26**	**AF1**	**1**	**DP4**	**1**	**0**	**9**	**14**
5258556		**G**	**C**	**130**	**DP**	**15**	**AF1**	**1**	**DP4**	**0**	**1**	**9**	**5**

References

Adams-Sapper, S., Diep, B.A., Perdreau-Remington, F. and Riley, L.W., 2013. Clonal composition and community clustering of drug susceptible and drug resistance *Escherichia coli* isolated from bloodstream infections. *Antimicrobial Agents and Chemotherapy*. 57(1). 490-497.

Adamus-Białek, W., Kubiak, A. and Czerwonka, G., 2015. Analysis of uropathogenic *Escherichia coli* biofilm formation under different growth conditions. *ACTABP*. 62(4). 765-771.

Agarwal, J., Srivastava, S. and Sing, M., 2012. Pathogenomics of uropathogenic *Escherichia coli*. *Indian Journal of Medical Microbiology*. 30(2). 141-149.

Bahrani-Mougeot, F.K., Buckles, E.L., Lockatell, C.V., Hebel, J.R., Johnson, D.E., Tang, C.M. and Donnenberg, M.S., 2002. Type 1 fimbriae and extracellular polysaccharides are preeminent uropathogenic *Escherichia coli* virulence determinants in the murine urinary tract. *Molecular Microbiology.* 45(4). 1079-1093.

Bankevich, A., Nurk, S., Antipov, D., Guervich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G., Alekseyev, M.A. and Pevzner, P.A., 2012. SPAdes: a new genome assembly algorithm and its application to single cell sequencing. *Journal of Computational Biology.* 19(5). 455-477.

Basic Local Alignment Search Tool., 2016. Blastx [online]. Available at: http://blast.ncbi.nlm.nih.gov

Bertelli, C. and Greub, G., 2013. Rapid bacterial genome sequencing: methods and applications in clinical microbiology. *Clinical Microbiology and Infection.* 19(9). 803-813.

Bevan, E.R., Jones, A.M. and Hawkey, P.M., 2017. Global epidemiology of CTX-M β-lactamases: temporal and geographical shifts in genotype. *Journal of Antimicrobial Chemotherapy.* 72. 2145-2155.

Bien, J., Sokolova, O. and Bozko, P., 2012. Role of uropathogenic Escherichia coli virulence factors in development of urinary tract infection and kidney damage. *International Journal of Nephrology*. Doi:10.1155.

Blango, M. and Mulvey, M., 2010. Persistence of Uropathogenic *Escherichia coli* in the Face of Multiple Antibiotics. *Antimicrobial Agents and Chemotherapy.* 54. 1855–1863.

Boyd, E.F. and Hartl, D.L., 1998. Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution. *Journal of Bacteriology*. 180(5). 1159-1165.

Branda, S.S., Vik, S., Friedman, L. and Kolter, R., 2005. Biofilms: the matrix revisited. *Trends Microbiology*. 13. 20-26.

Brumbaugh, A.R. and Mobeley, H.L.T., 2012. Preventing urinary tract infection: progress toward an effective *Escherichia coli* vaccine. *Expert Reviews.* 11(6). 663-676.

Brzuszkiewick, E., Brüggemann, H., Liesegang, H., Emmerth, M., Ölschläger, T., Nagy, G., Albermann, K., Wanger, C., Buchrieser, C., Emὸdy, L., Gottschalk, G., Kacker, J. and Dobrindt, U., 2006. How to become uropathogen: comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. *Proceedings of the National Academy of Sciences of the United states of America*. 103(34). 12879-12884.

BSAC., 2012. BSAC methods for antimicrobial susceptibility testing. Web Site for the British Society for Antimicrobial Chemotherapy Web.

BSAC., 2013. BSAC methods for antimicrobial susceptibility testing. Web Site for the British Society for Antimicrobial Chemotherapy Web [online]. Available at: http://bsac.org.uk/wp-content/uploads/2012/02/Version-12-Apr-2013_final1.pdf

Campbell, I., 2007. Chi-squared and Fisher–Irwin tests of two-by-two tables with small sample recommendations. *Statistics in Medicine.* 26(19). 3661-3675.

Carver, T., Thomson, N., Bleasby, A., Berriman, M. and Parkhill, J., 2009. DNAPlotter: circular and linear interactive genome visulasation. *Bioinformatics.* 25 (1). 119-120.

Cegelski, L., Pinkner, J., Hammer, N., Cusumano, C.K., Chorell, E., Aberg, V., Walker, J.N., Seed, P.C., Almqvist, F., Chapman, M.R. and Hultgren, S., 2009. Small molecules inhibitors target *Escherichia coli* amyloid biogenesis and biofilm formation. *Nature Chemical Biology.* 5. 913-919.

Center for Genomics Epidemiology., 2011. CGE [online]. Available at: https://cge.cbs.dtu.dk/services/

Cespedes, C., Saïd-salim, B., Miller, M., Lo, S., Kreiswirth, B.N., Gordon, R.J., Vavagiakis, P., Klein, R.S. and Lowy, F.D., 2005. The clonality of *Staphylococcus aureus* nasal carriage. *The Journal of Infectious Diseases.* 191. 444-452.

Chaudhuri, R.R. and Handerson, I.R., 2012. The evolution of the Escherichia coli phylogeny. *Infection, Genetics and Evolution*. 12. 214-226.

Choong, S. and Whitfield, H., 2000. Biofilms and their role in infections in urology. *BJUI International*. 86. 935-941.

Chung, J.C.S., Becq, J., Fraser, L., Schulz-Trieglaff, O., Bond, N.J., Foweraker, J., Bruce, K.D., Smith, G.P. and Welch, M., 2012. Genomic variation among contemporary *Pseudomonas aeruginosa* isolates from chronically infected cystic fibrosis patients. *Journal of Bacteriology.* 194(18). 4857-4866.

Clark, G., Paszkiewicz, K., Hale, J., Weston, J., Constantinidou, C., Penn, C., Achtman, M. and McNally, A., 2012. Genomic analysis uncovers a phenotypically diverse but genetically homogenous *Escherichia coli* ST131 clone circulating in unrelated urinary tract infections. *Journal of Antimicrobial Chemotherapy*. 67. 868-877.

CLSI., 2012. Clinical and Laboratory Standard Institute. Performance standard for antimicrobial disk susceptibility tests. Clinical and Laboratory Standard Institute.

Connell, I., Agace, W., Klemm, P., Schembri, M., Marild, S. and Svanborg, C., 1996. Type 1 fimbrial expression enhances *Escherichia coli* virulence for the urinary tract. *Proceedings of the National Academy of Sciences of the United States of America*. 93. 9827–9832.

Costerton, J.W., 1999. Introduction to biofilm. *International Journal of Antimicrobial Agents*. 11. 217-221.

Croxall, G., Hale, J., Weston, V., Manning, G., Cheetham, P., Achtman, M. and McNally, A., 2011. Molecular Epidemiology of Extraintestinal pathogenic *Escherichia coli* isolates from a regional cohort of elderly patients highlights the prevalence of ST131 strains with increased antimicrobial resistance in both community and hospital care settings. *Journal of Antimicrobial Chemotherapy*. 66. 2501-2508.

Croxen, M.A., Law, R.J., Scholz, R., Keeney, K.M., Woldarska, M. and Finlay, B.B., 2013. Recent advance in understanding Enteric pathogenic *Escherichia coli*. *Clinical Microbiology Reviews*. 26(4). 822-880.

Darch, S.E., McNally, A., Harrison, F., Corander, J., Barr, H.L., Paszkiewicz, K., Holden, S., Fogarty, A., Crusz, S.A. and Diggle, S.P., 2015. Recombination is a key driver of genomic and phenotypic diversity in a *Pseudomonas aeruginosa* population during cystic fibrosis infection. *Scientific Reports*. 5. 7649. DOI: 10.1038.

Didelot, X., Bowden, R., Wilson, D.J., Peto, T.E.A. and Crook, D.W., 2012. Transforming clinical microbiology with bacterial genome sequencing. *Reviews.* 13. 601-612.

Didelot, X., Walker, S., Peto, T.E., Crook, D.W. and Wilson, D., 2016. Within host evolution of bacterial pathogens. *Nature Reviews Microbiology.* 14(3). 150-162.

Dobrindt, U., 2005. Pathogenomics of *Escherichia coli*. *International Journal of Medical Microbiology*. 295. 357-371.

Dobrindt, U., Agerer, F., Michaelis, K., Janka, A., Buchrieser, C., Samuelson, M., Svanborg, C., Gottschalk, G., Karch, H. and Hacker, J., 2003. Analysis of genome plasticity in pathogenic and commensal *Escherichia coli* isolates by use of DNA arrays. *Journal of Bacteriology.* 185(6). 1831-1840.

Ejrnæs, K., 2011. Bacterial characteristics of importance for recurrent urinary tract infections caused by *Escherichia coli. Danish Medical Bulletin*. 58 (4). B4187.

El-Halfawy, O.M. and Valvano, M.A., 2015. Antimicrobial heteroresistance: an emerging field in need of clarity. *Clinical Microbiology Reviews*. 28.191–207. DOI:10.1128.

Erdem, A.L., Avelino, F., Xicohtencatl-Cortes, J. and Giron, J.A., 2007. Host protein binding and adhesive properties of H6 and H7 flagella of attaching and effacing *Escherichia coli*. *Journal of Bacteriology*. 189. 7426-7435.

Evolution and Genomics, 2011. bcfTools (online). Available at: http://evomics.org/resources/software/varaition/bcftools/

Feliziani, S., Marving, R.L., Luján, A.M., Moyano, A.J., Di Rienzo, J.A., Johansen, H.K., Molin, S. and Smania, A.M., 2014. Coexistence and within host evolution of diversified lineage of hypermutable *Pseudomonas aeruginosa* in long-term cystic fibrosis infections. *PLoS Genetics.* 10(10). DOI: 10.1371.

Flores-Mireles, A.L., Walker, J.N., Caparon, M. and Hultgren, S.J., 2015. Urinary tract infections: epidemiology, mechanism of infection and treatment options. *Nature Reviews Microbiology*. 13. 269-284.

Foweraker, J.E., Laughton, C.R., Brown, D.F.J. and Bilton, D., 2005. Phenotypic variability of *Pseudomonas aeruginosa* in sputa from patients with acute infective exacerbation of cystic fibrosis and its impact on the validity of antimicrobial susceptibility testing. *Journal of Antimicrobial Chemotherapy*. 55(6). 921-927.

Foxman, B., 1990. Recurring urinary tract infection: incidence and risk factors. *Journal of Clinical Epidemiology*. 80(3). 331-333.

Foxman, B., 2003. Epidemiology of urinary tract infection: incidence, morbidity and economic costs. *Disease-a-Month*. 49 (2). 53-70.

Foxman, B., 2010. The epidemiology of urinary tract infection. *Nature Reviews Urology*. 76(12). 653-660.

Galdbart, J., Morvan, A., Desplaces, N. and Solh, N.E., 1999. Phenotypic and genomic variation among *Staphylococcus epidermidis* strains infecting joint prostheses. *Journal of Clinical Microbiology.* 37(5). 1306-1312.

Gibreel, T.M., Dodgson, A.R., Cheesbrough, J., Fox, A.J., Bolton, F.J. and Upton, M., 2012. Population structure, virulence potential and antibiotic susceptibility of uropathogenic *Escherichia coli* from Northwest England. *Journal of Antimicrobial Chemotherapy*. 67. 346-356.

Goldhar, J., 1996. Nonfimbrial adhesions of *Escherichia coli. Advances in Experimental Medicine and Biology*. 408. 63-72.

Golubchik, T., Batty, E.M., Miller, R.R., Farr, H., Young, B.C., Larner-Svensson, H., Fung, R., Godwin, H., Knox, K., Votintseva, A., Everitt, R.G., Street, T., Cule, M., Ip, C.L.C.,

Didelot, X., Peto, T.E.A., Harding, R.M., Wilson, D.J., Crook, D.W. and Bowden, R., 2013. Within host evolution of *Staphylococcus aureus* during asymptomatic carriage. *PLoS ONE.* 8(5). DOI:10.1371.

Graham, J. C. and Galloway, A., 2001. ACP Best Practice No 167: the laboratory diagnosis of urinary tract infection. *Journal of Clinical Pathology*. 54. 911–919.

Gupta, K., Hooton, T.M., Naber, K.G., Wullt, B., Colgan, R., Miller, L.G., Moran, G.J., Nicolle, L.E., Raz, R., Schaeffer, A.J. and Sope, D.A., 2011. International clinical practice guidelines for the treatment of acute uncomplicated cystitis and pyelonephritis in women: A 2010 update by the Infectious Diseases Society of America and the European Society for Microbiology and Infectious Diseases. *Clinical Infectious Diseases*. 52. E103–E120.

Gurevich, A., Saveliev, V., Vyhhi, N. and Tesler, G., 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics.* 29(8). 1072-1075.

Hacker, J., 2002. Urinary tract infection: from basic science to clinical application. *Advance in Experimental Medicine and Biology*. 485. 1-8.

Han, C., Shiga, Y., Tobe, T., Sasakawa, C. and Ohtsubo, E., 2001. Structural and functional characterisation of IS679 and IS66 family elements. *Journal of Bacteriology*. 183(14). 4296-4304.

Hannan, T.J., Totsika, M., Mansfield, K., Moore, K.H., Schembri, M.A. and Hultgren, S.J., 2012. Host pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic *Escherichia coli* bladder infection. *FEMS Microbiology Reviews*. 36. 616-648.

Harris, S.R., Cartwright, E.J.P., TörökK, M.E., Holden, M.T.G., Brown, N.M., Ogilvy-Stuart, A.L., Ellington, M.J., Quail, M.A., Bentley, S.D., Parkhill, J. and Peaccock, S.J., 2013. Whole

genome sequencing for analysis of an outbreak of methicillin resistant Staphylococcus aureus: a descriptive study. *The Lancet Infection Diseases.* 13. 130-136.

Harris, S.R., Feil, E.J., Holden, M.T.G., Quail, M.A., Nickerson, E.K., Chantratita, N., Gardete, S., Tavares, A., Day, N., Lindsay, J.A., Edgeworth, J.D., Lencastre, H.D., Parkhill, J., Peacock, S.J. and Bentley, D., 2010. Evolution of MRSA during hospital transmission and intercontinental spread. *Science.* 327. 469-474.

Herrmann, B. and Burman, L.G., 1985. Pathogenesis of *Escherichia coli* cystitis and pyelonephritis: apparent lack of significance of bacterial motility and chemotaxis toward human urine. *Infection*. 13(1). 4-7.

Hilbert, DW., 2011. Uropathogenic *Escherichia coli*: the pre-eminent urinary tract infection pathogen. *Nova Science Puplishers, Inc*. 1-81. ISBN 978-1-61122-859-5.

Hooton, T.M., 2012. Uncomplicated urinary tract infection. *The New England Journal of Medicine*. 366(11). 1028-1037.

Hussain, A., Ewers, C., Nandanwar, N., Guenther, S., Jadhav, S., Wieler, L.H. and Ahmed, N., 2012. Multiresistant uropathogenic *Escherichia coli* from a region in India where urinary tract infections are endemic: genotypic and phenotypic characteristics of sequence type 131 isolates of the CTX-M-15 extended-spectrum-beta-lactamase-producing lineage. *Antimicrobial Agents Chemotherapy.* 56(12). 6358-6365.

Israel, D.A., Salama, N., Krishna, U., Rieger, U.M., Atherton, J.C., Falkow, S. and Peek, R.M., 2001. *Helicobacter pylori* genetic diversity within the gastric niches of a single human host. *Proceedings of the National Academy of Sciences of the United States of America.* 98(25). 14625-14630.

Jacobsen, S.M., Stickler, D.J., Mobley, H.L. and Shirtliff, M.E., 2008. Complicated catheter-associated urinary tract infections due to *Escherichia coli* and *Protus mirabilis*. *Clinical Microbiology Reviews*. 21(1). 26-59.

Jenkins, S. and Schuetz, A., 2012. Current Concepts in Laboratory Testing to Guide Antimicrobial Therapy. *Mayo for Medical Education and Research*. 87(3). 290-308.

Johnson, J.R., 1991. Virulence factors in *Escherichia coli* urinary tract infection. *Clinical Microbiology Reviews.* 4(1). 80-128.

Johnson, J.R. and Russo, T.A., 2005. Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. *International Journal of Medical Microbiology*. 295. 383-404.

Johnson, J.R., Delavari, P., Kuskowski, M. And Stell, A.L., 2001. Phylogenetic Distribution of Extraintestinal Virulence-Associated Traits in *Escherichia coli. The Journal of Infectious Diseases*. 183(1). 78-88.

Johnson, J.R., Johnston, B., Clabots, C., Kuskowski, M.A. and Castanheira, M., 2010. *Escherichia coli* sequence type ST131 as the major cause of serious multidrug resistance *E.coli* infection in the United States. *Clinical Infectious diseases*. 51(3). 286-294.

Johnson, R., Tiu, F. and Stamm, W., 1995. Direct Antimicrobial Susceptibility Testing for Acute Urinary Tract Infections in Women. *American Society for Microbiology*. 33(9). 2316-2323.

Johnson, R.J. and Stell, A.L., 2000. Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *The Journal of Infectious Diseases.* 181. 261-272.

Jones, C.H., Pinkner, J.S., Roth, R., Heuser, J., Nicholes, A.V., Abraham, S.N. and Hulgren, S.J., 1995. FimH adhesion of type 1 pili is assembled into a fibrillar tip structure in Enterobacteriaceae. *Proceeding of the National Academy of Science of the United States of America.* 92(6). 2081-2085.

Justice, S., Hung, C., Theriot, J.A., Fletcher, D.A., Anderson, G.G., Footer, M.J. and Hultgren, S.J., 2004. Differentiation and development pathways of uropathogenic *Escherichia coli* in urinary tract pathogenesis. *Proceedings of the National Academy of Sciences of the United States of America*. 101(5). 1333-1338.

Kallonen, T., Brodrick, H.J., Harris, S.R., Corander, J., Brown, N.M., Martin, V., Peacock, S.J. and Parkhill, J., 2017. Systemic longitudinal survey of invasive *Escherichia coli* in England demonstrates a stable structure only transiently disturbed by the emergence of ST131. *Genome Research.* 27. 1437-1449.

Kao, C., Lee, A., Huang, A., Song, P., Yang, Y., Sheu, S., Chang, W., Sheu, B. and Wu, J., 2014a. Heteroresistance of *Helicobacter pylori* from the same patient prior to antibiotic treatment. *Infection, Genetics and Evolution*. 23. 196-202.

Kao, C.Y., Lin, W.H., Tseng, C.C., Wu, A.B., Wang, M.C. and Wu, J.J., 2014b. The complex interplay among bacterial motility and virulence factors in different *Escherichia coli* infections. *European Journal of Clinical Microbiology and Infectious Diseases*. 33(12). 2157-2162.

Kaper, J.B., Nataro, J.P. and Mobley, H.L., 2004. Pathogenic *Escherichia coli*. *Nature Reviews*. 2. 123-140.

Kennemann, L., Didelot, X., Aebischer, T., Kuhn, S., Drescher, B., Droege, M., Reinnhardt, R., Meyer, T.F., Josenhans, C., Falush, D. and Suerbaum, S., 2011. *Helicobacter pylori*

genome evolution during human infection. *Proceedings of the National Academy of Sciences of the United States of America.* 18(12). 5033-5038.

Kivi, M., Rodin, S., Kupershmidt, I., Lundin, A., Tindberg, Y., Granström, M. and Engstrand, L., 2007. *Helicobacter pylori* genome variability in a framework of familial transmission. *BMC Microbiology.* 7(54). DOI: 10.1186.

Korotkova, N., Yarova-Yarovaya, Y., Tchesnokova, V., Yazvenko, N. and Carl, M.A., 2008. *Escherichia coli* DraE adhesion associated bacterial internalization by epithelial cells is promoted independently by decay accelerating factor and carcinoembryonic antigen related cell adhesion molecule binding and does not require the DraD invasion. *Infection and Immunity*. (76). 3869–3880.

Kucheria, R., Dasgupta, P., Sacks, S. H., Khan, M. S. and Sheerin, N. S., 2005. Urinary tract infections: new insights into a common problem. *Postgraduate Medical Journal*. 81(592). 83-86.

Kumar, S., Dave, A., Wolf, B. and Lerma, E.V., 2015. Urinary tract infections. *Disease-a-Month*. 61. 45-59.

Laing, C., Buchanan, C., Taboada, E.N., Zhang, Y., Kropinski, A., Villegas, A., Thomas, J.E. and Gannon, V.P.J., 2010. Pan-Genome sequence analysis using Panseq: an online toll for the rapid analysis of core and accessory genomic regions. *Bioinformatics.* 11(461). DOI: 10.1186.

Lane, M., Lockatell, V., Monterosso, G., Lamphier, D., Weinert, J., Hebel, J., Johnson, D. and Mobely, H., 2005. Role of motility in the colonisation of Uropathogenic *Escherichia coli* in the urinary tract. *American Society for Microbiology*. 73(11). 7644-7656.

Lane, M.C., Alteri, C.J., Smith, S.N. and Mobley, H.L., 2007. Expression of flagella is coincident with uropathogenic *Escherichia coli* ascension to the upper urinary tract.

Proceedings of the National Academy of Sciences of the United States of America. 104(42). 16669-16674.

Lau, S.H., Reddy, S., Cheesbrough, J., Bolton, F.J., Willshaw, G., Cheasty, T., Fox, A.J. and Upton, M., 2008. Major Uropathogenic *Escherichia coli* strain isolated in the northwest of England identified by multilocus sequence typing. *Journal of Clinical Microbiology*. 46(3). 1076-1080.

Leffler, H. and Svanborg-Eden, C., 1980. Chemical identification of a glycosphingolipid receptor for *Escherichia coli* attaching to human urinary tract epithelial cells and agglutinating human erythrocytes. *FEMS Microbiology Letters*. 8. 127–134.

Leflon-Guibout, V., Blanco, J., Amaqdouf, K., Mora, A., Guize, L. and Nicolas-Chanoine, M., 2008. Absence of CTX-M enzymes but high prevalence of clones including ST131 among faecal *Escherichia coli* isolates from healthy subjects living in the area of Paris, France. *Journal of Clinical Microbiology*. 46(12). 3900-3905.

Levert, M., Zamfir, O., Clermont, O., Bouvet, O., Lespinats, S., Hipeaux, M.C., Branger, C., Picard, B., Saint-Ruf, C. and Norel, F., 2010. Molecular and evolutionary bases of within-patient genotypic and phenotypic diversity in *Escherichia coli* extraintestinal infections. *PLoS Pathogens.* 6 (9). E1001125.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G. and Burbin, R., 2009. The sequence alignment/map format and SAMtools. *Bioinformatics.* 25(16). 2078-2079.

Li, H. and Durbin, R., 2009. Fast and accurate short read alignment with Burrows-wheeler transform. *Bioinformatics.* 25. 1754-1760.

Lieberman, T.D., Flett, K.B., Yelin, I., Martin, T.R., McAdam, A.J., Preiebe, G. and Kishony, R., 2014. Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressure. *Nature Genetics*. 46(1). 82-87.

Lieberman, T.D., Michel, J.B., Aingaran, M., Potter-Bynoe, G., Roux, D., Jr, M.R.D., Skurnik, D., Leiby, N., LiPuma, J.J., Goldberg, J.B., McAdam, A.J., Priebe, G.P. and Kishony, R., 2011. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nature Genetics*. 43(12). 1275-1280.

Lloyd, A. L., Rasko, D. A. and Mobley, H. L., 2007. Defining genomic islands and uropathogen-specific genes in uropathogenic *Escherichia coli*. *Journal of Bacteriology*. 189(9). 3532-3546.

Lo, Y., Zhang, L., Foxman, B. and Zollner, S., 2015. Whole genome sequence of uropathogenic *Escherichia coli* reveals long evolutionary history diversity and virulence. *Infection, Genetics and Evolution.* 34. 244-250.

Lund, B., Lindberg, F., Marklund, B.I. and Normark, S., 1987. The PapG protein is the alpha-D-galactopyranosyl-(1–4)–beta-D-galactopyranose-binding adhesion of uropathogenic *Escherichia coli. Proceedings of the National Academy of the Sciences of the United States of America*. 84. 5898–5902.

Lüthje, P. and Brauner, A., 2014. Virulence factors of Uropathogenic *E.coli* and their interaction with the host. *Advances in Microbial Physiology*. 65. 337-372.

Lüthje, P. and Brauner, A., 2016. Novel strategies in the prevention and treatment of urinary tract infections. *Pathogens.* 5(13). DOI: 10.3390.

Mabbett, A.N., Ulett, G.C., Watts, R.E., Tree, J.J., Totsika, M., Ong, C.Y., Wood, J.M., Monaghan, W., Looke, D.F., Nimmo, G.R., Svanborg, C. and Schembri, M.A., 2009.

Virulence properties of asymptomatic bacteriuria *Escherichia coli*. *International Journal of Medical Microbiology*. 299. 53-63.

Maclean, D., Jones, J.D.G. and Studholme, D.J., 2009. Application of next generation sequencing technologies to microbial genetics. *Nature Reviews Microbiology.* 7. 287-296.

Maluta, R.P., Logue, C.M., Cases, M.R.T., Meng, T., Guastalli, E.A.L., Rojas, T.C.G., Montelli, A.C., Sadatsune, T., Ramos, M.D.C., Nolan, L.K. and Silveira, W.D.D., 2014. Overlapped sequence types (STs) and serogroups of avian pathogenic (APEC) and human extra intestinal pathogenic (ExPEC) *Escherichia coli* isolated in Brazil. *PLoS One*. 9(8). DOI: 10.1371.

Manges, A.R., Johnson, J.R., Foxman, B., O'Bryan, T.T., Fullerton, K.E. and Rilley, L.W., 2001. Widespread distribution of urinary tract infections caused by a multidrug resistant *Escherichia coli* clonal group. *The New England Journal of Medicine*. 345(14). 1007-1013.

Mardis, E.R., 2011. A decade's perspective on DNA sequencing technology. *Nature*. 470. 198-203.

Martinez-Medina, M., Mora, A., Blanco, M., Lopez, C., Alonso, M.P., Bonacorsi, S., Nicolas-Chanoine, M.H., Darfeuille-Michaud, A., Garcia-Gil, J. And Blanco, J., 2009. Similarity and divergence among adherent-invasive *Escherichia coli* and extraintestinal; pathogenic *E.coli* strains. *Journal of Clinical Microbiology*. 47(12). 3968-3979.

Martinez, J., Mulvey, M., Schilling, J., Pinkner, J. and Hultgren, S., 2000. Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *The EMBO Journal*. 19(12). 2803-2812.

Martinez, J.L. and Baquero, F., 2002. Interaction among strategies associated with bacterial infection: pathogenicity, epidemicity and antibiotic resistance. *Clinical Microbiology Reviews*. 15(4). 647-679.

Marvig, R.L., Dolce, D., Sommer, L.M., Petersen, B., Ciofu, O., Campana, S., Molin, S., Taccetti, G. and Johansen, H.K., 2015. Within host microevolution of *Pseudomonas aeruginosa* in Italian cystic fibrosis patients. *BMC Microbiology.* 15(218). DOI: 10.1186.

McAdam, P.R., Holmes, A., Templeton, K.E. and Fitzgerald, J.R., 2011. Adaptive Evolution of *Staphylococcus aureus* during Chronic Endobronchial Infection of a Cystic Fibrosis Patient. *PLoS One*. 6(9). DOI: 10.1371.

McNally, A., Alhashash, F., Collins, M., Alqasim, A., Paszckiewick, K., Weston, V. and Diggle, M., 2013. Genomic analysis of Extra-intestinal pathogenic *Escherichia coli* urosepsis. *Clinical Microbiology and Infection*. 19(8). E328-E334.

Miles, A.A., Misra, S.S. and Irwin, J.O., 1983. The estimation of the bactericidal power of the blood. *Journal of Hygiene.* 38. 732-749.

Minardi, D., Anzeo, G.D., Conti, A. and Muzzonigro, G., 2011. Urinary tract infections in women: etiology and treatment options. *International Journal of General Medicine*. 4. 333-343.

Mobley, H.L.T. and Alteri, C.J., 2016. Development of a vaccine against *Escherichia coli* urinary tract infections. *Pathogens.* 5(1). DOI: 10.3390.

Mora, A., Lopez, C., Dahbi, G., Blanco, M., Blanco, J.E., Alonso, M.P., Herrera, A., Mamani, R., Bonacorsi, S., Moulin-Schouleur, M. and Blanco, J., 2009. Extraintestinal pathogenic *Escherichia coli* O1:K1:H7/NM from human and avian origin: detection of clonal groups B2 ST95 and D ST59 different host distribution. *BMC Microbiology*. 9(132). DOI: 10.1186.

Mulvey, M.A., 2002. Adhesion and entry of uropathogenic *Escherichia coli*. *Cellular Microbiology.* 4(5). 257-271.

Mulvey, M.A., Schilling, J.D. and Hultgren, S.J., 2001. Establishment of a persistence *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infection and Immunity*. 69.4572-4579.

Mwangi, M.M., Wei Wu, S., Zhou, Y., Sieradzki, K., Lencastre, H., Richardson, P., Bruce, D., Rubin, E., Myers, E., Siggia, E.D. and Tomasz, A., 2007. Tracking the in vivo evolution of multidrug resistance in *Staphylococcus aureus* by whole-genome sequencing. *Proceedings of the National Academy of Sciences of the United States of America*. 104(22). 9451-9456.

Nachin, L., Nannmark, U. and Nysrtöm, T., 2005. Differential roles of the universal stress proteins of *Escherichia coli* in oxidative stress resistance, adhesion and motility. *Journal of Bacteriology*. 187(18). 6265-6272.

Naves, P., Del Prado, G., Huelves, L., Gracia, M., Ruiz, V., Blanco, J., Rodriguez-Cerrato, V., Ponte, M.C. and Soriano, F., 2008. Measurements of biofilm formation by clinical isolates of *Escherichia coli* is method dependent. *Journal of Applied Microbiology*. 105 (2). 585-590.

NCCLS., 1992. National Committee for Clinical Laboratory Standards. Supplemental table M100-S4 to: Performance standards for antimicrobial disk susceptibility tests, 4th ed.

Nougayrède, J., Homburg, S., Taieb, F., Boury, M., Brzuszkiewicz, E., Gottschalk, G., Buchrieser, C., Hacker, J., Dobrindt, U. and Oswald, E., 2006. Escherchia coli induces DNA double strand breaks in eukaryotic cells. *Science.* 313(5788). 848-851.

O'Neill, M.B., Mortimer, T.D. and Pepperell, C.S., 2015. Diversity of *Mycobacterium tuberculosis* across evolutionary scales. *PLOS Pathogens*. 11(11). DOI:10.1371.

Orskov, I., Orskov, F., Birch-Anderson, A., Kanamori, M. and Svanborg Eden, C., 1982. O, K, H and fimbrial antigens in *Escherichia coli* serotypes associated with polynephritis and cystitis. *Scandinavian Journal of Infectious Diseases*. 33.18-25.

Oteo, J., Diestra, K., Juan, C., Bautista, V., Novais, A., Perez-Vazquez, M., Moya, B., Miro, E., Coque, T.M., Oliver, A., Canton, R., Navarro, F. and Campos, J., 2009. Extended spectrum β-lactamase producing *Escherichia coli* in Spain belong to a large variety of multilocus sequence typing types including ST10 complex/A, ST23 complex/A and ST131/B2. *International Journal of Antimicrobial Agents*. 34. 173-176.

Partridge, J.D., Nieto, V. and Harshey, R.M., 2015. A new player at the flagellar motor: FliL controls both motor output and Bias. *American Society for Microbiology.* 6(2). DOI: 10.1128.

Pan genomic sequence analysis., 2017. Pan seq [online] Available at: http://ifz.corefacility.ca/panseq/page/novel_full.html

Pappas, P.G., 1991. Laboratory in the diagnosis and management of urinary tract infections. *Medical Clinics of North America*. 75(2). 313-325.

Parthasarathy, G., Yao, Y. and Kim, K.S., 2007. Flagella promote *Escherichia coli* K1 association with and invasion of human brain microvascular endothelial cells. *Infection and Immunity*. 75. 2937-2945.

Pitout, J.D. and Laupland, K.B., 2008. Extended spectrum beta lactamase producing Enterobacteriaceae: an emerging public-health concern. *The Lancet Infectious Diseases*. 8(3). 159-166.

Pitout, J.D.D., 2012a. Extraintestinal pathogenic *Escherichia coli*: a combination of virulence with antibiotic resistance. *Frontiers in Microbiology*. 3(9). 1-7.

Pitout, J.D.D., 2012b. Extraintestinal pathogenic *Escherichia coli*: an update on antimicrobial resistance, laboratory diagnosis and treatment. *Expert Review of Anti-infective Therapy*. 10(10). 1165-1176.

Ponnusamy, P., Natarajan, V. and Sevanan, M., 2012. In vitro biofilm formation by uropathogenic *Escherichia coli* and their antimicrobial susceptibility pattern. *Asian Pacific Journal of Tropical Medicine*. 5(3). 210-213.

Pratt, L. and Kolter, R., 1998. Genetic analysis of *Escherichia coli* biofilm formation: roles of flagella, motility, chemotaxis and type I pili. *Molecular Microbiology*. 30(2). 285-293.

Roberts, J.A., Marklund, B.I., Ilver, D., Haslam, D., Kaack, M.B. and Baskin, G., 1994. The Gal a (1–4) Gal-specific tip adhesion of *Escherichia coli* P-fimbriae is needed for pyelonephritis to occur in the normal urinary tract. *Proceedings of the National Academy of Sciences of the United States of America*. 91. 11889–11893.

Rogers, B.A., Sidjabat, H.E. and Paterson, D.L., 2011. *Escherichia coli* O25b-ST131: a pandemic, multiresistant, community-associated strain. *Journal of Antimicrobial Chemotherapy.* 66(1). DOI: 10.1093.

Ronald, A., 2003. The aetiology of urinary tract infections: traditional and emerging pathogens. *Disease-a-Month*. 49.71-82.

Russell, P.W. and Orndorff, P.E., 1992. Lesion in two *Escherichia coli* type 1 pilus genes alter plius number and length without affecting receptor binding. *Journal of Bacteriology.* 174(18). 5923-5935.

Russo, T.A. and Johnson, J.R., 2000. Proposal for a new inclusive designation for extraintestinal pathogenic isolates of *Escherichia coli*: ExPEC. *The Journal of Infectious Diseases.* 181(5). 1753-1754.

Russo, T.A. and Johnson, J.R., 2003. Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. *Microbes and Infection*. 5. 449-456.

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A. and Barrell, B., 2000. Artemis: sequence visualisation and annotation. *Bioinformatics.* 16(10). 945-945.

Sanger.ac.uk., 2014. SMALT- Wellcome Trust Sanger Institute [online]. Available at: https://www.sanger.ac.uk/resources/software/smalt/#t_1

Schembri, M.A., Christiansen, G. and Klemm, P., 2001. Fim H-mediated autoaggregation of *Escherichia coli*. *Molecular Microbiology*. 41(6). 1419-1430.

Schmoll, T., Hoschiitzky, H., Morschhauser, J., Lottspeich, F., Jann, K. and Hacker, J., 1989. Analysis of genes coding for the sialic acid-binding adhesion and two other minor fimbrial subunits of the S-fimbrial adhesion determinant of *Escherichia coli*. *Molecular Microbiology*. 3(12). 1735-1744.

Schubert, S., Darlu, P., Clermont, O., Wieser, A., Magistro, G., Hoffmann, C., Weinert, K., Tenaillon, O., Matic, I. and Denamur, E., 2009. Role of intraspecies recombination in the spread if pathogenicity islands within the *Escherichia coli* species. *PLoS Pathogens*. 5(1). E1000257.

Seemann, T., 2014. Prokka: rapid prokaryotes genome annotation. *Bioinformatics*. 30(14). DOI: 10.1093.

Sieradzki, K., Leski, T., Dick, J., Borio, L. and Tomasz, A., 2003. Evolution of a Vancomycin-Intermediate *Staphylococcus aureus* Strain In Vivo: Multiple Changes in the Antibiotic Resistance Phenotypes of a Single Lineage of Methicillin-Resistant *S.*

aureus under the Impact of Antibiotics Administered for Chemotherapy. *Journal of Clinical Microbiology*. 41(4). 1687-1693.

Simms, A.N. and Mobely, H.L., 2008. Multiple genes repress motility in uropathogenic *Escherichia coli* constitutively expressing type 1 fimbriae. *Journal of Bacteriology*. 190(10). 3747-3756.

Skjøt-Rasmussen, L., Olsen, S.S., Jakobsen, L., Ejrnæs, L., Scheutz, F., Lundgren, B., Frimodt-Møller, N. and Hammerum, A.M., 2013. *Escherichia coli* clonal group A bactermia of urinary tract origin. *Clinical Microbiology and Infection*. 19(7). 656-661.

Smith, E.E., Buckley, D.G., Wu, Z., Saenphimmachak, C., Hoffman, L.R., D'Argenio, D.A., Miller, S.I., Ramsey, B.W., Speert, D.P., Moskowitz, S.M., Burns, J.L., Kaul, R. and Olson, M.V., 2006. Genetic adaption by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proceedings of the National Academy of Sciences of the United States of America*. 103(22). 8487-8492.

Sokurenko, E.V., Chesnokova, V., Dykhuizen, D.E., Ofek, I., Wu, X.R. and Krogfelt, K.A., 1998. Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesion. *Proceedings of the National Academy of Sciences of the United States of America*. 95. 8922–8926.

Soto, S., Marco, F., Guiral, E. and Vila, J., 2011. Biofilm formation in uropathogenic *Escherichia coli* strains: relationship with urovirulance factors and antimicrobial resistance. *Clinical Management of Complicated Urinary Tract Infection*. DOI: 10.5772.

Soto, S.M., Smithson, A., Martinez, J.A., Horcajada, J.P., Mensa, J. and Vila, J., 2007. Biofilm formation in uropathogenic *Escherichia coli* strains: relationship with prostatitis,

urovirulance factors and antimicrobial resistance. *The Journal of Urology*. 177(1). 365-368.

Spurbech, R.R. and Mobley, H.L.T., 2013. *Escherichia coli*: pathotypes and principles of pathogenesis. *Elsevier Inc, Second Edition*. 275-304.

Staples, M., Graham, R., Doyle, C., Smith, H. and Jennison, A., 2012. Prolonged and mixed non-O157 *Escherichia coli* infection in an Australian household. *Clinical Microbiology and Infection.* 18 (5). 140-143.

Stoesser, N., Sheppard, A.E., Moore, C.E., Golubchik, T., Parry, C.M., Nget, P., Saroeun, M., Day, N.P.J., Giess, A., Johnson, J.R., Peto, T.E.A., Crook, D.W. and Walker, A.S., 2015. Extensive within host diversity in fecally carried extended spectrum beta lactamase producing *Escherichia coli* isolates: implication for transmission analyses. *Journal of Clinical Microbiology.* 53(7). 2122-2131.

Subashchandrabose, S. and Mobley, H.L.T., 2015. Virulence and fitness determinant of uropathogenic *Escherichia coli. Microbiology Spectrum*. 3(4). DOI: 10.1128.

Szczepura, A.K., 1991. Efficiency in the pathology laboratories: a survey of operations management in NHS bacteriology. *Social Science and Medicine*. 33. 531-543.

Tabasi, M., Karam, M., Habibi, M., Yekaninejad, M. and Bouzari, S., 2015. Phenotypic assays to determine virulence factors of uropathogenic *Escherichia coli* (UPEC) isolates and their correlation with antibiotic resistance pattern. *Osong Public Health and Research Perspectives*. 6(4). 261-268.

Tartof, S.Y., Solberg, O.D., Manges, A.R. and Riley, L.W., 2005. Analysis of a uropathogenic *Escherichia coli* clonal group by multilocus sequence type. *Journal of Clinical Microbiology.* 43(12). 5860-5864.

Thomassen, M., Demko, C. and Boxerbaum, B., 1979. Multiple isolates of *Pseudomonas aeruginosa* with differing antimicrobial susceptibility patterns from patients with cystic fibrosis. *The Journal of Infectious Diseases*. 140. 873–880.

Tobes, R. and Pareja, E., 2006. Bacterial repetitive extragenic palindromic sequences are DNA targets for insertion sequence elements. *BMC Genomics*. 7(62). DOI: 10.1186.

Toval, F., Kohler, C.D., Vogel, U., Wagenlehner, F., Melllmann, A., Fruth, A., Schmidt, M.A., Karch, H., Bielaszewska, M. and Dobrindt, U., 2014. Characterization of *Escherichia coli* from hospital inpatients and outpatients with urinary tract infection. *Journal of Clinical Microbiology*. 52(2). 407-418.

Trautner, B.W. and Darouiche, R.O., 2004. Role of biofilm in catheter associated urinary tract infection. *American Journal of Infection Control*. 32(3). 177-183.

Tuchscherr, L., Medina, E., Hussain, M., Völker, W., Heitmann, V., Niemann, S., Holzinger, D., Roth, J., Proctor, R.A., Becker, K., Peters, G. and Löffler, B., 2011. *Staphylococcus aureus* phenotype switching: an effective bacterial strategy to escape host immunity response and establish a chronic infection. *EMBO Molecular Medicine*. 3. 192-141.

Vejborg, R.M., Hancock, V., Schembri, M.A. and Klemm, P., 2011. Comparative genomics of *Escherichia coli* strains causing urinary tract infections. *Applied and Environmental Microbiology*. 77(10). 3268-3278.

Walker, T.M., Ip, C.L.C., Harrell, R.H., Evans, J.T., Kapatai, G., Dedicoat, M.J., Eyre, D.W., Wilson, D.J., Hawkey, P.M., Crook, D.W., Parkhill, J., Harris, D., Walker, A.S., Bowden, R., Monk, P., Smith, E.G. and Peto, T.E.A., 2013. Whole genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. *The Lancet Infectious Diseases.* 13. 137-146.

Warner, D.F., Koch, A. and Mizrahi, V., 2015. Diversity and disease pathogenesis in *Mycobacterium tuberculosis. Trends in Microbiology*. 23(1). DOI: 10.1016.

Warren, R.M., Victor, T.C., Streicher, E.M., Richardson, M., Beyers, N., Pittius, N.C.G. and Helden, P.D., 2004. Patients with active tuberculosis often have different strains in the same sputum specimen. *American Journal of Respiratory and Critical Care Medicine.* 169. 610-614.

Welch, R.A., 2006. The genus *Escherichia*. *Prokaryotes*. 6. 60-71.

Wiles, T.J., Kulesus, R.R. and Mulvey, M.A., 2008. Origins and virulence mechanisms of Uropathogenic *Escherichia coli. Experimental and Molecular Pathology*. 85(1). 11-19.

Willner, D., Low, S., Steen, J.A., George, N., Nimmo, G.R., Schembri, M.A. and Hugenholtz, P., 2014. Single clinical isolates from acute uncomplicated urinary tract infections are representative of dominant in situ populations. *American Society for Microbiology*. 5(2). DOI: 10.1128.

Wolter, J.M., Kotsiou, G. and McCormack, G., 1995. Mixed morphotype testing of *Pseudomonas aeruginosa* cultures from cystic fibrosis patients. *Journal of Medical Microbiology*. 42. 220-224.

Wood, T.K., González Barrios, A.F., Herzberg, M. and Lee, J., 2006. Motility influences biofilm architecture in *Escherichia coli. Applied of Microbial Biotechnology*. 72(361). 361–367.

Wright, K.J., Seed, P.C. and Hultgren, S.J., 2005. Uropathogenic *Escherichia coli* Flagella Aid in efficient Urinary Tract Colonization. *Infection and Immunity*. 73(11). 7657–7668.

Young, B.C., Goluchik, T., Batty, E.M., Fung, R., Larner-Svensson, H., Votintseva, A.A., Miller, R.R., Godwin, H., Knox, K., Everitt, R.G., Iqbal, Z., Rimmer, A.J., Cule, M., Ip, C.L.C., Harding, R.M., Donnelly, P., Peto, T.E., Crook, D.W., Bowden, R. and Wilson, D.J., 2012. Evolutionary dynamics of *Staphylococcus aureus* during progression from

carriage to disease. *Proceedings of the National Academy of Sciences of the United States of America*. 109(12). 4550-4555.

Zdziarski, J., Brzuszkiewicz, E., Wullt, B., Liesegang, H., Biran, D., Viogt, B., Grönberg-Hemandez, J., Ragnarsdottir, B., Hecker, M., Ron, E.Z., Daniel, R., Gottschalk, G., Hacker, J., Svanborg, C. and Dobrindt, U., 2010. Host imprints on bacterial genomes-rapid, divergent evolution in individual patients. *PLoS Pathogens*. 6(8). E1001078.

Zdziarski, J., Svanborg, C., Wullt, B., Hacker, J. and Dobrindt, U., 2008. Molecular basis of commensalism in the urinary tract: low virulence or virulence attenuation?. *Infection and Immunity*. 76(2). 695-703.