# Host-bacterial interactions underlying recurrent urinary tract infections

**Aaron Tan Ming Zhi (BSc)**

Thesis submitted for the degree of Doctor of Philosophy

Biosciences Institute (Institute for Cell and Molecular Biosciences)

Faculty of Medical Sciences

Newcastle University

**August 2021**

# Abstract

Although most urinary tract infections (UTIs) are easily treated with antibiotics, many patients present with long-term bacterial colonisation of the urinary tract. Amongst these patients, some remain asymptomatic but others exhibit severe and frequent UTI symptoms. Currently, there is little understanding of the host-microbial interactions underpinning these clinical conditions. Hence further understanding of these interactions underlying the pathogenesis of asymptomatic bacteriuria (ASB) and recurrent UTIs is essential for effective clinical management. To explore these interactions further, genomic data from two clinical studies (AnTIC & BUTI) were mined and bacterial isolates exploited to determine the effects of antibiotic prophylaxis (AnTIC) and host-specific factors (BUTI) on long-term bacterial colonisation.

Using bioinformatic approaches, the AnTIC dataset was analysed to investigate the impact of antibiotic prophylaxis on bacterial colonisation of the urinary tract. The findings suggested that antibiotic prophylaxis reduced the frequency of symptomatic UTIs by not only clearing bacteria from the urinary tract, but also in a subset of patients through stabilising *Escherichia coli* colonisation. An association between stable colonisation and a reduction in UTI frequency was also observed in the BUTI study, but with a focus on *Proteus mirabilis* rather than *E. coli*. In these patients, host factors appeared important with patients characterised by a TLR1 1805T SNP associated with stable and clinically asymptomatic *P. mirabilis* colonisation. *In vitro* experiments using bladder RT4 cells and TLR agonists supported a role for TLR1 in modulating the immune response, which may play a role in ASB.

Analysis of *E. coli* isolated from AnTIC patients supported a link between stable long-term bacterial colonisation and increased antibiotic resistance. In the UK, nitrofurantoin underpinned by the low incidence of resistance amongst uropathogens, is prescribed for managing UTIs. *In vitro* experiments using mutagenised clinical *E. coli* isolates were performed to investigate potential factors driving nitrofurantoin resistance. In the absence of nitrofurantoin, bacterial growth data demonstrated that nitrofurantoin resistant ($Nit^R$) strains grew as well as sensitive strains ($Nit^S$ doubling time= 21.9 ± 0.9 min; $Nit^R$ mutants= 25.7 ± 0.4 min; $Nit^R$ natural= 25.7 ± 1.1 min). However, in the presence of nitrofurantoin (8 – 16 µg/ml), resistant strains showed a fitness advantage, with the average doubling time of these mutants reduced by 35% ($p<0.001$). These results in combination with clinical pharmacokinetics data suggest that urinary nitrofurantoin concentrations in prophylaxis therapy driving fitness via selective advantage, which potentially results in the selection of $Nit^R$ isolates.

Taken together, these studies show that investigating both host and microbial components and how they interact are essential to understanding the pathogenesis of UTI and will help improve the management of recurrent UTI sufferers.

## Acknowledgements

I would like to take this opportunity to express my immense gratitude for the unwavering support of my supervisors, Phil Aldridge and Judith Hall, throughout this PhD project. Judith has done an invaluable job of keeping me scientifically grounded both in thinking and writing, especially the latter due to my tendency of writing my scientific reports like a Jane Austen novel and explaining things like Humphrey Appleby, which is evidenced by how this sentence is written. Special thanks go to Phil for being a mentor and a good friend, I can always count on him to be there for me in both fair and stormy times. In all, I couldn't have asked for a better supervisory team and they will be sorely missed.

I wish to also show my appreciation to my colleagues who have assisted me in my research endeavours. Chris Boyles, for patiently sharing his wet-lab expertise with me during the first year of my PhD. Catherine Mowbray, for the nice chats we had in the office and for teaching me how to do cell culture and RT-PCR. Maxime Vallee, for going through absolute hell by sorting out the raw AnTIC bacterial samples with Phil, and bringing in the much-needed funding to perform whole genome sequencing on bacterial isolates. Indeed, without the whole genome sequence data and isolates from the AnTIC study, Chapters 3 and 4 of this thesis would not have existed.

I should also not forget my friends at Jesmond Parish Church; Philip Dove, KC, Emily, Adam, Jonathan, Chris, Wendy, Henry, Dave, Andy, Nicole, Katie, Susan, Ben, Jerry, Dominic, Nick, and Paul; for keeping me sane during the lockdowns and making my long stay in Newcastle an absolute joy. I also owe a debt of gratitude to Newcastle University for subsidising the cost of my studies by providing me with the NUORS scholarship. Finally, and most importantly, I want to thank my parents and my sister who have always been there for me through thick and thin during the past 26 years, I wouldn't be where I am now if it weren't for them.

# Table of contents

# List of Figures

# List of Tables

# Abbreviations

| | |
|---|---|
| AMP | Antimicrobial peptide |
| AMR | Antimicrobial resistance |
| AnTIC | Antibiotic Treatment for Intermittent Bladder Catheterisation |
| ASB | Asymptomatic bacteriuria |
| BLAST | Basic local alignment search tool |
| BUTI | Biomarkers in Urinary Tract Infection study |
| *cat* | Chloramphenicol resistance gene |
| CAUTI | Catheter-associated UTI |
| cgMLST | Core genome multi locus sequence typing |
| CISC | Clean intermittent self-catheterised |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| *dfr* | Dihydrofolate reductasegene |
| DSB | Double-stranded break |
| EC | *E. coli* |
| ELISA | Enzyme Linked Immunosorbent Assay |
| ExPEC | Extraintestinal pathogenic *E. coli* |
| GP | General practitioner |
| IBC | Intracellular bacterial colonies |
| IFN | Interferon |
| IgA | Immunoglobulin A |
| LB | Luria broth |
| LBP | LPS-binding protein |
| LPS | Lipopolysaccharide |
| LRR | Leucine-rich repeat |
| MDR | Multi drug resistant |
| MHB | Muller Hinton broth |
| MIC | Minimum inhibitory concentration |
| MLST | Multi-locus sequence typing |
| mRNA | Messenger RNA |
| MW | Molecular weight |
| *nfsA* | Nitroreductase A gene |
| *nfsB* | Nitroreductase B gene |
| NGS | Next-generation sequencing |
| NICE | National Institute for Health and Care Excellence |
| Nit | Nitrofurantoin |
| $Nit^R$ | Nitrofurantoin resistant |
| $Nit^S$ | Nitrofurantoin sensitive |
| OLC | Overlap layout consensus |
| PAMP | Pathogen-associated molecular pattern |
| PBS | Phosphate-buffered saline |
| PCR | Polymerase chain reaction |
| PM | *P. mirabilis* |
| PRR | Pathogen recognition receptor |

| | |
|---|---|
| rUTI | Recurrent urinary tract infection |
| sgRNA | Single guide RNA |
| SNP | Single nucleotide polymorphism |
| ST131 | Sequence type 131 |
| TLR | Toll-like receptor |
| TNF | Tumour necrosis factor |
| TPF | Type P fimbriae |
| UPEC | Uropathogenic E. coli |
| UTI | Urinary tract infection |
| WT | Wild-type |

# Chapter 1 Introduction

## *1.1 Urinary tract infection (UTI)*

Urinary tract infection (UTI) is an umbrella term describing infections of the urinary tract, and impacting either the urethra, bladder, ureter, and kidneys [1]. Most infections are limited to the urethra and bladder (lower urinary tract) causing lower UTI, but it can spread to the ureters and kidneys (upper urinary tract) causing upper UTI (Table 1) [2]. The symptoms associated with UTIs are driven by the inflammatory responses and dependent on which part of the urinary tract is infected [3]. Individuals with lower UTI often present with symptoms of urethritis (inflammation of the urethra) and cystitis (inflammation of the bladder). Urethritis is accompanied by a burning sensation during urination while cystitis is characterised by (i) increased frequency and urgency of urination, (ii) bloody/cloudy/dark/odorous urine, (iii) pelvic pain in females and rectal pain in males. Individuals with upper UTI present with symptoms of ureteritis (inflammation of the ureter) and/or pyelonephritis (inflammation of the kidney), which are typified by (i) pain and tenderness in the upper back and sides and (ii) fever-like symptoms, which include high temperatures, chills, nausea, and vomiting [4]. An untreated upper UTI can lead to bacteraemia (presence of bacteria in the bloodstream) due to the rich vasculature of the kidneys, which can eventually turn into life-threatening urosepsis [5].

UTIs commonly affect females, with fifty to eighty percent of this population experiencing an UTI episode throughout their lifetime [6]. The prevalence of UTIs generally increases with age barring a spike in young females aged 14-24 [7]. Compared to females, males are less likely to experience UTIs due to their longer urethra, which makes it more difficult for pathogens to infect the male urinary tract [8]. Besides sex and age, there are other risk factors contributing to UTIs which include (i) pregnancy, (ii) menopause and hormonal disorders, (iii) sexual intercourse, (iv) use of topical spermicidal agents, (v) drug-associated immunosuppression, (vi) diabetes, (vii) neurological disorders or physical abnormalities affecting the genitourinary tract, and (viii) catheter usage [9]. The presence of certain risk factors (risk factors vi-viii) can influence the clinical diagnosis and management of UTI [2]. Patients who present with UTI and have these risk factors are usually classified as having complicated UTI (risk factors vi-viii) or catheter-associated UTI (risk factor viii)

(Table 1) whereas patient without these risk factors are clinically classified as having uncomplicated UTI (Table 1) [2] .

Most of the aforementioned risk factors can also predispose an individual to recurrent UTI (rUTI), which is either defined as ≥ 2 symptomatic infections within 6 months or ≥ 3 symptomatic infections within 12 months (Table 1) [10]. Recurrent UTI is common even among individuals with no apparent risk factors, as an example, a 6-month longitudinal study of healthy college-aged women found a 27% recurrence rate [11]. The high incidence of antibiotic resistance among uropathogens adds another layer of complication to the management of rUTI as it can result in treatment failures [12,13]. Nevertheless, continuous low-dose antibiotic prophylaxis has so far proven to be an effective therapeutic option to manage both complicated and uncomplicated rUTI [14,15].

Individuals can also present with bacteria in their urine (bacteriuria) without any symptoms that are associated with an UTI. This is known as asymptomatic bacteriuria (ASB) and is particularly common among elderly individuals [16]. The incidence of ASB is 15% or greater in both sexes from the age 65 to 80 and can be as high as 40-50% after the age of 80 [17]. With the exception of pregnant women and patients who will undergo surgery involving the urogenital tract, current recommendations advise that individuals presenting with ASB should not be treated with antibiotics [16]. Bacterial strains that are associated with ASB may even be beneficial as their presence may reduce the incidence of UTI by preventing pathogenic strains from invading and establishing themselves in the urinary tract [18]. Interestingly, a study reported that the deliberate inoculation of one such strain (*E. coli* 83972) in patients who were refractory to conventional therapy proved to be successful at reducing UTI frequency [19].

### *1.1.1 Diagnosis of UTI*

As highlighted earlier, UTIs are clinically classified according to: (i) the extent of anatomical involvement, (ii) the presence of an abnormality in the urinary tract or catheter usage, and (iii) frequency of symptomatic infections (Table 1).

**Table 1 Clinical classification of UTI.**

| Criteria | Clinical classification | Detail |
|---|---|---|
| Extent of anatomical involvement | Lower UTI | Infections that are limited to the lower urinary tract |
| | Upper UTI [a] | Infections that involve the upper urinary tract in addition to the lower urinary tract |
| Presence of an abnormality in the urinary tract or catheter usage | Uncomplicated UTI | Infections in patients with a normal urinary tract |
| | Complicated UTI [b] | Infections in patients with an abnormality [c] that affects the genitourinary tract or who are catheter users. |
| | Catheter-associated UTI (CAUTI) | Infections in patients who are catheter users. |
| Frequency of symptomatic infections | Acute UTI | <3 symptomatic infections within 12 months or<br><2 symptomatic infections within 6 months |
| | Recurrent UTI (rUTI) or chronic UTI | ≥ 3 symptomatic infections within 12 months or<br>≥ 2 symptomatic infections within 6 months |
| | Asymptomatic bacteriuria (ASB) | Asymptomatic infection of the urinary tract |

[a] Due to the anatomical structure of the urinary tract and the pathophysiology of UTI, it is rare to encounter an upper UTI that does not also involve the lower urinary tract. [b] It should be noted that UTIs in pregnant females and UTIs in healthy males are clinically treated as complicated UTIs. [c] Examples of such abnormalities include physical abnormalities like urolithiasis/nephrolithiasis (bladder/kidney stones), benign prostatic hyperplasia (enlargement of the prostate), ureteral stents, pelvic organ prolapse; physiological abnormalities like diabetes and immunosuppression; and neurological abnormalities that affect bladder function (neurogenic bladder) like spina bifida, cerebral palsy, stroke, Parkinson's disease, multiple sclerosis, and spinal cord injuries

In daily practice, physicians rely on common clinical presentations including frequency, dysuria, urgency, suprapubic pain, fever, haematuria, and cloudy/odorous urine to diagnose UTIs (Figure 1). Ideally, while dipsticks are easy to use, such evaluations should be reserved for more ambiguous cases. For example, when the patient presents with less than three mild UTI-linked symptoms (Figure 1). However, in actual clinical practice, most physicians tend to use the dipstick on a routine basis for diagnostic confirmation [20,21].

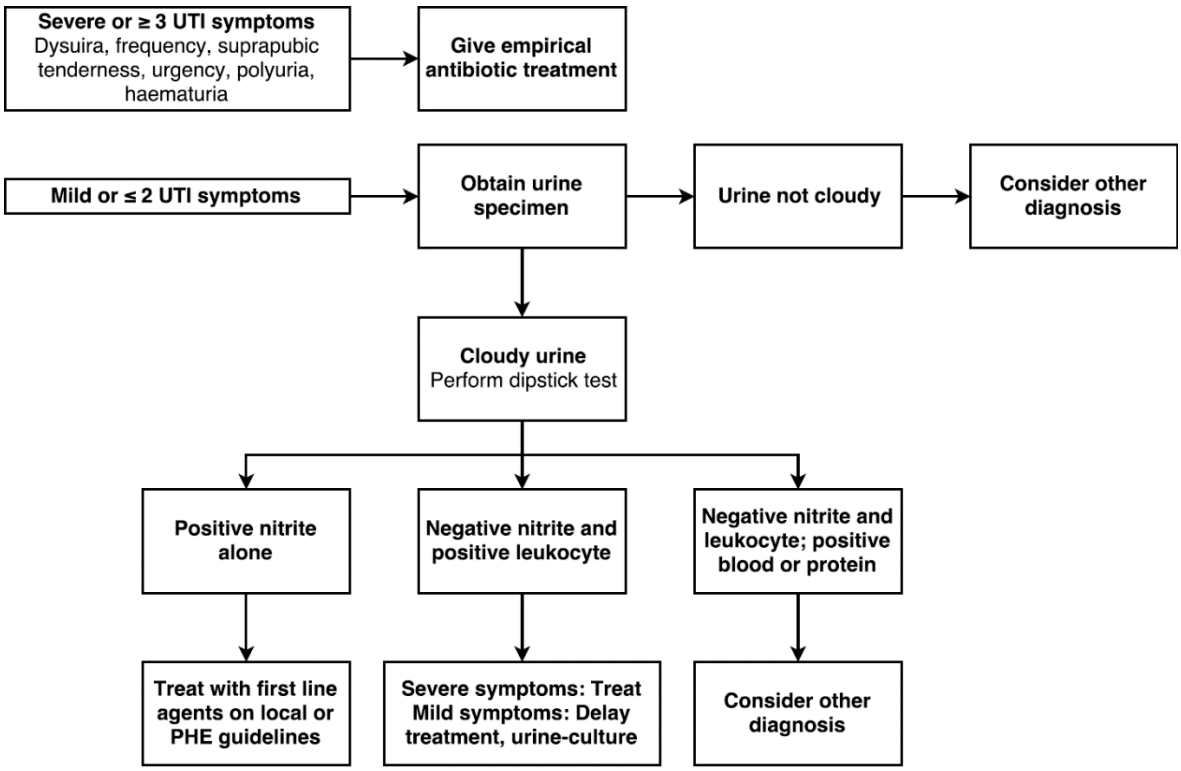



**Figure 1: Clinical workflow of managing acute UTI patients in the primary care setting.** Adapted from Public Health England guidelines [22].

The dipstick test uses strips containing reagent pads that can semi-quantitatively assess nitrite (Gram-negative uropathogen metabolite), leukocyte esterase (inflammatory marker), protein content and blood. Of these four markers, nitrite and leukocyte esterase are commonly used in tandem as positive predictors of an UTI[23]. However, their low sensitivity (positive leukocyte esterase & nitrite: 22.9%, positive nitrite alone: 42.9%) [24] means that a negative result for either nitrite or leukocyte esterase cannot be used to confidently rule out UTI in symptomatic patients, necessitating urine culture for full diagnosis and effective management [25,26]. Nevertheless, compared to urine culture, dipsticks have the advantage in terms of cost, speed, and simplicity [27]. Their high specificity (positive leukocyte esterase & nitrite: 99.1%, positive nitrite alone: 98%) [24] also enables physicians to confidently and immediately prescribe antimicrobials to patients who do test positive (Figure 1) without jeopardising antibiotic stewardship.

The gold-standard for UTI diagnosis in the UK is urine culture of clean-catch midstream urine on Cysteine Lactose Electrolyte Deficient (CLED) or chromogenic agar plates (bioMérieux CHROMID® CPS® Elite) [28]. The diagnostic threshold used in UK laboratories is $10^4$ - $10^5$ Colony Forming Units per millilitre (CFU/ml) of a single uropathogenic species (Table 2) [28,29]. Culture positive samples that (i) fail to meet the CFU threshold, or (ii) have more than one uropathogenic species, or (iii) do not have

any of the listed uropathogenic species are classified as being contaminated and clinically disregarded [28,29] (Table 2).

**Table 2 Organisms implicated in uncomplicated and complicated UTI.** [28]

| | |
|---|---|
| **Uncomplicated UTI** | *Escherichia coli, Proteus mirabilis, Staphylococcus saprophyticus, Streptococci spp., Enterococci spp., Candida spp.* |
| **Complicated UTI** | *E. coli, Klebsiella spp., Enterobacter, Proteus spp., Enterococcus spp., Pseudomonas aeruginosa, Staphylococcus aureus,* coagulate negative *Staphylococci* |

### *1.1.2 Management of UTI*

According to the National Institute for Health and Care Excellence (NICE) guidelines, it is recommended that immediate antibiotic treatment should be prescribed for all UTI cases [30]. In non-pregnant women with mild UTI symptoms, patients can be given an antibiotic prescription, to be used if there is no improvement in 48 hours or if the symptoms worsen [30]. In addition to antibiotics, patients are usually advised to maintain adequate hydration and practice proper hygienic practices with regards to the genitourinary area (examples include wiping from front to back after defaecation, not delaying urination, and avoid wearing occlusive underwear) [31]. Patients may also be advised to consume cranberry products or D-mannose supplements as a preventative measure, especially for rUTI [31]. However, a Cochrane review has already deemed cranberry products to be ineffective [32] whereas the efficacy of D-mannose supplements is still under investigation due to the lack of high quality studies [33].

As for antibiotic management, oral trimethoprim or nitrofurantoin are generally the first-line antibiotic for managing lower UTIs whereas oral cefalexin or ciprofloxacin are used for managing upper UTIs [31] (Table 3). It should be noted that there is a high incidence of trimethoprim and β-lactam resistance among uropathogens, which has been increasing over the past few decades [13]. In response to this, several guidelines have been recently revised since 2017 [34] to promote nitrofurantoin in place of trimethoprim as a first-line therapy for managing uncomplicated lower UTIs [30,31]. The UK public health service (National Health Service, NHS) has taken additional measures to accelerate this transition [35]. One such measure was the introduction of antibiotic prescription targets via the NHS England Quality Premium Scheme [36]. This had the benefit of decreasing trimethoprim prescription among general practices by 53.6% and increasing nitrofurantoin prescription by 68.1% between 2015 and 2019

[35]. Unlike other antibiotics, the incidence of nitrofurantoin resistance has not significantly increased since its introduction [37].

Other antibiotics, oral amoxicillin or pivmecillinam hydrochloride, may be used as second line options in the event of treatment failure [31] (Table 3). Due to the high incidence of amoxicillin resistance among uropathogens, it should however, only be used if the antibiotic sensitivity report indicates that the infecting organism is susceptible [31]. Intravenous antibiotics are generally reserved for patients who are severely unwell or unable to take oral treatment [31] (Table 3). Such patients are usually suffering from an upper UTI and are at a higher risk of developing sepsis [38].

**Table 3 General antibiotic prescription guideline for the management of UTI.**

| | |
|---|---|
| **Lower uncomplicated acute UTI** | First line: Oral nitrofurantoin or trimethoprim<br>Second line: Oral nitrofurantoin, fosfomycin, pivmecillinam hydrochloride, or amoxicillin |
| **Lower complicated acute UTI** | First-line: Oral nitrofurantoin<br>Second-line: Oral amoxicillin or cefalexin |
| **Lower acute CAUTI** | First-line: Oral amoxicillin, nitrofurantoin, or trimethoprim<br>Second-line: Oral pivmecillinam hydrochloride |
| **Upper acute UTI** | First-line: Oral cefalexin or ciprofloxacin<br>Second-line: Intravenous amikacin, ceftriaxone, cefuroxime, ciprofloxacin, or gentamicin |
| **Upper acute CAUTI** | First-line: Oral cefalexin, ciprofloxacin, co-amoxiclav, or trimethoprim<br>Second-line: Intravenous amikacin, ceftriaxone, cefuroxime, ciprofloxacin, or gentamicin |
| **Lower uncomplicated rUTI** | Unlicensed: Vaginal oestrogen (only postmenopausal women), single-dose antibacterial prophylaxis |
| **Asymptomatic bacteruria** | Should not be managed with antimicrobials unless the patient is a pregnant female (oral amoxicillin, nitrofurantoin, or cefalexin) |

Adapted from the 81st edition of the British National Formulary [31]

### *1.1.3 Known shortcomings of current diagnostic and management techniques*

Accurate diagnosis is the cornerstone of effective management of any infectious disease. The standard diagnostic workflow (Figure 1) that is commonplace in most UTI guidelines is geared towards managing acute cases, but fails to address the additional complexity associated with recurrent infections [22]. It was only in 2018, that a national, evidence-based guideline on how to effectively diagnose and manage rUTI was released [39]. However, the main hurdle to effective rUTI management lies

within the diagnostic tests themselves. For example, the diagnostic threshold for urine cultures in UK laboratories is $10^4$ - $10^5$ CFU/ml of a single uropathogenic species, which is arguably high for a symptomatic UTI.

The $10^4$ - $10^5$ CFU/ml threshold was derived from a 1957 pyelonephritis study [40]. The authors utilised urine samples from a cohort of pregnant females with acute, severe pyelonephritis as the basis for their proposed diagnostic threshold [40]. Despite this sample being unrepresentative of those with lower urinary tract symptoms, this threshold has been adopted and is still being used as the standard for UTI diagnosis in most countries as of writing [41,42]. Unfortunately, this threshold has also impacted the urinary dipstick test, which was developed and calibrated using microscopic and microbiological reference standards [43]. For example, the nitrite component of the dipstick test selectively screens for nitrate reducing bacteria, of which most are Gram-negatives [44]. Furthermore, the nitrite threshold utilised in dipsticks is derived from the same $10^5$ CFU/ml cut-off proposed by the 1957 pyelonephritis paper [41]. Another diagnostic component is leukocyte esterase, which is a surrogate marker for microscopic pyuria, i.e. presence of white blood cells in the urine. Unfortunately, microscopic pyuria by itself is yet another surrogate marker for bacterial infection [41]. Multiple studies have demonstrated that both dipstick and microscopic pyuria are poor indicators of UTI, with the dipstick being a poor surrogate marker of leukocyte presence [25,43,45].

Polymicrobial growth in urine cultures is often referred to as sample contamination [46]. Interestingly, one interpretation of Koch's original postulates to understand disease pathogenesis has resulted the proposal whereby one pathogen = one disease [47]. This is evidenced in the 1957 pyelonephritis paper where the authors assumed that coliform organisms were the dominant pathogens while dismissing polymicrobial growth as likely contamination [40]. However, recent studies have found that most UTI cases are associated with the presence of two or more bacterial species, with monomicrobial colonisation being the minority in terms of prevalence [46,48–50]. Indeed, this concept of polymicrobial colonisation in disease pathogenesis has already been highlighted in other studies investigating clinical conditions such as bacterial vaginosis, inflammatory bowel syndrome, dental caries, and otitis media [51,52].

However, these 1957 principles are still impacting UTI diagnosis. For example, it is standard practice for medical laboratories to immediately dismiss urine samples with

(i) sub-threshold levels of bacteriuria; (ii) presence of Gram-positive bacteria; or (iii) polymicrobial growth as being “contaminated” [53–57]. Such samples are usually withheld from further processing and analysis, thereby preventing physicians from acquiring useful clues, e.g. antimicrobial resistance, that could be used to inform therapeutic decisions. Even if a laboratory decides to proceed with the full analysis, there is still a risk that the responsible uropathogen(s) escape detection due to the limitations inherent to current urine culture protocols [48,58,59].

As a result of such diagnostic practices, rUTI patients can be generally poorly managed, leading to poorer outcomes and diminished quality of life [60,61]. Without any accurate information to base therapeutic decision making on, antibiotic prescription becomes a wearisome cycle of trial and error for the patient. This translates into more GP visits, accelerated development of antibiotic resistance, and eroded patient confidence [41]. In the end it is the patients that suffer as they are compromised by insensitive tests based on obsolete models, resulting in untreated or mistreated infections [62]. It can be therefore argued that there is therefore an urgent need to revisit and revise the aetiological and pathological mechanisms underlying UTI. Only then, can the current system of diagnosing and managing UTI, especially rUTI be reformed and improved.

## *1.2 Host defences*

### *1.2.1 Anatomy of the urinary tract and the external genitourinary region within the context of UTI*

A general understanding of the normal anatomy of the male and female urinary tract is required in order to appreciate the aetiology and pathophysiology of UTI. The urinary tract is a contiguous, hollow-organ system that collects, transports, stores, and expels urine in a periodic and highly-coordinated manner. Each kidney is composed of nephrons which filter metabolic by-products and toxins from the blood. These filtered waste products diluted in water comprise urine, which is transported from the kidneys into the bladder via a pair of long fibromuscular tubes called ureters (Figure 2A). In the ureters, urine is transported via (i) peristaltic activity during periods of low urine flow or (ii) passive pressure gradient in an open system during periods of high urine flow. There is always a constant flow of urine from the kidneys to the bladder in a healthy urinary tract.

The bladder is a hollow muscular organ that temporarily stores urine from the kidneys (Figure 2). The bladder has three openings, which comprise a pair of ureteric orifices where urine from the ureters drains into the bladder and the internal urethral orifice where urine leaves the bladder via the urethra during urination (Figure 2A). The ureteric orifices have mucosal flaps that act as valves which prevent the backflow of urine into the ureters. During urination, the smooth muscle of the bladder contracts while both the internal and external urethral sphincters relax, forcefully ejecting urine from the bladder into the urethra and out to the external environment. Frequent micturition due to bladder filling and voiding produces high shear forces in the bladder and urethra [63]. This is remarkably effective at flushing out uropathogens that lack the necessary adhesion proteins that facilitate binding to uroepithelial cells.

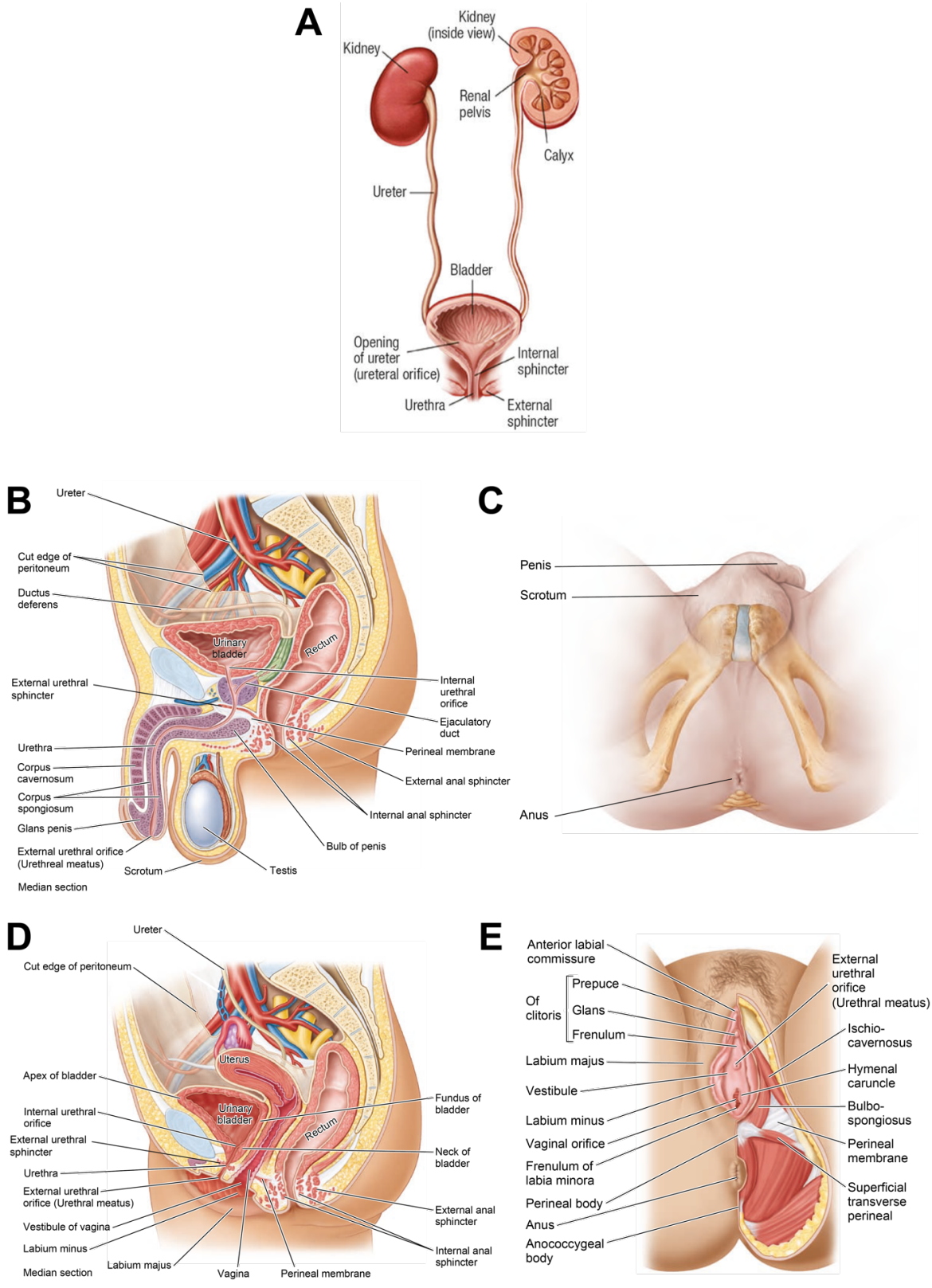



**Figure 2 Anatomy of the male and female urinary tract. (A)** Anatomical overview of the urinary tract showing its major components **(B)** Median section view of the male genitourinary system. Note the length of the male urethra compared to the female urethra in Figure 2D **(C)** Inferior view of the male perineum, note the distance between the anus and the penis where the external urethral orifice is situated. **(D)** Median section view of the female genitourinary system. (E) Inferior view of the female perineum. Note the proximity of the external urethral orifice (urinary meatus) to the vaginal orifice and the proximity of the vaginal orifice ot the anus. Adapated from Moore Essential Clinical Anatomy, Pearson Human Anatomy and Physiology, and Oncology Pro [64–66]

The main difference between the male and female urinary tract is the urethral length (Figure 2B,D). In males, the urethra is usually between 13 - 20cm long, whereas in females, the urethra is considerably shorter at around 3.8 - 5.1cm long [8]. The male urethra runs vertically through the prostate (3-4cm), then through external urethral sphincter (2-2.5cm), then through the corpus spongiosum of the penis (~15cm), slightly dilating in the glans penis before terminating at the external urethral orifice (urethral meatus) [8] (Figure 2B). The female urethra passes obliquely from the bladder to the urethral meatus along the anterior vaginal wall, with the distal two-thirds of the urethra being surrounded by the external urethral sphincter [8] (Figure 2D).

The distal end of the urethra (outermost 1.5cm section of urethra) and the urethral meatus in both males and females are particularly rich in microflora. In females, the external urethral meatus opens into the vulval vestibule, immediately in front of the vaginal orifice (Figure 2E). The vulval vestibule is a microflora-rich region which has been strongly implicated in the pathogenesis of recurrent UTI [67]. This region remains poorly researched, with recent publications focusing on the perianal, vaginal or periurethral microflora.

In healthy, pre-menopausal females of reproductive age, the vaginal microflora is dominated by a few *Lactobacillus species*, which includes *L. crispatus*, *L. gasseri*, *L. jensenii*, and *L. iners*. During puberty, the rise in oestrogen levels promote glycogen deposition in the vaginal epithelial cells as well as their proliferation (Figure 3) [68]. Vaginal epithelial cells are constantly shed (desquamated), desquamated cells undergo a lytic process which releases the glycogen deposits into the environment [68]. Glycogen is metabolised by the *Lactobacilli* via anaerobic fermentation, producing lactic acid as the primary metabolic by-product [68]. Lactic acid creates a low pH environment in the vagina, making it a hostile environment for most pathogenic bacteria (Figure 3) [69–71]. This is further enhanced by the secretion of antimicrobial compounds by *Lactobacilli* which includes bacteriocins, hydrogen peroxide and surfactants [71].

Should the vaginal lactobacilli be diminished, the vagina can become a reservoir for *E. coli* and other uropathogenic species (Figure 3). This can increase predisposition to a variety of urogenital diseases, especially UTI due to the proximity of the vagina to the urethral meatus in females (Figure 2E). Vaginal colonisation by *E. coli* can

subsequently seed the urinary bladder with these pathogens, prompting an episode of rUTI . This explains why postmenopausal women are at a higher risk of developing rUTI as the loss of oestrogen can cause the vaginal *Lactobacilli* population to diminish (Figure 3). Along this line, the use of intravaginal oestrogen has been proven to be effective at preventing/minimising rUTI in this population [72].

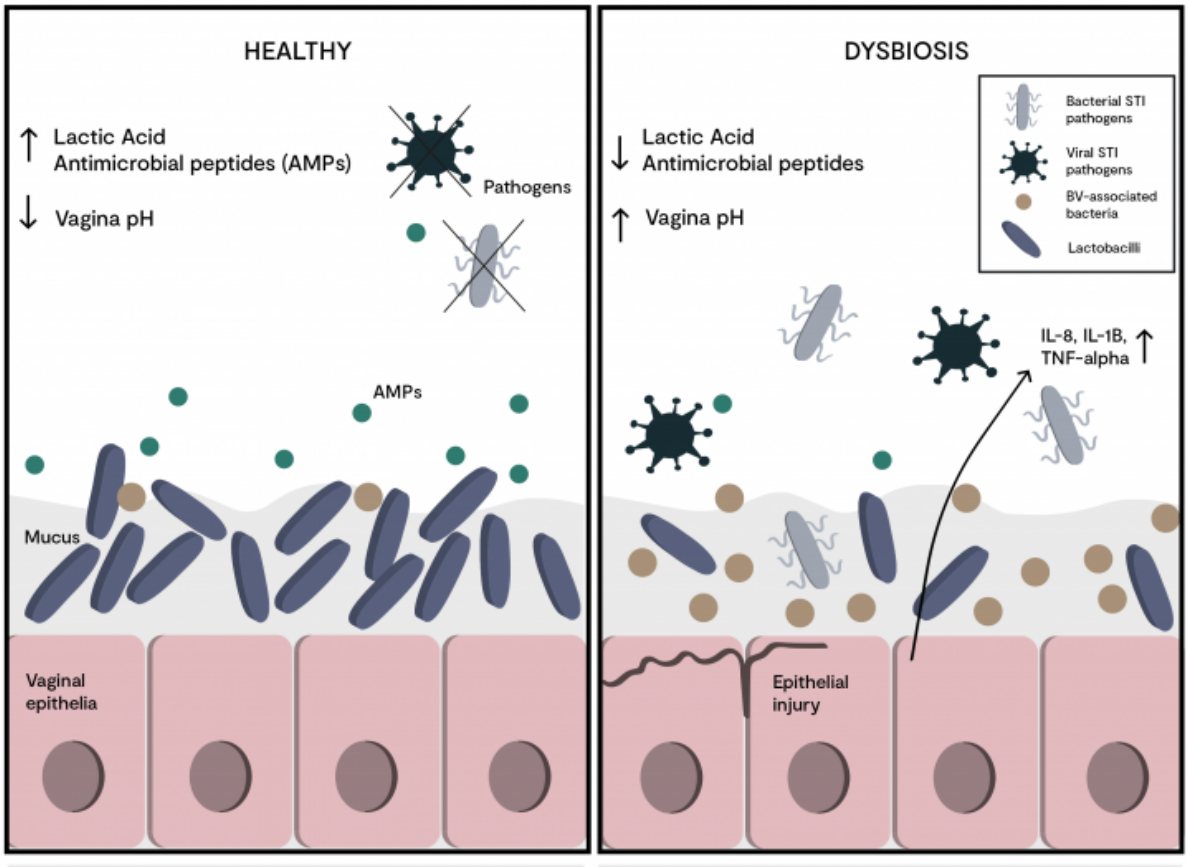



**Figure 3 Healthy vs dysbiotic vaginal microflora.** In a healthy vaginal microflora, *Lactobacilli* creates an inhospitable environment for most pathogens by generating (i) lactic acid which lowers the pH of the vaginal environment and (ii) antimicrobial peptides (bacteriocin) and compounds (hydrogen peroxide) that kill pathogens by disrupting their cellular membrane. In a dysbiotic vaginal environment, the diminishment of *Lactobacilli* lowers the production of lactic acid which in turn raises vaginal pH. Since the production of antimicrobial peptides and hydrogen peroxide is dependent on a low pH environment, the rise in vaginal pH also decreases the production of these compounds. This allows pathogenic bacteria and viruses to colonise the vagina, inducing a pro-inflammatory response (via the secretion of IL-8, IL-1 β, and TNF-alpha) that can perpetuate the dysbiotic condition. Adapted from a figure obtained from the Invivo Healthcare website. [73]

In males, the external urethral meatus is located at the tip of the glans penis that is normally covered by the prepuce (foreskin) in uncircumcised males (Figure 2B). The inner prepuce is a mucous membrane that can provide a favourable environment for microorganisms [74–76]. Unlike the vaginal microflora, the preputial microflora is more of an infection reservoir rather than protective [74]. Compared to circumcised males, uncircumcised males are 9.1 times more likely to be hospitalised with severe UTI, with long-term health complications [77]. However, owing to the low incidence of UTI

among males combined with the ease and acceptability of circumcision, there has been little interest to investigate the role of the preputial microflora in the pathogenesis of UTI among males.

The proximity of the female urethral orifice to the anus (Figure 2D, E) increases the likelihood of microbial transfer, especially during defaecation. UTI is generally thought to begin with the extraneous transfer of gut-dwelling uropathogens from the perianal to the periurethral region as a result of unhygienic practices [78]. Although it has been suggested that gut-dwelling bacteria may intrinsically migrate from the perianal to the periurethral region via the perineal region [79], existing evidence does not unanimously support this behaviour [80]. Case in point, a clinical study demonstrated that the application of antimicrobials to the perineum failed to reduce the risk of rUTI [81]. In males, the unintentional transfer of microbes from the perianal to the periurethral region is almost impossible to accomplish as the periurethral orifice is located in the penis, which is located anteriorly, thereby greatly increasing the distance between the anus and the periurethral orifice (Figure 2B, C). However, physical handling of the penis is inevitable during urination, and this may transfer skin-dwelling or even gut-dwelling organisms onto the prepuce. Nevertheless, the long male urethra usually acts as a sufficiently robust barrier to prevent these organisms from reaching the bladder.

### *1.2.2 Histology of the urothelium and its constitutive defences*

The urothelium is an epithelial tissue that lines most of the urinary tract, including the renal pelvis, ureters, and the bladder in both males and females [82]. A portion of the urethra is also covered by this tissue, but the coverage length differs between both sexes [83,84]. In females, the innermost two-thirds of the urethra is lined by the urothelium [83] whereas in males, the length of the urethra is that lined by the urothelium is proportionally shorter as it only covers the pre-prostatic and prostatic urethra [84].

The main function of the urothelium is to form a robust yet distensible barrier that can accommodate large changes in urine volume without any leakage between the blood-urine barrier. It is stratified and is composed of three layers: basal, intermediate, and umbrella (Figure 4). Compared to the gut epithelia, the turnover rate of the urothelium is slower, taking around 3-6 months [85,86]. However, the urothelium can display enormous regenerative capability when it is damaged as

evidenced by its rapid restoration of integrity within days of significant damage [87]. This is an especially important property given the cytotoxic nature of urine [88].

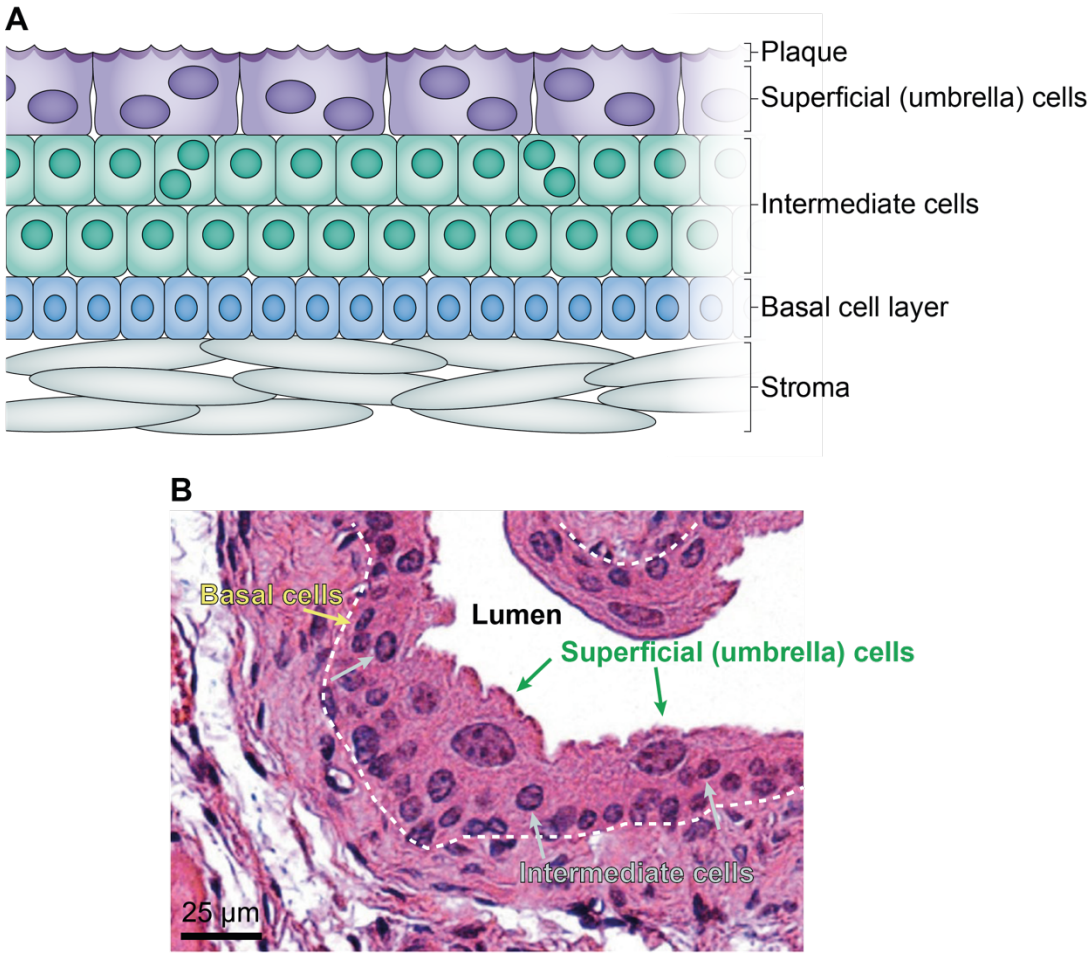



**Figure 4 Structure of the urothelium. (A)** The urothelium is composed of three distinct cell populations: basal, intermediate and superficial (umbrella). Umbrella cells secrete an apical membrane plaque composed of uroplakin proteins at the luminal surface. **(B)** Staining with haematoxylin and eosin shows large, binucleated umbrella cells (green arrows), intermediate cells (grey arrows), and basal cells (yellow arrows) at 40x magnification. The white dashed line is the urothelial basement membrane. Adapted from Jackson et al. (2020) [89]

The basal cell layer is the innermost layer and comprises a single-cell stratum that is intimately associated with the underlying capillary bed (Figure 4). Basal cells are mononucleated, ~10 µm in diameter, and are connected to the more superficial intermediate layer via desmosomes to the deeper basement membrane by hemidesmosomes [90]. Pluripotent stem cells likely reside in this layer and are responsible for maintaining normal cell turnover and rapid epithelial regeneration when damaged [91]. However, there is no strong evidence supporting their presence in this layer due to a lack of biochemical markers that are sufficiently specific to discriminate these cells from transit-amplifying cells, which arise from stem cells and contribute to normal turnover [91].

Superficial to the basal cell layer is the intermediate cell layer, which is one to two layers thick in rodents (the primary animal model that is used to study UTI) and up to five strata thick in humans [90] (Figure 4). Intermediate cells are usually pyriform (pear-shaped), mononucleated, ~10-15 μm in diameter, and are joined to each other and the superficial umbrella layer via desmosomes and gap junctions [90] (Figure 4). Intermediate cells in the most superficial stratum are partially differentiated and can rapidly differentiate into replacement umbrella cells when the overlying umbrella cells are lost or damaged [92].

The outermost layer is composed of umbrella/facet cells that are a single layer of highly differentiated and polarised cells [90] (Figure 4). These cells are either mono- or multinucleated, polyhedral, and ~25-250 μm in diameter, and their morphology changing depending on the filling state of the bladder [90] (Figure 4). The most striking feature of these cells is that they are covered by crystalline uroplakin protein plaques, which form a highly impermeable barrier against liquids, toxins, and microbes [93] (Figure 4). In an empty bladder, these plaques are found in cytoplasmic vesicles just below the apical membrane that faces the urine environment [93] (Figure 4). As the bladder fills, these vesicles are trafficked to the apical membrane where they release uroplakins [93]. The uroplakin plaques are integrated into the apical membrane to offset the increased permeability brought about by the distension of the urothelium caused by bladder filling, thereby maintaining the urine-blood barrier [93].

The barrier-like function of the umbrella layer is further supplemented by a mucous-like layer that is composed of glycosaminoglycan (GAG) [94]. This layer is particularly important in preventing bacterial adhesion to the uroepithelium as uroplakin is the binding target for uropathogens expressing type 1 fimbriae [94,95]. *In vitro* bladder models demonstrated that the removal of this layer led to a significant 50-fold increase in *E. coli* adherence [96]. *In vivo* murine studies found that mice with a disrupted mucous layer experienced increased *E. coli* burden in their bladders 4-7 days post-infection compared to control mice [97].

### *1.2.3 Urothelial recognition of uropathogens*

The urothelium is not just a passive barrier, but an active sentinel of the innate immune response as it is able to recognise and respond to the presence of uropathogens in the urinary tract. Pathogen recognition is mediated via Toll-like receptors (TLRs), a family of Pathogen recognition receptors (PRRs) that detect

highly conserved molecular compounds, which are exclusively associated with microbial pathogens. These compounds are termed pathogen-associated molecular patterns (PAMPs) and are detected by different members of the TLR family (Table 4).

**Table 4 Members of the TLR family.**

| TLR Member | Target Ligand | Organism |
|---|---|---|
| TLR1* | Triacyl lipoprotein | Bacteria (Gram -ve) |
| TLR2* | Lipoteichoic acid<br>Peptidoglycan<br>Zymosan<br>Lipomannans<br>Cell-wall B-glucans | Bacteria<br>Fungi |
| TLR3 | Double-stranded RNA | Intracellular bacteria, Viruses |
| TLR4 | Lipopolysaccharide (LPS), Type 1 fimbriae, and Type P fimbriae (TPF) | Bacteria (Gram -ve) |
| TLR5 | Flagellin | Motile bacteria |
| TLR6* | Diacyl lipoprotein | Bacteria (Gram +ve) |
| TLR7 | Single-stranded RNA | Viruses |
| TLR8 | Single-stranded RNA | Viruses |
| TLR9 | Unmethylated CpG-DNA | Viruses, Intracellular bacteria |
| TLR 10 | Unknown | N/A |

TLRs highlighted in green are implicated in the urothelial response. *Unlike the rest of TLR family, TLR1-2 and TLR2-6 form heterodimers instead of homodimers [98]

There are ten TLRs in humans and twelve in mice, with their expression differing depending on anatomical location and cell type [99]. It should be noted that the expression and functionality of the TLRs between human and mice can greatly differ even when comparing similar cell/tissue types [100,101]. This is an important consideration as much of the authoritative literature on UTIs has been derived using murine models despite the fact that mice are immunologically different from humans, especially with regards to the urinary tract. The implication is that there is a substantial difference in the immune response towards uropathogens between human and mice, which will also affect the behaviour of the invading/colonising uropathogens [102]. With regards to UTIs, TLRs 2,4 and 5 are the most relevant members that contribute to the urothelial and immune response [98]. These detect microbial components and virulence factors that are commonly associated with uropathogens [98].

TLRs consist of an extracellular leucine-rich repeat region that is responsible for PAMP recognition and an intracellular cytoplasmic tail containing a Toll-IL-1 receptor (T1R) domain that interacts with intracellular signalling molecules involved in the inflammatory response (Figure 5). Other proteins (co-receptors/accessory proteins) may interact with TLRs, via the leucine-rich repeat (LRR) and transmembrane helix, during ligand binding (Table 5) [103]. Some co-receptors like LPS-binding protein and MD2 are indispensable for TLR function while others including CD36 and CD14 are not as vital but can enhance/modulate TLR binding and response [104] (Table 5). These co-receptors improve the specificity of ligand detection and may even influence the early innate immune response [105] (Table 5)

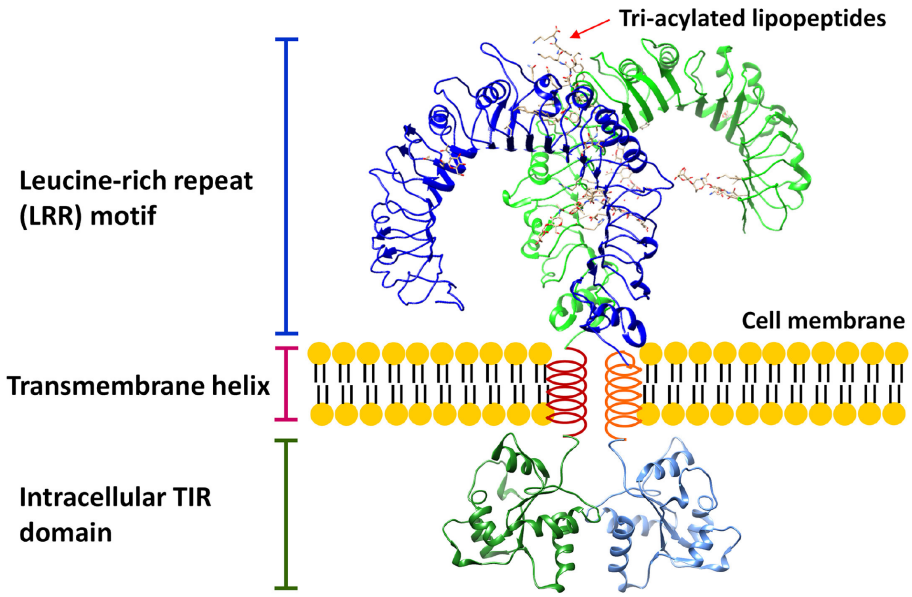



**Figure 5 TLR structure.** All TLRs have the same conserved structural features: (i) leucine-rich repeat (LRR) motif, (ii) transmembrane helix, (iii) intracellular TLR domain. The LRR structure is responsible for the recognition of PAMP molecules, the LRR of the TLR showed in the figure above is based on the model of the TLR1-2 heterodimer interacting with its agonist, a tri-acylated lipopeptide molecule. The transmembrane helix has four functions: (a) dimerisation of TLR molecules, (b) trafficking of TLR molecules to the cell membrane, (c) anchoring of TLR molecules to the cell membrane (d) interaction with appropriate co-receptors. The intracellular Toll-IL-1 receptor (T1R) domain interacts when intracellular signalling molecules involve in the inflammatory response when a TLR dimer binds to its PAMP [106–108]. Figure obtained from Gao et al. 2017.

**Table 5 TLR co-receptors involved UTI innate immunity** [105]

| TLR co-receptor | Associated TLR | Function |
|---|---|---|
| LPS-binding protein (LBP) | TLR4 (direct)<br>TLR2,3,4,7,8,9 (indirect) | LPS delivery to TLR4<br>Lipotechoic acid, peptidoglycan, and lipopeptide delivery to CD14 |
| MD2 | TLR4 | Required for cell-surface expression of TLR4<br>Formation of TLR4-LPS complex |
| CD36 | TLR2,4,6 | Enhances immune response to certain TLR2/6 ligands (lipotechoic acid) |
| CD14 | TLR2,3,4,7,8,9 | Increase the sensitivity of TLRs towards a vast range of ligands<br>Necessary for adaptor protein signalling (TRIF/Myd88) |

TLR activation is induced via ligand binding, which induces receptor dimerisation, and facilitates close contact of the cytoplasmic T1R domains [103] (Figure 6). This results in the recruitment of MyD88/TRIF adaptor proteins, which in turn recruit protein kinases IRAK1/4. The protein kinases associate with TRAF6 resulting in a kinase cascade which culminates in the activation of IKK, which phosphorylates and degrades IkB, the inhibitor of NFkB [109]. NF-kB is the master regulator of inflammation and its activity is constitutively inhibited in homeostatic cells via the NFkB-IkB

complex [109]. The dissociation of NFkB from this complex allows it to enter the nucleus, which promotes the expression of genes encoding inflammatory mediators that activate multiple antimicrobial activities within and outside the urothelium [109] (Figure 6).

The inflammatory mediators include cytokines, histamine, kinins, complement factors, clotting factors, nitric oxide, and proteases. Within the urinary tract, these mediators can irritate the bladder urothelium, causing erythematous swelling and ulceration in more severe infections. These responses are responsible for the main symptoms of UTI: pain during urination, urgency, increase frequency of urination, and dysuria. Inflammatory mediators can also enter into systemic circulation, causing low-grade fever but this is relatively uncommon for individuals presenting with lower UTI. Individuals with upper UTI tend to present with more severe symptoms, which includes high-grade fever, as the renal epithelia can elicit a more robust immune response compared to that of the bladder [110]. Aggressive or poorly managed upper UTIs can develop into life-threatening sepsis as the rich vasculature of the kidneys provides a pathway for uncontained renal infections to spread into systemic circulation.

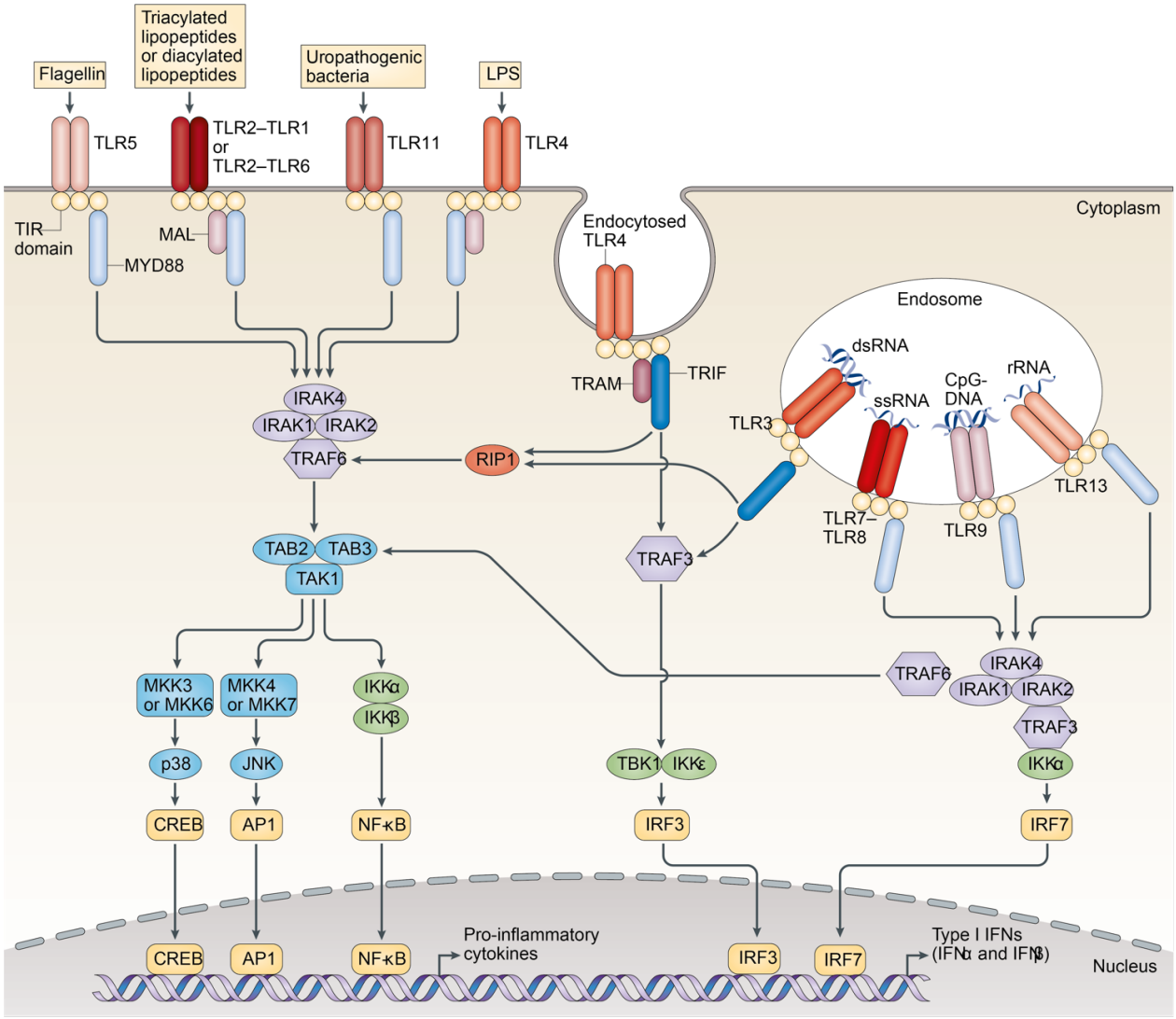



**Figure 6 Signalling pathway of TLRs.** TLR signalling is initiated when an appropriate ligand binds to a receptor, inducing dimerisation. This causes the Toll-IL1-Receptor (T1R) domains to engage the MYD88, MAL, or TRIF-TRAM adapter protein(s) which usually recruit protein kinases IRAK1, IRAK2, IRAK4. For exogenous TLRs (TLR1,2,4,5,6), the IRAK1-IRAK2-IRAK4 complex associates with TRAF6 ubiquitin ligase, activating TRAF6 and causing it to be polyubiquitinated, creating a scaffold to recruit the TAB2-TAB3-TAK1 complex. The scaffold also allows TAK1 to interact with the IRAK complex, thereby activating TAK1. TAK1 interacts with certain mitogen-activated protein kinases (MAPKs) such as p38 and JNK which activate the CREB and AP1 transcription factors that transcribe cytokine genes. Most importantly, TAK1 also activates the IKK complex that is composed of IKKα, IKKβ, and IKKγ (not shown). TAK1 activates IKKβ which then phosphorylates and degrades IkB (not shown), the constitutive inhibitor of NF-kβ. This allows NF-kβ to migrate into the nucleus, where it can promote the transcription of genes for pro-inflammatory cytokines (TNF-α, IL-1β, IL-6, IL-8). For endogenous TLRs (endocytosed TLR4, TLR3,7,8,9), instead of TRAF6, TRAF3 is the main ubiquitin ligase. In the case of TLR3 and endocytosed TLR4, the TRIF adapter protein can interact directly with TRAF3 to recruit the TBK1-IKKε complex that activates the IRF3 transcription factor. The IRF3 response induced by endocytosed TLR4 is relatively weak compared to that induced by TLR3, and its functional role *in vivo* is still unknown. For TLR7,8,9, the Myd88 adapter protein is recruited instead of TRIF. Myd88 recruits the IRAK1-IRAK2-IRAK4 complex that can physically associate with IRF7 via TRAF3 and IKKα. Thea activation of IRF3 and IRF7 induces the production of type 1 interferon (IFNs) such as IFNα and IFNβ that mediate anti-viral responses. Note that TLR11 and TLR13 have only been found in mice. LPS: lipopolysaccharide, dsDNA: double-stranded DNA, dsRNA: double-stranded RNA, ssRNA: single-stranded RNA, TIR: Toll-IL1-Receptor, IRF: interferon regulatory factor, Myd88: myeloid differentiation primary response 88, MAL: Myd88-adaptor-like protein, TRIF: TIR-domain-containing adapter inducing IFN-B, IFN: interferon [103,111]

### *1.2.4 Urothelial response against uropathogens*

Inflammatory mediators comprise of chemokines and cytokines such as Interleukin 1β (IL1β), IL6, IL8, and tumour necrosis factor alpha (TNFα) [103]. Collectively, they promote the early recruitment and activation of innate immune cells into the urinary tract to mediate bacterial clearance. In addition to attracting immune cells to the urinary tract, these mediators also help promote the synthesis and secretion of antimicrobial peptides (AMPs) from the urothelium [112]. AMPs such as β-defensin 1 and cathelicidin (LL-37) are constitutively secreted by the urothelium but their level of expression is greatly enhanced by TLR activation [113].

AMPS such as the defensins and cathelicidin, being cationic and hydrophobic, specifically target the negatively charged outer membrane of bacteria [114]. Their primary antimicrobial mechanism lies in their ability to insert into and disrupt these membranes, essentially, they create membrane pores which kill the bacteria [113]. Host cells are not impacted due to their neutrally charged plasma membrane [114]. At higher concentrations, AMPs can act as chemoattractants for phagocytic cells by interacting with their N-formyl-methionyl peptides (fMLP) receptors [115]. It should be noted that this chemoattractive property is only effective in the micromolar range, which is substantially higher than observed for traditional chemokines such as IL8, which exerts its chemoattractive effects in the picomolar range [112]. While IL8 can function as a systemic chemoattractant, the chemoattractive property of AMPs is limited to the local urothelium [112]. This not only limits the intensity of phagocytic cell recruitment to the urothelium as their presence can be destructive if unregulated, but also focuses their phagocytic activity to regions where there is high bacterial burden, ensuring efficient bacterial clearance while limiting collateral damage to residential epithelial cells [112].

Complementing AMP activities are iron- and siderophore binding proteins such as lactoferrin and lipocalin 2 [116]. The primary function of these proteins is to limit the availability of iron in the urinary tract, making it a more hostile environment for the growth of uropathogens [117]. Iron is essential in many metabolic processes which means that it is a crucial micronutrient for pathogen survival. Lactoferrin is a potent iron chelator which directly limits the availability of iron in the urinary tract [118]. Lipocalin 2 sequesters certain siderophores (iron-chelating compounds used by bacteria), thereby interfering with bacterial iron uptake [119]. These proteins are

produced by both the urothelium and immune cells, with their secretion being upregulated during an active bacterial infection [120].

At a last resort, umbrella cells can be exfoliated to remove adherent bacteria from the urothelium [121]. This is accompanied by the rapid differentiation of partially differentiated intermediate cells that reside directly underneath the umbrella cells, thereby maintaining the urine-blood barrier [122]. However, exfoliation may be a double-edged sword as it creates an opening for certain uropathogens to invade the deeper layers of the urothelium [123]. Indeed, murine studies have demonstrated that *E. coli* is able to induce the apoptosis of umbrella cells so that they can invade and disseminate within the intermediate layer where they form quiescent intracellular reservoirs that are potentially responsible for driving recurrent infections [124,125].

### *1.2.5 Innate immune response against uropathogens*

Neutrophils, recruited by the CXC chemokine 1 (CXCL1) and IL8 chemokines produced by the urothelium, are the first responders to uropathogens in the urinary tract and are primarily responsible for bacterial clearance [126–128]. In murine models, they can be detected in urine as early as two hours post-infection, with their numbers peaking by six hours [129]. Neutrophil recruitment into the urinary tract is directly proportional to bacterial burden, this behaviour is exploited in dipstick tests to assess the severity of an infection (Section 1.1.1) [130]. Bacterial clearance is facilitated by the release of reactive oxygen species and other cytotoxic products [131]. However, these antimicrobial mechanisms are also toxic to the urothelium, which is why excessive neutrophil activity/response can cause urothelial scarring thereby potentially predisposing the bladder to persistent infections [132–134]. To minimise damage, activated neutrophils usually migrate across the urothelium and into the urine environment before carrying out their antimicrobial functions [121].

Macrophages are key to directing the activity of neutrophils and other immune cells in the urinary tract [135]. There exists a substantial population of resident macrophages in the submucosa of the urinary tract, and numbers can be further increased via recruitment from the bloodstream following an infection [136]. During an infection, two main macrophages subsets will coordinate the immune response: the “sentinel” lymphocyte antigen 6 complex negative ($LY6C^-$) macrophage and the “vanguard” lymphocyte antigen 6 complex positive ($LY6C^+$) [135]. $LY6C^-$ are residential to the urinary tract and secrete CXCL1 and macrophage migration inhibitory factor (MIF) to

recruit neutrophils; and CC-chemokine ligand 2 (CCL2) to recruit LY6C$^{+}$ macrophages (Figure 7) [135].

Upon reaching the basal urothelium, LY6C$^{+}$ macrophages then secrete TNF to induce the LY6C$^{-}$ macrophages to secrete CXC chemokine 2 (CXCL2) [135], which triggers the spontaneous production of matrix metalloproteinase 9 (MMP9) among neutrophils, thereby inducing their migration across the urothelium (Figure 7) [137]. The neutrophils are closely escorted by the LY6C$^{+}$ macrophages which regulate their potent antimicrobial activity by keeping them in close proximity before targeting pathogens in the urine environment [135]. This ensures that the urothelium is mostly spared from the cytotoxic effects of the antimicrobial compounds that are released by the neutrophils [138].

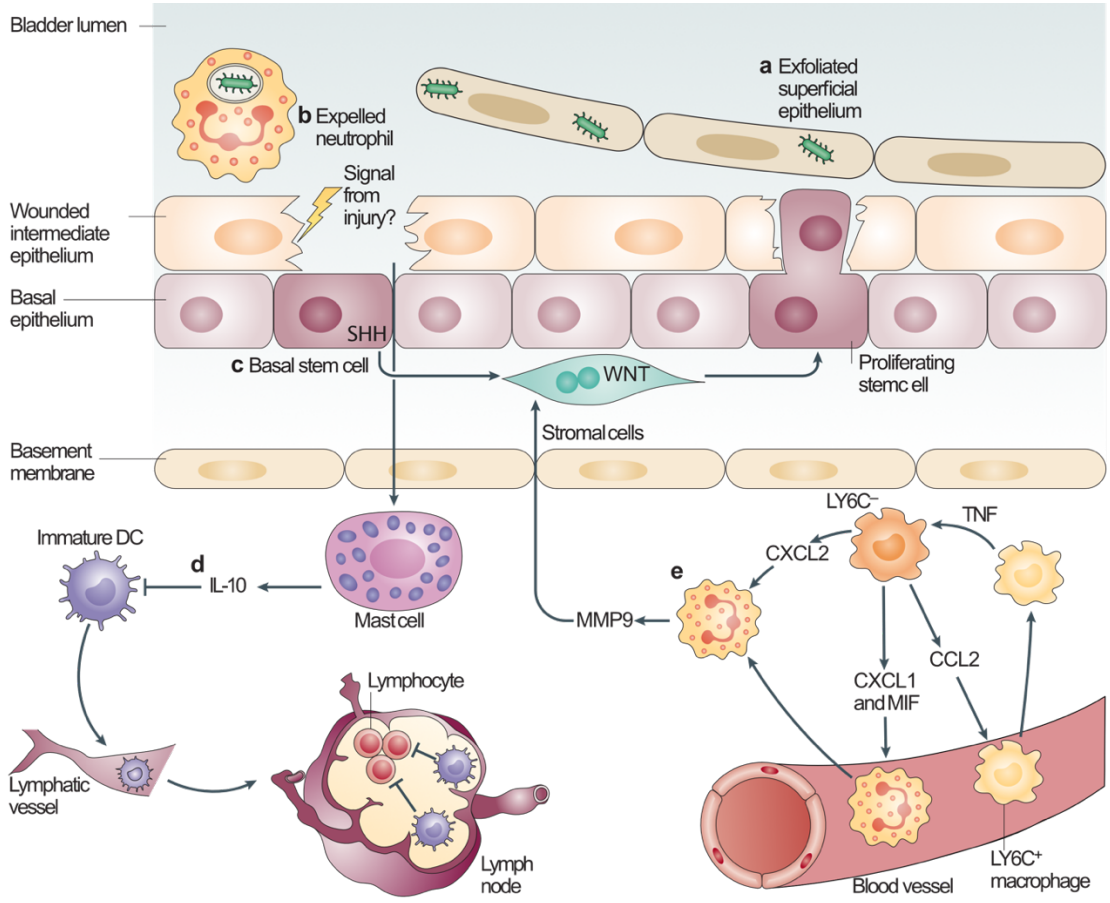



**Figure 7 Innate immune response of the urinary tract. (a)** Exfoliation of bladder urothelial cells can clear adhered uropathogens but it can compromise the urothelial barrier if it is too extensive. **(b)** The action of chemokines promote the migration of neutrophils into the urine, minimising tissue damage from their cytotoxic granules. **(c)** Bladder resident mast cells mediate pro-inflammatory responses during the early stage of an infection but will switch to mediate anti-inflammatory responses when the urothelium is damaged in order to facilitate epithelial regeneration. **(e)** Resident LY6C$^-$ macrophages recruit LY6C$^+$ macrophages and neutrophils from the bloodstream by releasing CC-chemokine ligand 2 (CCL2), chemokine CXC-chemokine ligand 1 (CXCL1) and macrophage inhibitory factor (MIF). LY6C$^+$ macrophages secrete tumour necrosis factor (TNF), which acts on local LY6C$^-$ macrophages to promote the production of CXCL2, which acts on neutrophils to induce the expression of matrix metalloproteinase 9 (MMP9) which promotes their migration across the basal membrane**.** Reproduced from Abraham SN et al. (2015) [121]

Another key resident immune cell are mast cells, which play a pivotal role in directing the innate immune response during an infection. During the early phase of infection, mast cells adopt a pro-inflammatory phenotype, releasing copious amounts of TNF, histamine, and several other chemokines into the urine [139,140]. At a later stage, typically 6-12 hours post-infection, mast cells adopt an anti-inflammatory phenotype, producing IL10 to supress the inflammatory response (Figure 7) [110]. This switch appears to be synchronised to the breakdown of the urothelium and may facilitate its regeneration following bacterial clearance [110]. However, the immunosuppressive activity of mast cells can impede bacterial clearance within the lower urinary tract, especially if these are induced before an infection is completely cleared [110]. For example, a study by Zhu et al. (2020) demonstrated a significant association

between increased IL10 levels and chronic UTIs in mice models [141]. In line with this, a longitudinal clinical study using older adults suggested that UPEC might be able to exploit this mechanism to ensure its persistence in the urinary tract [142]. Importantly, resident mast cells create a tolerogenic environment in the lower urinary tract that prevent or dampen any adaptive immune responses [110].

### *1.2.6 Adaptative immune response against uropathogens*

The limited adaptive immune response of the lower urinary tract probably explains the high recurrence of UTIs following bladder infection [110]. During the later stage of infection, high concentrations of IL10 in the bladder prevent the activation of dendritic cells but still allows for their migration into the nearby lymph nodes[110]. These immature dendritic cells are regulatory and supress the activation of the adaptive immune response [135]. Such a tolerogenic phenomenon is unique to the bladder; a necessary compromise to maintain the integrity of the urothelial lining and prevent aberrant adaptive responses to urine contents [110,143]. However, infections that progress to the kidneys will induce an adaptive immune response as evidenced by the presence of antibodies specific for the infecting organism in patients with pyelonephritis or upper UTI [110].

### *1.2.7 Polymorphisms of the innate immune response and UTI*

The central role of TLRs in initiating the innate immune response explains the growing interest in genetic polymorphisms affecting either their expression and/or function. A skewed inflammatory response because of altered TLR function may very well influence disease progression and outcome [144]. Certain single nucleotide polymorphisms (SNPs) affecting TLR1, TLR2, TLR4 and TLR5 are associated with altered disease outcomes or predispositions in relation to UTIs (Table 6).

**Table 6 A summary of the Toll-like Receptor (TLR) single nucleotide polymorphisms (SNPs) which are implicated in either predisposition or protection to/from urinary tract infections (UTIs).**

| TLR SNP | Functional allele | Non-functional allele |
|---|---|---|
| TLR1 G1805T | TLR1 1805T<br>(Co-dominant, Protective) | TLR1 1805G<br>(Co-dominant, Susceptible) |
| TLR2 G2258A | TLR2 2258G<br>(Dominant, Protective) | TLR2 2258A<br>(Recessive, Susceptible) |
| TLR4 A896G | TLR4 896A<br>(Recessive, Protective*) | TLR4 896G<br>(Dominant, Susceptible*) |
| TLR5 C1174T | TLR5 1174C<br>(Recessive, Protective) | TLR5 1174T<br>(Dominant, Susceptible) |

For the TLR1 G1805T SNP (S602I, rs5743618), the TLR1 1805G allele is associated with a non-functioning TLR1 whereas the TLR1 1805T allele is associated with a functioning TLR1 [145]. The G and T alleles are co-dominant, which means that TLR1 signalling is highest for individuals with the 1805TT genotype followed by 1805GT and 1805GG [146]. The 1805G allele alters the transmembrane domain of the TLR1 receptor (Figure 8) in a manner that impedes trafficking from the Golgi apparatus to the cell membrane, thereby lowering surface expression of TLR1 [147]. Clinically, the 1805TT genotype is associated with protection from pyelonephritis (upper UTI) compared to the GT and GG genotypes [146].

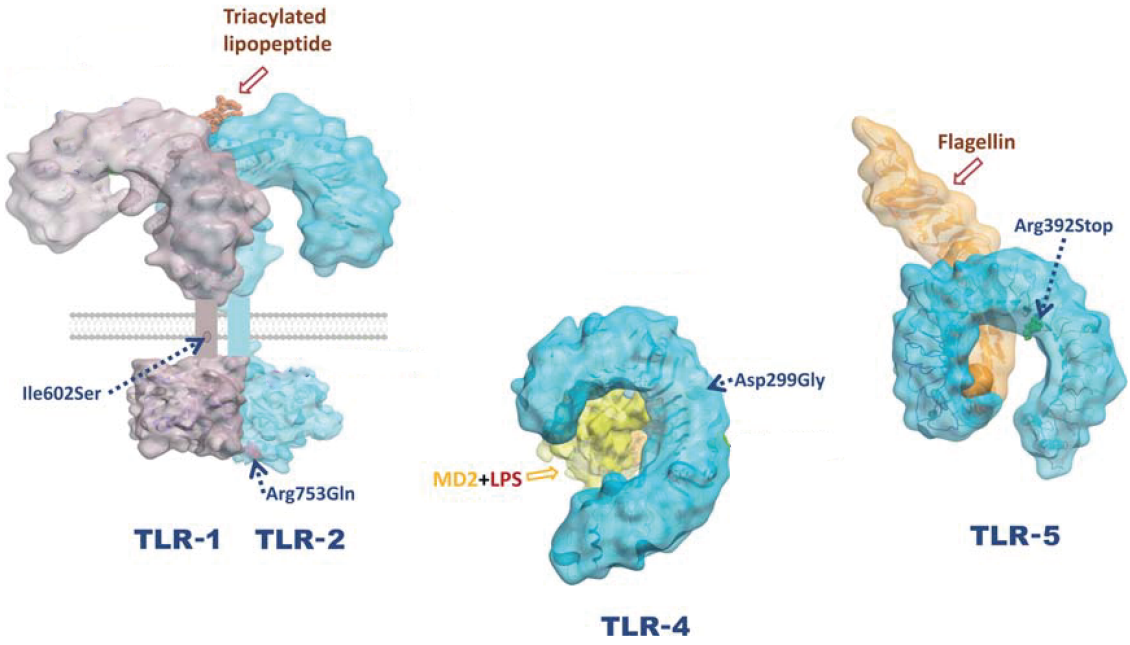



**Figure 8 Three-dimensional structure of Toll-like receptor (TLR) proteins showing the positions of the amino acids which change due to single nucleotide polymorphisms (SNPs) that are relevant to UTI.** For TLR1 and TLR2, the structure of the entire protein is shown whereas for TLR4 and TLR5, only the LRR domain of these proteins are shown. The ligand for each TLR protein is also shown. Figure adapted from Skevaki et al. 2015 [145]

Next is the TLR2 G2258A (R753Q, rs5743708) SNP, with the autosomal dominant TLR2 2258G allele being associated with normal TLR2 function, while the autosomal recessive TLR2 2258A allele is associated with deficient TLR2 function [148]. The TLR2 2258A allele alters the intracellular T1R domain (Figure 8) in a manner that not only interferes with the ability of TLR2 to heterodimerise with TLR1 and TLR6 but also with the recruitment of Myd88 adapter protein, a crucial component of the downstream signalling pathway that activates the inflammatory response, following ligand binding (Figure 6) [145,149]. This was demonstrated in transfected renal-derived HEK293 cells expressing TLR2 2258A which were found to elicit a less potent NFkβ inflammatory response, when measured using a reporter gene, when challenged with TLR2 PAMPs compared to cells expressing TLR2 2258G [150]. The 2258A allele is associated with susceptibility to UTIs in children and ASB in adults [151,152].

The TLR4 A896G (D299G, rs4986790) SNP is the most widely studied TLR SNP owing to the role of TLR4 in detecting LPS, which is a major component of the outer membrane of Gram-negative bacteria, which are associated with many serious infections [153]. The TLR4 896A allele is autosomal recessive and associated with normal TLR4 function whereas the TLR4 896G allele is autosomal dominant and associated with blunted TLR4 signalling as it alters the LRR region (Figure 8) in a way that interferes with ligand binding [145,154,155]. There are many contradictory

findings on the role of this SNP in altering susceptibility to various Gram-negative bacterial infections. There are several reasons for this including, (i) most studies do not take into account other TLR SNPs, even within the same TLR molecule; (ii) there is a lack of standardisation with regards to the usage of *in-vitro* systems, and most do not accurately model primary cell conditions [156]; (iii) use of small sample sizes. The findings generated from *in-vitro* systems can be particularly confusing as the expression of TLR genes can substantially differ between cell lines, even when similar cell types are compared [157,158]. In the context of UTI, one study found a strong association between the TLR4 896G allele and protection from rUTI among Caucasian adult females [159]. In contrast, other studies found a strong association between the TLR4 896G allele and increased susceptibility to lower acute UTIs among East Asian and Hungarian populations [160,161].

The final and perhaps most important SNP with regards to UTI is TLR5 C1174T (R392STOP, rs5744168), with the autosomal recessive TLR5 1174C SNP being associated with normal TLR5 function and the autosomal dominant TLR5 1174T SNP being associated with abrogated TLR5 signalling [162]. The TLR5 1174T allele changes the arginine at amino acid 392 into a stop codon, resulting in a truncated TLR5 protein that lacks the transmembrane and T1R domains [145] (Figure 8). Several studies have demonstrated an association between the TLR5 1174T allele and increased susceptibility to rUTI [159,163]. *In vitro* data generated from proliferating and differentiated normal urothelial cells strongly suggests that TLR5 signalling predominates the bladder urothelial response [158]. Supporting this, murine studies have shown the importance of bacterial motility, which is mediated by flagellin expression, a TLR5 agonist, in UPEC pathogenicity [164]. Wild-type flagellated UPEC regularly outcompeted non-flagellated mutants, with flagellar gene expression coinciding with the ability of UPEC to ascend into the ureter and kidneys [164,165].

## *1.3 Aetiology and pathophysiology of UTI*

### *1.3.1 Microbiological epidemiology of UTI*

UTIs are caused by both Gram-negative and Gram-positive bacteria, as well as by some fungi. The predominant causative organism associated with both uncomplicated and complicated UTI is uropathogenic *Escherichia coli* (UPEC) (Figure 9). For uncomplicated UTI, UPEC is identified in 75% of cases followed by *Klebsiella pneumoniae* (6%), *Staphylococcus saprophyticus* (6%), *Enterococcus spp.*

(5%), *group B Streptococcus* (3%), *Proteus mirabilis* (2%), *Pseudomonas aeruginosa* (1%), *Staphylococcus aureus* (1%), and *Candida spp* (1%) (Figure 9). For complicated UTI, UPEC is similarly common (65%) followed by *Enterococcus spp.* (11%), *K. pneumoniae* (8%), *Candida spp.* (7%), *S. aureus* (3%), *P. mirabilis.* (2%), *P. aeruginosa* (2%). and *GBS* (2%) (Figure 9). The predominance of UPEC as the causative organism of UTI means that most of the available literature concerning the microbiological aspects of UTI is focused on this species.

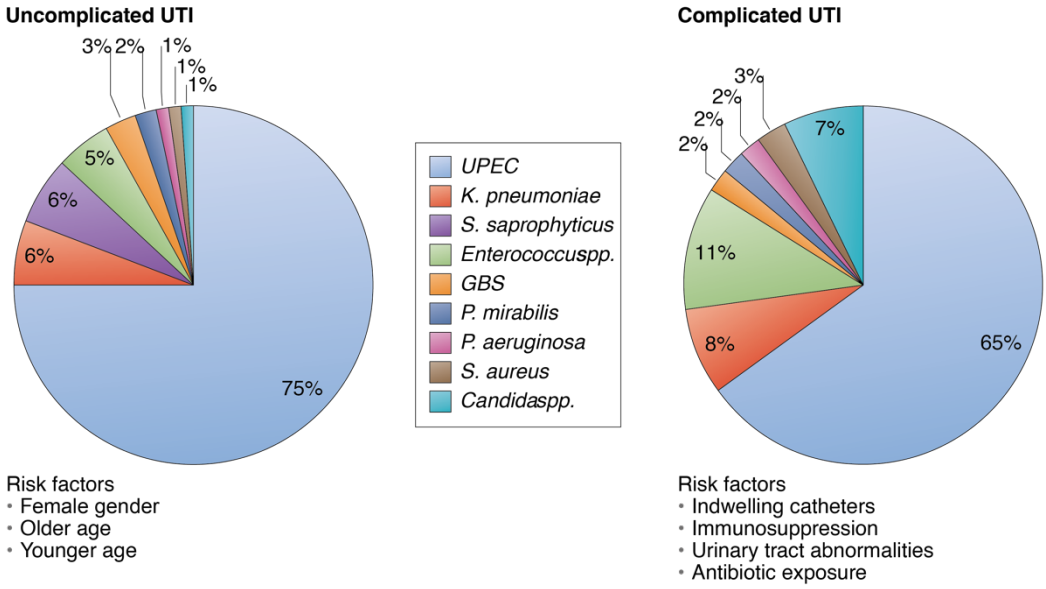



**Figure 9 Microbiological breakdown of urinary tract infections based on prevalence.** *GBS: group B Streptococcus.* Reproduced from Flores-Mireles et al. (2015) [166]

Uropathogens have evolved and acquired numerous virulence factors to subvert key components of the host defence system in order to invade and colonise the urinary tract. However, there is no single nor set of virulence factors that make uropathogens distinct from other pathotypes, especially among *E. coli* isolates. This implies that most uropathogens are opportunistic rather than professional colonisers of the urinary tract. Nevertheless, there are some virulence factors which are associated with an increased probability of successful colonisation.

## *1.4 Escherichia coli*

*Escherichia coli* (*E. coli*) is a diverse bacterial species found naturally in the intestinal tract of all humans and many other mammalian species. Although usually commensal, there are *E. coli* subsets capable of causing enteric diseases and other subsets that cause extra-intestinal diseases, including UTIs. Therefore, *E. coli* isolates can be broadly categorised into three major groups based on their ability to cause disease: (i) commensal or non-pathogenic *E. coli*, (ii) intestinal pathogenic *E.*

*coli*, and (iii) extraintestinal pathogenic *E. coli* (ExPEC) [167]. There are six main "pathotypes" that cause intestinal diseases: enteropathogenic *E. coil* (EPEC), enterohaemorrhagic *E. coil* (EHEC), enterotoxigenic *E. coli* (ETEC), enteroaggregative *E. coli* (EAEC), enteroinvasive *E. coli* (EIEC), and diffusively adherent *E. coli* (DAEC) (Figure 10). Each pathotype is associated with specific virulence factors, which are collectively responsible for their disease-causing ability and the accompanying symptoms (Figure 10).

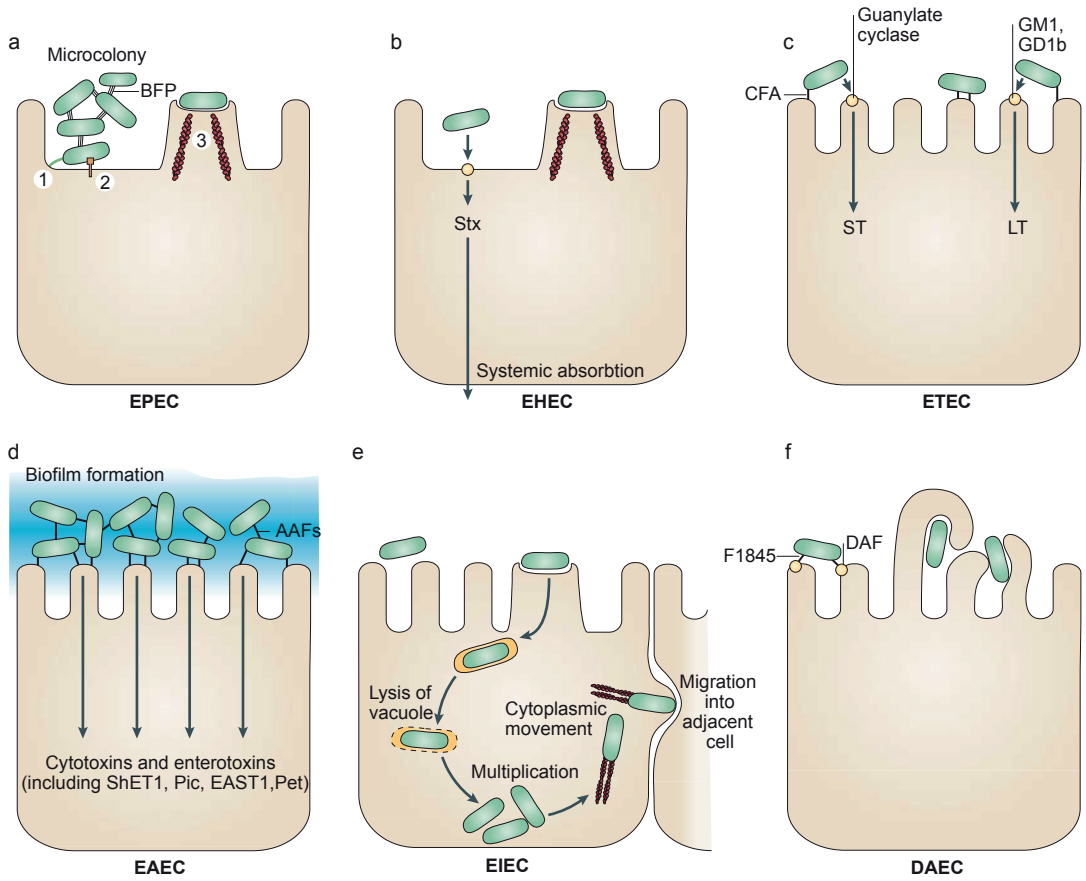



**Figure 10 Pathotypes of intestinal pathogenic *E. coli*.** (a) EPEC adhere to small bowel enterocytes which destroys the normal microvillar architecture and induces the characteristic attaching and effacing lesion. 1. Initial adhesion, 2. Protein translocation by type 3 secretion, 3. Pedestal formation. (b) EHEC also induce a similar attaching and effacing lesion but in the colon. The distinguishing feature of EHEC is the secretion of Shiga toxin (Stx) which can lead to potentially life-threatening complications if it gets into systemic circulation. (c) ETEC adheres to small bowel enterocytes and secretes heat-labile (LT) and/or heat-stable (ST) enterotoxins. (d) EAEC adheres to both small and large bowel epithelia by forming a thick biofilm. It expresses secretory enterotoxins and cytotoxins. (e) EIEC invades the colonic epithelial cells, lyses the phagosome, and moves intracellularly by nucleating actin microfilaments. It can also move into other cells by direct cell-to-cell spread. (f) DAEC induces a signal transduction effect in small bowel enterocytes that results in the growth of long finger-like cellular projections which wrap around the bacteria. AAF: aggregative adherence fimbriae, BFP: bundle forming pilus, CFA: colonisation factor antigen, DAF: decay-accelerating factor, EAST1: enteroaggregative E. coli ST1, LT: heat-labile enterotoxin, ShET1: *Shigella* enterotoxin 1, ST: heat-stable enterotoxin

In contrast, *E. coli* that cause UTIs have yet to be grouped into different pathotypes, instead a general term is used for these *E. coli* isolates: uropathogenic *E. coli* (UPEC), belonging to the wider group of extraintestinal pathogenic *E. coli* (ExPEC)

isolates of which there are other two pathotypes: sepsis-associated *E. coli* (SEPEC) and neonatal meningitis-associated *E. coli* (MNEC). Although there are many virulence factors which are strongly implicated in the establishment of UTIs, there is neither a single nor a combination of virulence factors that can be used to reliably distinguish UPEC isolates from the other listed pathotypes [168]. UPEC isolates are thus highly heterogenous, but phylogenetic analyses of these isolates reveal that most of them originate from a specific phylogroup and appear to be clonal [168].

*E. coli* has a strong phylogenetic structure, and its isolates are usually classified into the following phylogroups (also known as clades): A, B1, B2, C, D, E, F, and G [169] (Figure 11). In the UK/Europe, most UPEC isolates belong to phylogroup B2 and less frequently to phylogroups D, A, and B1 [170]. Phylogroups A and B1 are thought to be associated with commensal extraintestinal strains whereas phylogroup E is associated with the intestinal pathogens. *E. coli* isolates from phylogroup B2 are significantly associated with rUTIs and tend to possess a wider repertoire of virulence factors [171] compared to the other phylogroups (D, A, B1) [170]. However, the carriage of specific virulence factors or a combination thereof is not associated with any particular phylogroup nor clonal group [170]. B2 isolates are also more likely to persist in the urinary tract, causing long-term infections or relapses. On the other hand, D isolates are associated with a higher likelihood of cure or reinfection with a different isolate, implying that D isolates are not able to colonise the urinary tract as robustly compared to B2 isolates [170]. However, B2 isolates were found to be more susceptible to antimicrobials compared to isolates belonging to phylogroups A and D. Indeed, A and D isolates have a higher carriage of antibiotic resistance genes due to the greater incidence of IncH and IncI plasmids in these groups [170]. Interestingly, the higher carriage of antibiotic resistance genes among A and D isolates was not associated with any significant changes in disease outcome [170].

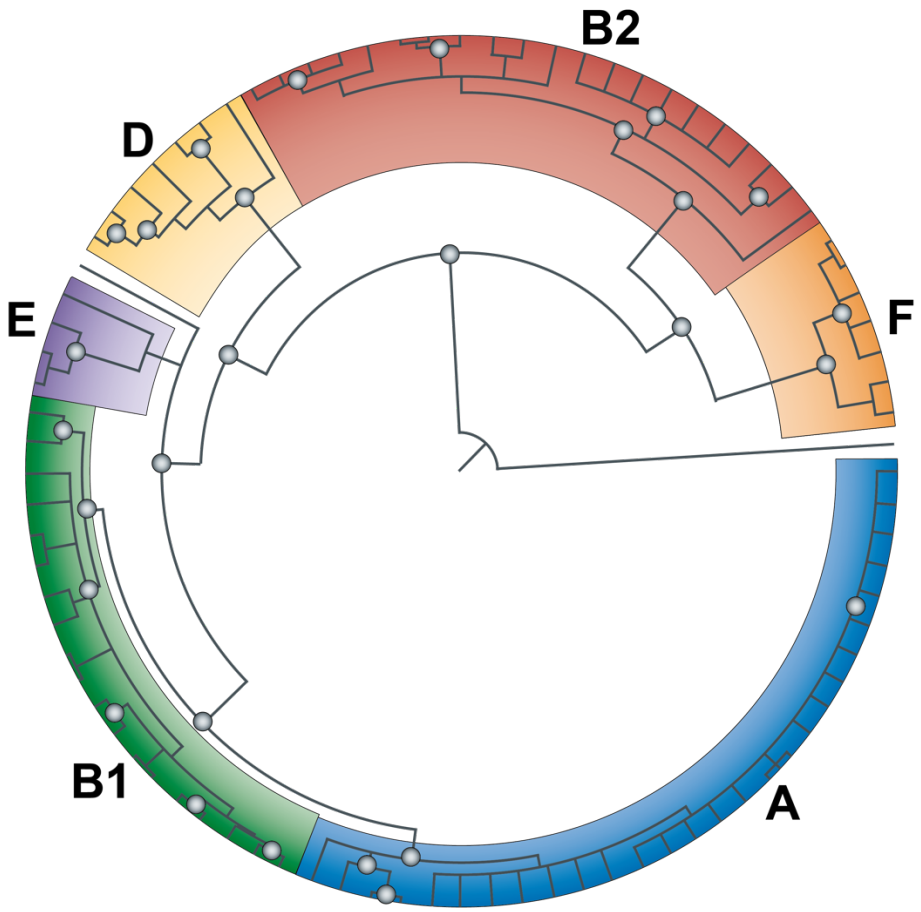



**Figure 11 The major phylogroups of *E. coli*.** The phylogenetic tree was generated using 72 strains from the *Escherichia coli* reference collection (ECOR) and 15 gnome reference strains. The colours indicate the six main phylogroups: A, B1, B2, D, E, and F. Phylogroups C and G are not shown due to the limited phylogenetic resolution of this tree. Adapted from Tenaillon et al. (2010) [172]

*E. coli* has a clonal population structure, meaning that there is strong non-random association of alleles and that most recovered isolates belong to only a few of all possible genotypes [173]. Nevertheless, *E. coli* is highly heterogenous and this characteristic is driven by the high level of homologous recombination (gene conversion) within its chromosome, which occurs as frequently as mutation [173]. The average fragment length involved in these recombination events is estimated to be between 50bp and 2-4kbp [173]. Although these recombination events are scattered over the chromosome, there are two “hotspots” located at the O-antigen biosynthesis gene cluster and the *fim* operon [173]. These recombination “hotspots” are also integration “hotspots” which means that homologous recombination can involve the acquisition of large DNA fragments [173]. Interestingly, recombination does not degrade the phylogenetic signal (owing to the small length of the recombined fragment), which facilitates the reconstruction of a meaningful phylogeny [173].

A clone is composed of almost indistinguishable, or very closely related isolates that originate from a common ancestor. Currently, the most commonly used method to

define *E. coli* strains is multi-locus sequence typing (MLST), with a clone being represented by a unique sequence type (ST). However, EHEC researchers and clinicians remain partial to the serotype designations derived from the somatic (O-serogroup, lipopolysaccharide component), capsular (K, exopolysaccharide component), and flagellar (H) antigens, which correspond to the recombination hotspots of *E. coli*. It should be noted that strains of a single sequence type can still differ substantially in terms of gene repertoire due to the high recombination rate in *E. coli* [173]. In the UK, UPEC isolates are predominated by the following clonal groups in decreasing order: ST73, ST131, ST69, ST95, ST10, and ST127 [174].

Clonal group ST131 is of particular interest as the associated isolates are often highly virulent and resistant to multiple antibiotic classes. In fact, ST131 is the predominant *E. coli* lineage amongst ExPEC isolates and has managed to achieve global prevalence presumably due to heavy use of fluroquinolones and extended spectrum cephalosporins in the 1980s [175]. Although the core genome of ST131 is remarkably stable, its dynamic accessory genome is characterised by frequent recombination events involving mobile genetic elements that carry AMR or virulence gene, allowing it to adapt to a variety of ecological niches [176].

This genetic diversity amongst clinical ST131 isolates is maintained by the effects of negative frequency dependent selection (NFDS), which prevents a single successful isolate from becoming too dominant within the ST131 population [177]. This maintains a diverse repertoire of genes within the ST131 pangenome, increasing the probability of successful adaptation even in the face of drastic ecological changes, i.e. widespread antibiotic usage [177]. The emergence of multi-drug resistant (MDR) ST131 isolates is primarily driven by the diversity of colonisation factors within ST131, which facilitate persistent colonisation of different niches within a human host. Persistent colonisation is strongly associated with the acquisition of antibiotic resistance as the fitness advantage gained from resistance increases with the duration of colonisation [178]. Therefore, the robust genetic diversity of ST131 and its ability to persistently colonise human hosts; combined with the past 60 years of widespread antibiotic usage and mass human travel, has led to the emergence, selection, and global dissemination of MDR ST131 isolates[175].

### *1.4.1 Pathogenesis of E. coli colonisation*

The gastrointestinal tract is widely accepted as a reservoir for most uropathogenic bacteria, and it is believed that healthy humans have a reservoir of UPEC strains mostly belonging to phylogroup B2 and to a lesser extent, phylogroup D [179,180]. Most uropathogenic strains possess a repertoire of virulence factors encoded in pathogenicity islands that bestows the ability to infect an immunocompetent host [181]. Nevertheless, the transfer of uropathogens from the perianal to the periurethral area is usually the result of unhygienic practices [78]. Although it has been suggested that gut-dwelling bacteria may intrinsically migrate from the perianal to the periurethral region via the perineal region [79], existing evidence do not unanimously support this behaviour [80]. Case in point, a clinical study demonstrated that the application of antimicrobials to the perineum failed to reduce the risk of rUTI [81].

Colonisation of the periurethral area and urethral ascension allow bacteria to infect the bladder (Figure 12). The bladder environment selects for the expression of type 1 fimbriae which facilitates the attachment of *E. coli* to the high-mannose type *N*-glycans of uroplakin plaques present on the luminal surface of umbrella cells (Figure 12) [182]. This attachment prevents *E. coli* isolates from being flushed out by micturition, allowing them to persist in the bladder environment (Figure 13A). However, this is not always a benign process for the bladder epithelium as murine models have demonstrated that some strains can cause apoptosis and exfoliation while others can invade it to form intracellular bacterial colonies (IBC) (Figure 12, Figure 13A).

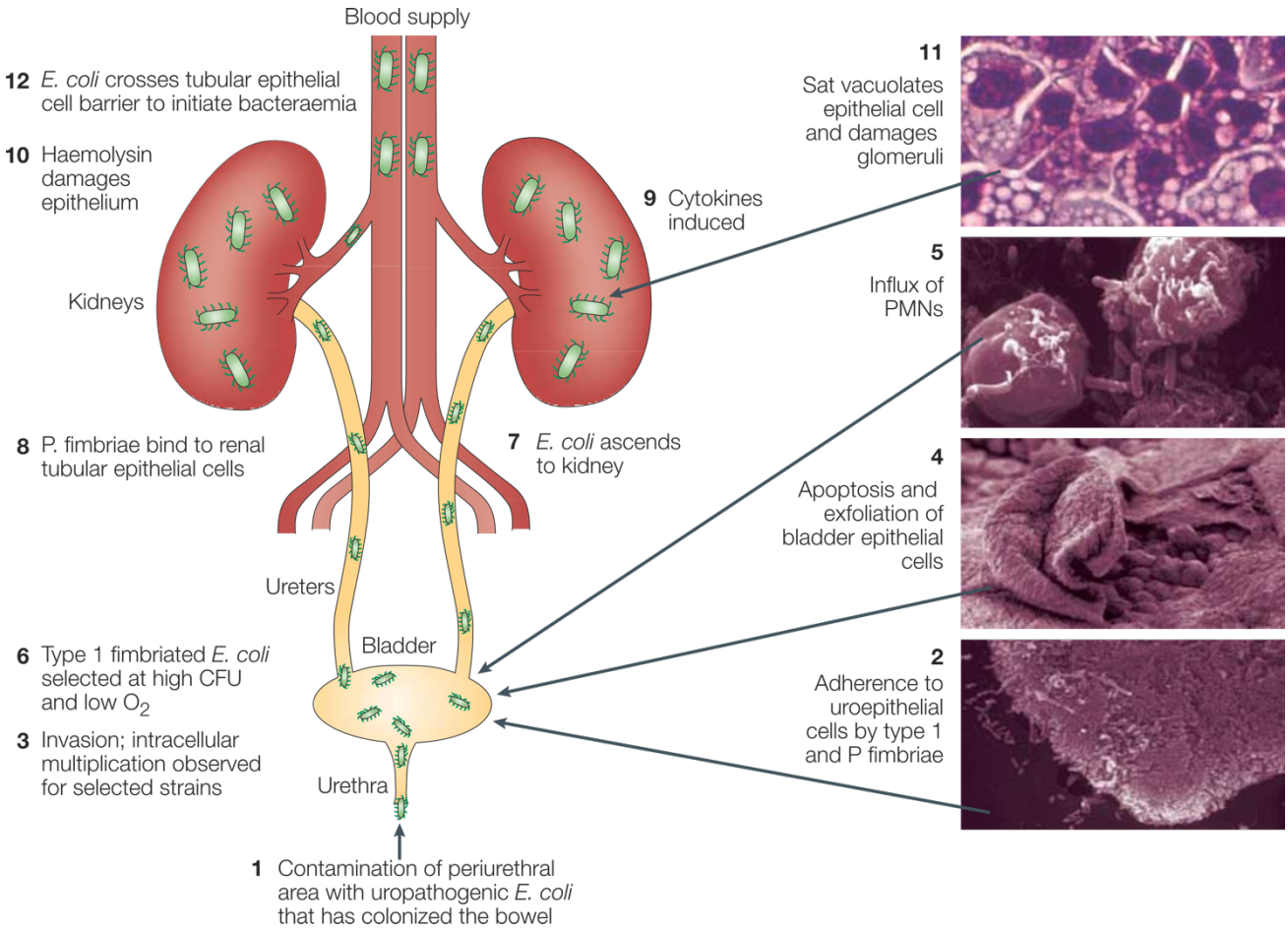



**Figure 12 Pathogenesis of UTI caused by UPEC.** This figure shows the different stages of a UTI. CFU, colony-forming units; PMN, polymorphonuclear leukocytes. Taken from Kaper et al. 2004 [181]

The intracellular location of IBCs makes them resistant to antibiotic treatment and host clearance, providing a safe space for bacterial proliferation. Eventually, the bacterial population within the IBC reaches a certain threshold, causing the bacteria contained within it to detach and burst out into the bladder lumen (Figure 13A). The escaped bacteria then reattach to the epithelium and start another round of IBC formation (Figure 13A). Multiple IBC cycles occur during an infection, which has the effect of stripping the superficial urothelium in some parts of the bladder, thereby exposing the underlying intermediate layers (Figure 13A). UPEC isolates which bind to these intermediate layers will not form IBCs but will instead form quiescent intracellular reservoirs (QIRs), consisting of 4-10 inactive bacteria within membrane-bound vesicles that can remain viable for months. The reactivation of bacteria within these QIRs is hypothesised to cause recurrent UTIs (Figure 13A). Interestingly, the findings that led to the development of the IBC and QIR theories were mostly derived from murine studies or human cell models that do not accurately reflect the human urothelium [183,184]. Indeed, the IBC lifecycle has yet to be satisfactorily demonstrated within human subjects *in-vivo* [183], leaving extracellular reservoirs (vaginal

colonisation, periurethral colonisation, urinary stones) as the predominant source of recurrent infections in humans [71,185–189].

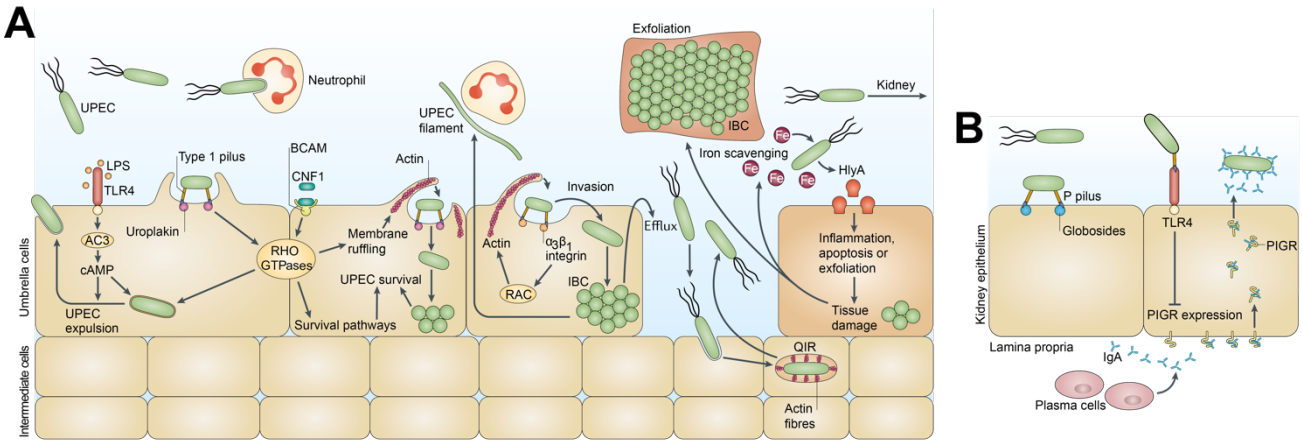



**Figure 13 Pathogenesis of *E. coli* colonisation.** (A) In the bladder, uropathogenic *E. coli* (UPEC) expresses type 1 fimbriae/pilus. The adhesion located at the tip of the pilus (FimH) binds to the mannosylated uroplakins and integrins that cover the surface of umbrella cells. In murine cells, uroplakin binding by FimH induces actin rearrangement via activation of RHO-family GTPases, resulting in bacterial invasion. Once inside the host cell, UPEC can evade host defences and antibiotic treatment. However, lipopolysaccharide (LPS) from UPEC is sensed by TLR4, inducing the production of cyclic AMP (cAMP) via the activation of adenylyl cyclase 3 (AC3). This causes the expulsion of UPEC that are still in vesicles. UPEC is thought to subvert this mechanism by escaping into the cytoplasm where it forms intracellular bacterial colonies (IBCs). Maturation of IBCs causes bacterial dispersal, which frees the bacteria within it to invade other host cells, enabling the IBC cycle to perpetuate. Alternatively, UPEC can establish quiescent intracellular reservoirs (QIRs) by invading the freshly exposed intermediate cells. QIRs consist of 4-10 non-replicating bacteria within membrane-bound vesicles that can remain viable for months. In addition to the proposed IBC and QIR cycles, UPEC possess other pathogenic mechanism to survive in the harsh bladder environment. UPEC may secrete α-haemolysin (HylA), a cytotoxin that induces the lysis of urothelial cells via pore formation, which releases iron and nutrient for UPEC to utilise. Siderophores expressed by UPEC allow the bacteria to scavenge iron and thus survive in the iron-poor environment of urine. Cytotoxic necrotising factor 1 (CNF1) remodels the urothelium by inducing constitutive activation of RHO GTPasesRAC1, RHOA and cell division control 42 (CDC42). (B) UPEC isolates that reach the kidneys depend on the expression of pyelonephritis-associated (P) fimbriae/pili, which binds to the globoside-containing glycolipids lining the renal tissue. The adhesin at the end of the P fimbriae (PapG) can also interact with TLR4, to reduce the expression of polymeric immunoglobulin receptor (PIGR), impairing the transport of immunoglobulin A (IgA) across the renal epithelium, slowing down bacterial clearance. Figure reproduced from Flores-Mireles et al. 2015 [166]

In strains that are associated with cystitis (lower UTI), there is continuous expression of type 1 fimbriae, which confines the infection to the bladder [190]. Strains which cause pyelonephritis (upper UTI) usually possess an invertible element that controls the expression of type 1 fimbriae [191]. At some point in the infection, the invertible element turns to the “off” position, thereby reducing the expression of type 1 fimbriae [191]. This frees the *E. coli* isolates from the bladder epithelium, allowing them to ascend via the ureters to the kidneys (Figure 12). In the kidneys, the *E. coli* isolates attach to the renal epithelium by expressing P fimbriae which attaches to glycosphingolipid receptors (Figure 13B) [192]. Certain virulence factors like haemolysin can damage the

renal epithelium, allowing the bacteria to breach the renal epithelium and enter the bloodstream, leading to bacteraemia (Figure 12).

### *1.4.2 UPEC virulence factors that contribute to UTIs*

Fimbriae or pili are short filamentous organelles that facilitate adhesion to various surfaces. UPEC strains generally have a higher carriage of fimbrial gene clusters compared to intestinal or commensal strains. Type 1 and P fimbriae are the among the best characterised variants in the context of UTI. Type 1 fimbrial genes are found in most *E. coli* isolates whereas P fimbrial genes are exclusively limited to pathogenicity islands distributed among UPEC strains [193,194].

Both type 1 and P fimbriae are structurally composed of a rod and a tip section (Figure 14A,B). The fimbrial tip contains an adhesion protein (FimH for type 1 fimbraie and PapG for P fimbriae) that mediates binding to the host surface. The adhesion protein is connected to the fimbrial rod via linker subunits in a strictly defined order. For type 1 fimbriae, one copy of the FimH subunit is connected to one copy of FimG followed by one copy of FimF, which is then connected to the FimA subunits that make up the Type 1 fimbria rod (Figure 14A) [195]. For P fimbria, one copy of the PapG subunit is connected to one copy of PapF, followed by 5-10 copies of PapE, and then one copy of PapK, which is connected to the PapA subunits that make up the P fimbria rod (Figure 14B) [195]. The fimbrial structure is constructed via a chaperone-usher system, consisting of a periplasmic chaperone (FimC/PapD) and an outer membrane usher protein (FimD/PapC) (Figure 14) [195].

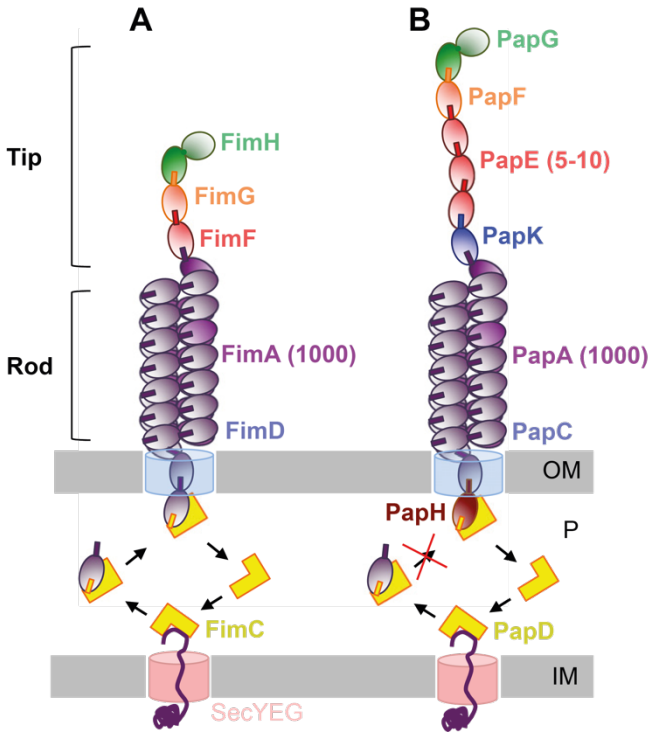



**Figure 14 Structure of the Type 1 (A) and P (B) fimbria.** This figure shows the arrangement of subunits that make up the type 1 and P fimbriae. Subunits are located to the periplasm (P) via the Sec general secretory pathway across the inner membrane (IM). The subunits are folded by an enzyme (disulphide oxidoreductase, DsbA) and transported to the usher protein by a chaperone (FimC/PapD). At the usher protein, the subunit is transferred to the periplasmic N-terminal domain of the usher and incorporated into the fimbria via donor-strand exchange. Once the fimbria has been fully constructed, the fimbria is capped at the base by a termination subunit (FimI for Type 1 fimbria and PapH for P fimbria, note that FimI is not shown as its existence is yet to be confirmed). Figure reproduced from Lillington et al. 2014. [195]

The FimH tip of type 1 fimbria binds to mannosylated residues present on uroplakin plaques, which cover the luminal surface of umbrella cells that comprise the most superficial layer of the urothelium. This binding prevents UPEC from being flushed out by high shear forces generated during micturition, allowing them to ascend the urethra and into the bladder [196]. FimH is also associated with IBC formation in mice as it facilitates UPEC invasion of bladder urothelial cells [197]. In the kidneys, murine studies demonstrated that type 1 fimbria promote interbacterial binding, which results in the formation of biofilm-like communities within renal tubules, which can cause obstructions [198].

The PapG tip of P fimbria binds to glycosphingolipids found in the renal epithelium, specifically the P blood-group antigen [195]. P fimbria also regulate bacterial motility via PapX, a non-structural gene found in the P fimbrial gene cluster [199]. PapX is a crucial component in the crosstalk between fimbria-mediated adherence and flagella-mediated motility among UPEC isolates [199]. Overexpression of PapX supresses

flagellar synthesis by binding to the promoter sequence of *flhDC,* the master regulator of flagellar expression, thereby inhibiting *flhDC* expression [199].

The pathogenesis of an UTI, which necessitates urethral ascension supported the hypothesis that flagellar expression and hence motility is a critical virulence factor for UPEC and other uropathogens. This organelle has been extremely well studied, as evidenced by the plethora of quality papers discussing its assembly, structure, mechanism, and regulation [200]. In the context of UTI, the most important aspects of this organelle are its regulation and structure.

The bacterial flagellum is a long helical surface organelle that is composed of polymerised subunits of flagellin encoded by *fliC* [200]*.* TLR5 recognises a highly conserved amino acid motif (amino acids 89-96) within flagellin that is crucial for the formation of flagellar filament and motility [201]. The recognition of flagellin by TLR5 is permissive to variations in amino acid content within the recognition motif, allowing TLR5 to recognise a broad range of flagellin molecules [202]. Moreover, most point mutations within this recognition motif substantially reduce bacterial motility. The structure requirements for flagellar motility are more rigid compared to that required for TLR5 recognition. This asymmetry might explain why most bacterial species, especially *E. coli*, have failed to evade TLR5 recognition while maintaining adequate flagellar-mediated motility.

However, UPEC can avoid TLR5 recognition by downregulating flagellar expression while in the bladder environment [203]. Interestingly, the relationship between the expression of type 1 fimbriae and motility is unidirectional as flagellar expression does not have any effects on fimbrial expression, whereas fimbrial expression dramatically affects flagellar expression and hence motility. This regulation of motility by the type 1 fimbrial system is not limited to UPEC isolates as constitutive expression of type 1 fimbriae in a laboratory *E. coli* K-12 strain also resulted in decreased motility [204].

Several studies have demonstrated that flagellar expression confers a significant fitness advantage in murine models [205]. Moreover, transient flagellar expression might be necessary at specific stages of an UPEC infection to facilitate efficient colonisation of the urinary tract, especially when an isolate competing with other uropathogens for the limited nutrients present in urine or when ascending to the upper urinary tract [165,206]. However, flagellum-mediated motility is not mandatory for

successful colonisation and is potentially detrimental for persistent colonisation as its component protein (flagellin, FliC) is a potent activator of the bladder innate immune response mediated through TLR5 signalling [207]. Since most UPEC isolates display downregulated flagellar expression, it suggests that any potential competitive advantage that could be derived from flagellar-mediated motility is heavily offset by the potent immune response, and hence bacterial clearance that would be induced by its expression [208].

## 1.5 *Proteus mirabilis*

*Proteus mirabilis* is well known to microbiologists as the bacterial species that swarms across agar surfaces in a striking bulls'-eye pattern, overtaking any other species present [209]. This Gram-negative rod belongs to the same bacterial order (Enterobacteriaceae) as *E. coli*, to the family of Morganellaceae, and of the Proteus genus [210]. There are six species within this genus: *Proteus mirabilis, Proteus vulgaris, Proteus penneri, Proteus cibarius, Proteus terrae*, and *Proteus hauseri* [210]. In the clinical setting, *P. mirabilis* is the most commonly encountered species, mainly in the context of UTIs, particularly catheter associated UTI (CAUTI) [210].

*P. mirabilis* is capable of causing symptomatic infections in both the lower and upper urinary tract. It is also present in cases of asymptomatic bacteriuria (ASB), especially among the elderly and individuals with type 2 diabetes [211,212]. Furthermore, long-term *P. mirabilis* infections can result in the formation of urinary stones due to its ability to produce urease. *P. mirabilis* can be often isolated from the gastrointestinal tract, but it is unknown whether it is a commensal, a pathogen or a transient organism. The majority of UTI cases that are associated with *P. mirabilis* are thought to result from its ascension from the gastrointestinal into the urinary tract. However, person-to-person transmission may also be a major contributor, especially for catheterised individuals in healthcare settings [213].

*P. mirabilis* causes between 1-10% of all UTI cases, varying depending on the geographical location of the study, the type of samples collected, and the characteristics of the patients examined [214]. This species is more frequently encountered among patients presenting with complicated UTIs, particularly patients with either spinal cord injury and/or anatomical abnormality of the urinary tract, and is a major contributor to CAUTI, causing around 10-44% of long-term CAUTIs in the United States [215]. *P. mirabilis* is also frequently encountered in Gram-negative

bacteraemia (5-20% incidence), especially in patients with concurrent UTI [216]. Geriatric patients presenting with *P. mirabilis* bacteraemia are at a particularly high risk of developing life-threatening urosepsis, with mortality rates being as high as 50% [217].

### *1.5.1 Pathogenesis of P. mirabilis colonisation*

When an indwelling catheter is inserted, it bypasses many of the host defences against urinary tract infections [218]. Compared to normal urination and flushing of the urethra at regular intervals, the catheter continuously drains the bladder as urine accumulates [218]. This does not provide the volume nor force required to effectively flush the urinary tract [218]. Furthermore, the design of the traditional Foley catheter and its drainage tubing results in urine retention within the bladder (10-100ml), providing a reservoir for bacterial colonisation [219]. Catheter insertion can easily damage the fragile urothelium if inserted improperly, since the urothelium is only 5-7 cell layers thick [220,221]. Even if the catheter is properly inserted, its presence can elicit a robust inflammatory response which may remodel the urothelium and deposit fibrinogen on the catheter surface, making the urothelium and catheter surface more hospitable for bacterial colonisation [222]. Rough surfaces on the catheter can also facilitate the formation of bacterial biofilm as well as causing damage to the ureteral lining during insertion or withdrawal [223].

The presence of an indwelling catheter provides a solid surface that promotes swarming behaviour in *P. mirabilis*. When *P. mirabilis* is grown on a solid surface, it grows in place for a while before differentiating into swarmer cells (swarming). Swarmer cells move forward as a population across the surface for some distance and then reverts to a shorter morphotype called swimmer/vegetative cells (Figure 15). *P. mirabilis* moves across many solid surfaces by repeating this process, resulting in the characteristic bulls' eye pattern when grown on agar media. Swarmer cells are morphologically and phenotypically distinct from swimmer/vegetative cells as they are multinucleated, filamentous (20-80 μm long), and hyperflagellated. Unlike swimmer cells, swarmer cells move together as a group, forming rafts of parallel cells (Figure 15). Although swarming behaviour has been observed for other bacterial species, it is remarkably robust in *P. mirabilis* as swarming will occur on most solid surfaces unless specific inhibitors are added. Amongst catheterised patients, this swarming behaviour allows *P. mirabilis* to rapidly ascend from the periurethral region to the bladder by migrating along the catheter surface. By swarming along the

catheter surface instead of the urothelium that lines that urethra and bladder, *P. mirabilis* can also avoid being detected by host immune receptors including TLR5 and thus minimise the probability of triggering an immune response that would result in its clearance from the urinary tract.

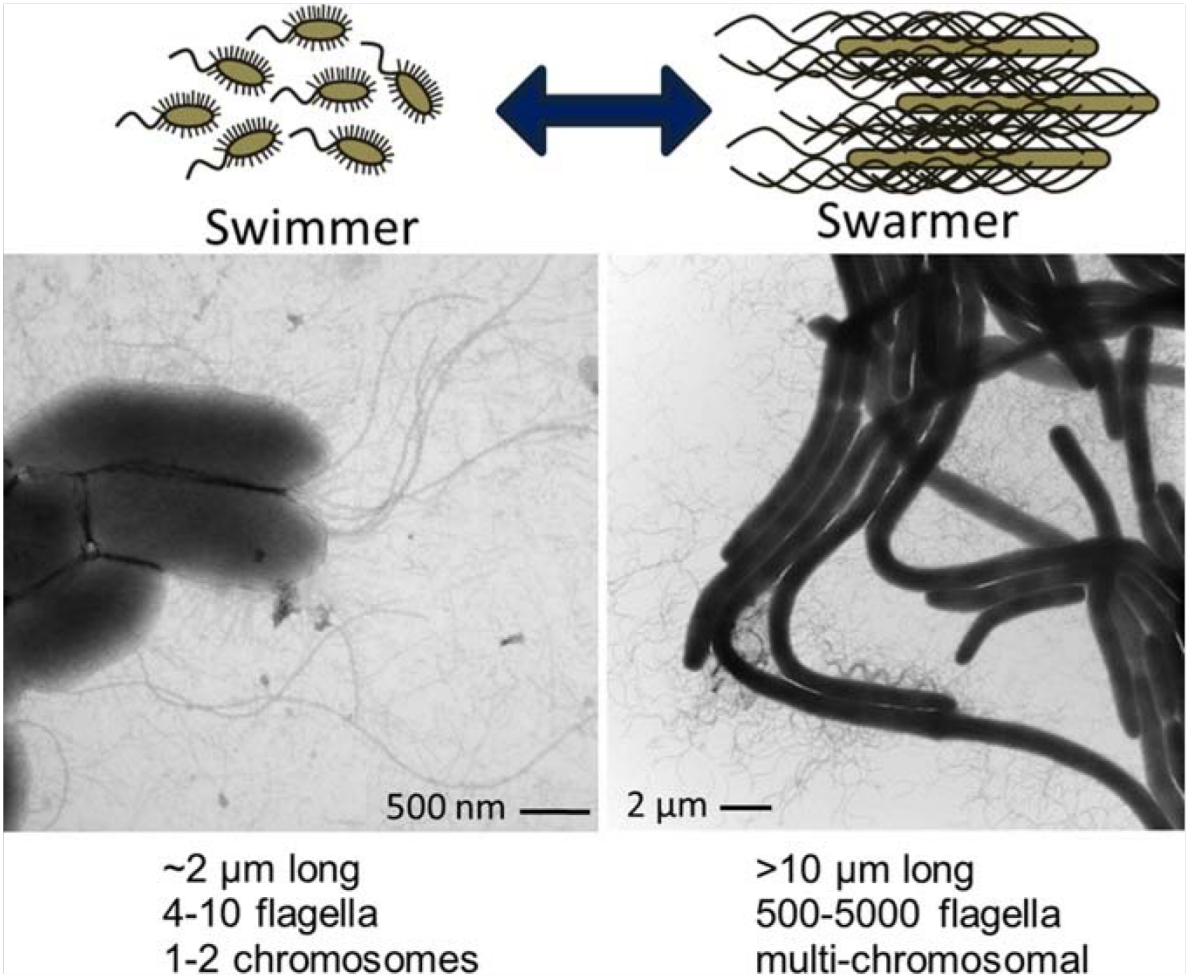



**Figure 15 Swimming and swarming phenotype in *P. mirabilis*.** The left picture shows a transmission electron micrograph (TEM) of swimmer cells. The right picture shows swarmer cells displaying an elongated morphology with an abundance of flagella. Reproduced from Schaffer JN et al. (2013) [209]

*P. mirabilis* can also form a crystalline biofilm on the catheter surface, which is dependent on the expression of mannose-resistant Proteus-like (MR/P) fimbria (Figure 16). MR/P fimbriae are functionally analogous to type 1 fimbriae in *E. coli* as they also mediate attachment to the umbrella cells of the urothelium (Figure 16).

The ability of *P. mirabilis* to form a crystalline biofilm is derived from its ability to produce urease. Urease hydrolyses urea, that is abundant in urine, into ammonia and carbon dioxide. The generation of ammonia increases the local pH, which can lead to the precipitation of calcium and magnesium ions in the urine. This induces the formation of calcium phosphate (apatite) and magnesium ammonium phosphate

(struvite) crystals, which are then trapped in the biofilm via the production of extracellular polymeric substances by bacteria (Figure 16). This crystalline biofilm protects the bacterial community from the immune system and from antibiotics. Over time, expansion of the crystalline biofilm can block the catheter, preventing proper urine drainage. Individuals with long-term catheterisation (> 28 days) tend to experience catheter blockage from crystalline deposits generated from the urease activity of *P. mirabilis*. This can result in vesicoureteral reflux, in which urine from the bladder flows back up towards the ureters and into the kidneys. Vesicoureteral reflux promotes the development of an upper UTI from a lower UTI that can eventually lead to bacteraemia, septicaemia, shock, and death.

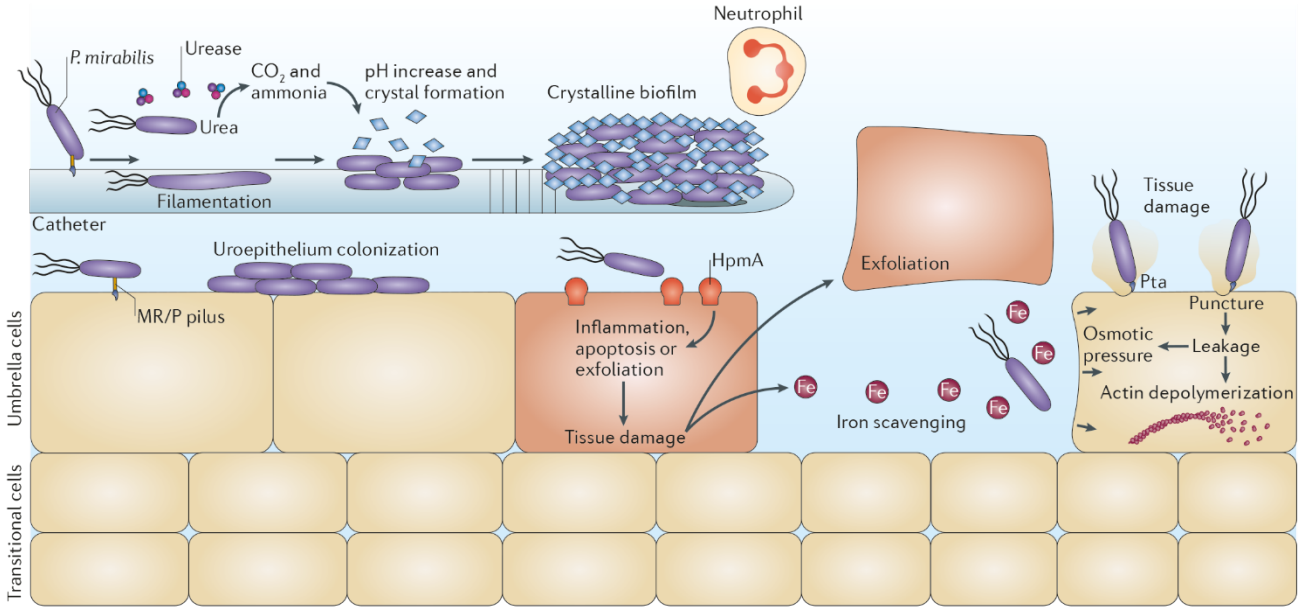



**Figure 16 Pathogenesis of *Proteus mirabilis* colonisation in UTI.** The expression of mannose-resistant Proteus-like (MR/P) fimbria is important for *P. mirabilis* to cause catheter-associated urinary tract infection (CAUTI). MR/P fimbria mediates attachment to the catheter/urothelium and promotes the biofilm formation during the initial stages. Urease production raises urinary pH via the production of ammonia from the hydrolysis of urea. This induces the precipitation of calcium and magnesium ions in the urine, resulting in crystal formation. These crystals can be incorporated into biofilms via the production of extracellular polymeric substances, transforming them into crystalline biofilms. Crystalline biofilms protect the bacterial community within them form the host immune response and from antibiotics. Additionally, these structures impede proper urine drainage, resulting in vesicoureteral reflux in which urine flows from the bladder into the ureters and kidneys, which promotes the colonisation of the upper urinary tract by *P. mirabilis*. This can eventually lead to pyelonephritis, bacteraemia, septicaemia, and shock. P. mirabilis can also produce bacterial toxins such as haemolysin (HpmA) and Proteus toxic agglutin (Pta), which can destroy tissues and promote bacterial dissemination to the kidneys. HpmA inserts itself into the cell membrane, allowing nutrient to leak out but this also destabilise the host cell, causing tissue damage and exfoliation. Pta creates a puncture in the host plasma membrane, which causes the cytosol to leak out. This depolymerises the actin filaments and generates osmotic stress, thus compromising the structural integrity of the cell. Reproduced from Flores-Mireles AL et al. (2015). [166]

*P. mirabilis* is the only bacterial species positively associated with catheter obstruction and this is probably linked to its remarkably vigorous urease activity [224]. In addition to being involved in the formation of crystalline biofilms, the highly efficient urease enzyme is linked to the formation of urinary stones [225]. These stones can impede urinary flow and cause tissue damage, especially during extended periods of *P. mirabilis* colonisation [226]. For *P. mirabilis* and other uropathogens, these stones may provide refuge against antibiotics and the immune system[227]. Furthermore, the bacteria embedded within these stones can constitute an extracellular reservoir of infection, allowing long-term persistence in the urinary tract [228].

The indirect role of urease in contributing to the persistence of *P. mirabilis* colonisation via the generation of urinary stones has been highlighted through murine studies. Mice infected with urease mutant strains had significantly lower bacterial counts in their urine, bladder, and kidneys compared to mice infected with wild-type

strains [229]. Urease production also appears to be a fitness factor; when the urease knockout mutant and wild-type parental strains were co-challenged in mice, the mutant strain was mostly unrecoverable [230]. However, urease expression by itself is not necessary for colonisation as the urease mutants are still capable of bladder colonisation albeit with lesser efficacy compared to their parent strain [231,232].

### *1.5.2 P. mirabilis virulence factors that contribute to UTI*

Like UPEC, *P. mirabilis* is flagellated, bestowing it with motility and chemotaxis. Unlike UPEC, *P. mirabilis* encodes two flagellins, FlaA and FlaB. FlaA is the predominant variant while FlaB is normally silent [233]. Recombination can occur between the homologous regions of *flaA* and *flaB,* resulting in the expression of hybrid FlaAB flagellin [233]. The recombinant flagellin is functional and antigenically distinct from FlaA, which could serve as a method of evading the immune response during UTI [233]. However, it remains unknown whether there are any differences in the TLR5 response between FlaA and FlaAB, especially given the versatility of TLR5 at recognising a wide range of bacterial flagellins [234].

Flagellar expression is crucial for swarming motility in *P. mirabilis*. During the initial transition from the swimming to swarming phenotype, the transcription of flagella is greatly increased and remains high throughout the swarming phase (Figure 15) [209]. Flagellar expression is also vital for the swarming differentiation process as non-motile mutants do not swarm, fail to elongate into a filamentous morphology and do not undergo polyploidy to become multinucleated cells [235]. This is because swarming is dependent on the flagella acting as a sensor to detect any changes in flagellar rotation. Solid or viscous surfaces are sensed by restricted flagellar rotation, which is believed to transmit signals that transcribe genes associated with swarming [236].

*In vitro* studies have demonstrated that *P. mirabilis* is able to use its flagella to effectively invade cultured cells derived from the urinary tract [237]. The findings suggest that the invasiveness of *P. mirabilis* is comparable to *Salmonella typhimurium* or *Salmonella typhi,* which are highly invasive bacterial species. Cellular invasion could be responsible for driving bacteraemia once the infection reaches the kidneys or facilitating the formation of intracellular bacterial communities during bladder/kidney infections. However, cultured cells such as EJ/28 are fundamentally different from primary tissue and the ability of *P. mirabilis* to invade urothelial/renal cells has yet to be proven from clinical samples [237,238].

As with *E. coli*, flagellar expression in *P. mirabilis* is also regulated by fimbrial genes. The *mrp* operon that encodes the MR/P fimbria also encodes a regulatory element, MrpJ. MrpJ represses motility by binding to the *flhDC* promoter thereby repressing the downstream target of *flhDC*, i.e. *flaA* that encodes the primary filament subunit of flagella in *P. mirabilis* [239]. MrpJ is crucial for *P. mirabilis* virulence as demonstrated by a murine study in which mice were transurethrally co-challenged with both wild-type and *mrpJ* null mutants [239]. Lower numbers of *mrpJ* mutants were recovered from the bladder and kidney tissues of the infected mice [239]. Hence, the repression of flagellar expression appeared to be crucial in determining the success of both *P. mirabilis* and *E. coli* isolates in invading the urinary tract. As with *E. coli*, data from a murine study by Zunino *et al* suggests that flagellar expression is not a mandatory virulence factor in *P. mirabilis* as non-motile isolates were as infectious as motile isolates in invading the bladder and kidneys [240]. However, the murine model employed in that study was more representative of uncomplicated UTI than CAUTI as indwelling catheters were not used post-inoculation. This could have altered host-microbial dynamics in a manner which diminishes the fitness advantage of flagellar-mediated motility due to the lack of a solid, abiotic surface for *P. mirabilis* to express its swarming phenotype; thereby diminishing the fitness advantage of flagellar-mediated motility and giving the false impression that motility is a dispensable feature in *P. mirabilis* infections. Although this may indeed be the case for the infrequent uncomplicated UTI cases that are caused by *P. mirabilis*, this is not clinically representative of most *P. mirabilis* infections as this species is far more commonly associated with CAUTI.

MR/P fimbria is another virulence factor that is commonly expressed by *P. mirabilis*. Its primary function resides in biofilm formation by mediating fimbria-fimbria interaction between cells. However, MR/P expression is not mandatory for biofilm formation as MR/P- cells are still able to generate a biofilm, albeit at a slower rate [241]. Unlike *E. coli*, colonisation of *P. mirabilis* is limited to the luminal surface of the urothelium and intracellular colonisation has rarely been observed in mice, this behaviour is associated with MR/P expression [241]. MR/P+ cells are usually localised to the intact bladder urothelium while MR/P- cells exclusively colonise the exfoliated areas of the urothelium, specifically the lamina propria [241]. The expression of the MR/P fimbria in wild-type populations is heterogenous, with a strong bias favouring its expression in the urethral and bladder environments [241]. The colonisation

behaviour of *P. mirabilis* overwhelmingly favours extracellular colonisation and is generally restricted to the most superficial layer of the urothelium.

This colonisation pattern combined with the vigorous urease activity of *P. mirabilis* create a highly alkaline extracellular niche that is highly conducive for *P. mirabilis*. Most uropathogens do not cope well in a high pH i.e. alkaline environment and this can be exploited by *P. mirabilis* as a competitive advantage [242]. *In vitro* growth competition assays in which *P. mirabilis* was grown in competition with common uropathogens including *E. coli* and *E. faecalis* demonstrated that *P. mirabilis* not only outcompeted these species, but also killed them outright, presumably a consequence of the alkaline environment resulting from its activity (Boyles, Hall, & Aldridge unpublished). The rise in pH can also promote the formation of urinary stones that is beneficial to *P. mirabilis* colonisation as discussed in the previous section (1.5.1).

To summarise, flagella-mediated motility and crystalline biofilm formation via urease production are central to the pathogenesis of *P. mirabilis* in causing complicated UTI or CAUTI within a catheterised urinary tract [243]. In contrast, motility is dispensable for *E. coli*, which is more predominant amongst non-catheterised individuals [205]. The periodic and powerful flushing action of micturition favours adhesion over motility, making FimH-mediated adhesion central to the pathogenesis of *E. coli* colonisation in causing uncomplicated UTI [244]. In catheterised individuals with indwelling catheters, continuous urine flow diminishes the flushing force of micturition [245], presumably making adhesion less advantageous compared to motility. This explains why *E. coli* and *P. mirabilis* have different host preference and colonisation strategy despite infecting the same anatomical site.

## *1.6 A critique of murine models that predominate existing UTI research and findings*

Almost all of the current understanding underlying UTI was derived from findings generated from murine models. It is already well documented that this model may not always be the most suitable when it comes to studying human diseases [246]. In the context of UTI, there are several key differences between the human and murine urinary tracts that further compound the weaknesses of the murine model. Firstly, the hydrodynamics of micturition is drastically different between small (<3kg body mass) and large mammals. In small mammals like mice, urination is a high-speed event lasting between 0.01-2s whereas in large mammals (humans), urination lasts for

around 21s [247]. This 21s duration remains remarkably constant for all animals with a body mass ≥ 3kg [247]. Small animals micturate by expelling urine drops termed dripping urination while large animals do so by expelling a steady stream of urine termed jetting urination [247]. This difference can be explained by physics, larger animals have longer urethras and thus higher gravitational force and increased flow rates, allowing for a constant stream of urine to be expelled [247]. On the other hand, smaller animals have much smaller urethras, which create high viscous and capillary forces that limit their urine flow to discrete droplets [247]. These differences in the functional physiology of micturition between mice and human will likely select for different virulence factors and behaviours in the context of UTI. This could explain the conflicting findings surrounding which bacterial virulence factors are crucial in the pathogenesis of UTIs.

Secondly, the urothelium in mice is substantially different when compared to humans, with a thinner transitional (intermediate) urothelium (1-3 cell layers in mice vs 5-7 cell layers in humans) [248]. This means that the lamina propria is more easily exposed in mice compared to humans, predisposing murine models to more severe disease progression. The lamina propria is a thin layer of connective tissue that makes up the most superficial layer of the stroma and lies beneath the basal layer of the urothelium. Furthermore, intermediate cells in mice express uroplakins, which is not observed in human cells [248,249]. This is a crucial difference as QIR formation, proposed as a potential mechanism underling recurrent UTIs, is dependent on the interaction between FimH on UPEC and uroplakin on the intermediate cells. Therefore, QIR formation may not necessarily occur in humans due to the lack of uroplakin expression within the intermediate layer of the human urothelium. Indeed as mentioned earlier, IBC and QIR structures have yet to be confidently observed in clinical samples.

Thirdly, there is a difference in the immune response between humans and murine models. Certain murine strains e.g. CH3/HeJ mice, are hyporesponsive to LPS as they possess non-functioning TLR4, and these strains are highly susceptible to UPEC UTI [246]. Moreover, the expression of TLRs in the murine urothelium is markedly different from humans. Mice express TLR11 which is not expressed in humans and TLR11 has been suggested to play a role in protecting the urinary tract against UPEC [250]. For the shared TLRs (TLR 1-10), TLR4 expression and signalling predominates the urothelial immune response in mice whereas TLR5 signalling

predominates in humans [158]. Although TLR4 is present on the luminal surface of human umbrella cells, the CD14 co-receptor is not expressed, thus preventing LPS from inducing TLR4 signalling in humans [156]. However, P or type 1 fimbriae are able to induce TLR4 signalling in human urothelial cells [156]. Although there are studies that purport TLR4 as the dominant urothelial response in humans, these studies derived their findings from immortalised cell lines, particularly higher-grade cell lines (T24) which display more myeloid characteristics such as the expression of CD14 [158].

Another factor is the method in which UTIs are induced in murine models. Transurethral instillation is the most common method, which involves inserting a urinary catheter such that the catheter tip is directly in the bladder (Figure 17). The inoculum, typically consisting of $10^7$ - $10^9$ colony forming units, is slowly delivered via the inserted catheter. Transurethral instillation of bacterial isolate bypasses the need for urethral ascension, which pathologically does not reflect an uncomplicated UTI. The implication is that certain virulence factors will appear to be dispensable for urinary tract colonisation as isolates lacking these factors will still be able to colonise the bladder in these study models.

Furthermore, the number of bacteria used in an inoculum is usually orders of magnitude greater than that typically present during a UTI episode in humans ($10^3$ – $10^5$ CFU/ml, $10^5$ -$10^7$ CFU in a full human bladder assuming a capacity of 400-500ml or $10^4$ – $10^5$ CFU in an emptied human bladder, assuming a 100ml residual volume). Considering that mice are usually forced to urinate prior to transurethral instillation, this means that the initial bacterial density within the urinary tract immediately after instillation is phenomenally high at $10^9$ – $10^{11}$ CFU/ml, and sixty times greater than that of a typical human infection, assuming a generous residual volume of 60 µl [251].

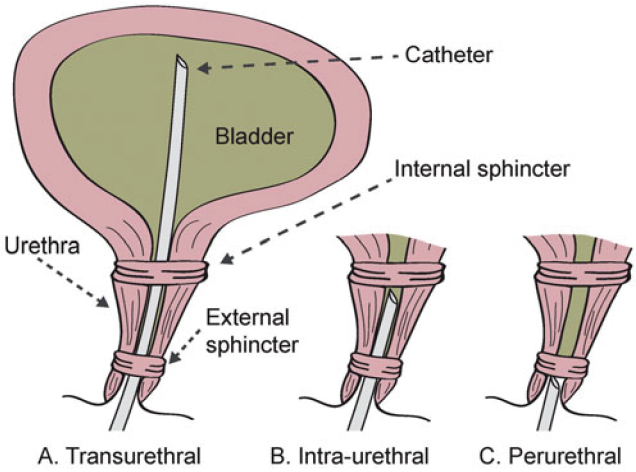



**Figure 17 Inoculation techniques that are commonly used to investigate UTI in murine models.** The diagram illustrates the placement of the catheter tip when inoculation is performed via the (A) transurethral, (B) intra-urethral, or (C) periurethral methods. The transurethral method is the most commonly used of the three. Figure reproduced from Carey AJ et al. 2014 [246]

Therefore, UTI findings generated from murine studies should be viewed with caution as they may not exactly model the human disease. An alternative to animal models are cell culture models, which are occasionally used in lieu of murine models for cost and ethical reasons. However, such models are also unable to accurately model the urinary tract as almost all of them are limited to using superficial urothelial cells. Although uropathogens primarily interact with these cells during the initial stages of a UTI, these cells are not the only components of the urinary tract and its immune response. For example, there are resident immune cells that regulate the bladder immune response during the late stage of UTI and more importantly, the physiological process of micturition that occurs intermittently in animals but is completely absent in these models. These features cannot be easily replicated with current cell culture techniques and technologies, and when combined with the aforementioned weakness of the murine model, they further strengthen the argument that large animal models are the only viable way by which UTI can be accurately modelled and studied. Although murine and cell culture models may be more affordable and acceptable on the surface, they are a false economy as they have failed to generate accurate, robust, and translatable data with regards to the human UTIs.

Ideally, future studies on UTI pathogenesis should be restricted to large mammals (>3kg) featuring either porcine, canine or feline models with bacteria being introduced via periurethral instillation in which the tip of the catheter only reaches the periurethral region (Figure 17) [3,252,253]. Of the large animal models, the porcine model

appears to be most promising as it displays a high degree of similarity to the human immune and physiological responses during a UTI [252].

## 1.7 Resistance against clinically relevant antibiotics among uropathogens

Trimethoprim, nitrofurantoin, ciprofloxacin, and cefalexin are the antibiotics prescribed for the management of UTI in the UK (Table 7) [30]. With regards to these antibiotics, the incidence of resistance among UPEC isolates is the highest for trimethoprim (35%), followed by ciprofloxacin (11%), cefalexin (9%), and nitrofurantoin (2%) [254].

**Table 7 Antibiotic dosing regimens for managing UTIs in the UK.**

| | **Lower UTI** | **Upper UTI** | **CAUTI** | **rUTI (prophylaxis)** |
|---|---|---|---|---|
| Trimethoprim | 200mg, 2x daily, 3 days | 200mg, 2x daily, 14 days | 200mg, 2x daily, 7 days | 100mg, 1x daily, indefinite |
| Nitrofurantoin | 50mg, 4x daily, 3 days | N/A | 50mg, 4x daily, 7 days | 100mg, 1x daily, indefinite |
| Ciprofloxacin | 250mg, 2x daily, 3 days | 500mg, 2x daily, 7 days | 500mg, 2x daily, 7 days | N/A |
| Cefalexin | 500mg, 2x daily, 3 days | 500mg, 2-3x daily, 7-10 days | 500mg, 2-3x daily, 7-10 days | 125mg, 1x daily, indefinite* |

Adapted from the British National Formulary [31]. *Cefalexin is unlicensed for rUTI prophylaxis

### 1.7.1 Trimethoprim resistance

In the last few decades, the increasing incidence of trimethoprim resistance among UPEC isolates has led to its replacement by nitrofurantoin as the first line antimicrobial for managing UTI [34]. Trimethoprim disrupts the tetrahydrofolate synthesis pathway by specifically targeting the bacterial dihydrofolate reductase, thereby competitively inhibiting the reduction of dihydrofolic acid to tetrahydrofolic acid (Figure 18) [255]. Tetrahydrofolic acid (encoded by *folA*) is an essential component in the synthesis of thymidine, which is used in the synthesis of bacterial DNA [255]. Trimethoprim is often used in conjunction with sulfamethoxazole due to their synergistic effects. Sulfamethoxazole inhibits dihydropteroate synthase (encoded by *folP*), an enzyme involved further upstream in the tetrahydrofolate synthesis pathway (Figure 18) [255].

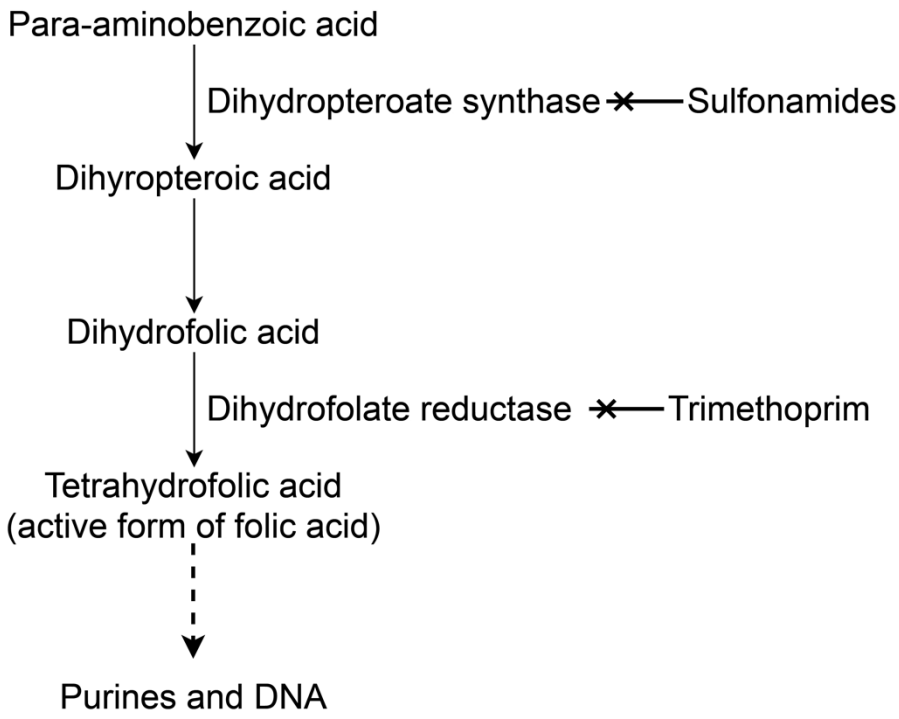



**Figure 18 How sulphonamides and trimethoprim disrupt the folic acid synthesis pathway.** Adapted from Masters PA et al. 2003 [255]

Some bacteria e.g. *Campylobacter jejuni and Helicobacter pylori*, are innately resistant to trimethoprim as they do not express dihydrofolate reductase and thus do not offer a target for the drug [256]. For organisms that express dihydrofolate reductase, chromosomal resistance can be acquired either by overexpression to bypass the competitive inhibition by trimethoprim or by point mutations within *folA* that prevent trimethoprim binding, the latter is particularly common amongst *S. pneumoniae* [256]. For UPEC, plasmid-borne resistance is perhaps the most common mechanism by which trimethoprim and sulphonamide resistance is acquired [256]. Resistance is mediated by the mobilisation of foreign resistant variants of dihydrofolate reductase (*dfr*) or dihydropteroate synthase (*sul)* within the integrons of conjugative plasmids [256]. To date, there are at least 36 resistant *dfr* variants and three resistant *sul* variants but little is known about the genetic origins of even the earliest variants [256]. The heavy usage of trimethoprim and the low fitness cost of maintaining these mutant genes probably explains the explosive increase in trimethoprim resistance over the last few decades [256]. In response to this, coordinated measures have been taken to reduce trimethoprim usage in the UK.

### *1.7.2 Nitrofurantoin resistance*

Nitrofurantoin is a broad-spectrum antibiotic that has been used clinically since the mid-1950s to manage uncomplicated UTIs [257]. Its popularity waned in the 1970s with

the introduction of trimethoprim and modern β-lactam antibiotics, which were more suitable for individuals with decreased renal function [257,258]. However, there has been renewed interest in nitrofurantoin due to the increased resistance against trimethoprim and β-lactams observed in uropathogens over the last few decades [257]. It is effective against a variety of common uropathogenic bacterial species, i.e. *E. coli, Enterococcus spp., Enterobacter spp., Klebsiella spp.* ,making it a versatile choice for treating UTIs [259].

The exact mechanism(s) by which nitrofurantoin exerts its antimicrobial effects is still unknown. Most uropathogens possess genes encoding nitroreductases that convert nitrofurantoin into electrophilic intermediates that attack bacterial ribosomal proteins, thereby inhibiting protein synthesis [259]. These intermediates can also interfere in nucleic acid synthesis and aerobic metabolism (citric acid cycle) when nitrofurantoin is present at high concentrations [259]. There is also evidence that unreduced nitrofurantoin is bactericidal as antibacterial activity has been observed under conditions where nitroreductase activity was inhibited [259]. Uropathogens that lack nitroreductase genes, e.g. *P. mirabilis*, are innately resistant against nitrofurantoin.

#### 1.7.2.1 Pharmacokinetics of nitrofurantoin

When used for UTI management, nitrofurantoin is taken orally in the form of tablets or liquid suspension. The majority of nitrofurantoin absorption occurs rapidly in the small intestine, and the drug is rapidly excreted in bile (10%) and urine (90%), resulting in a short serum half-life of 20 minutes [260]. Up to 25% of an oral dose is excreted as unmetabolised nitrofurantoin in the urine, with the remainder being enzymatically degraded [261].

Nitrofurantoin comes in two formulations: immediate-release (Macrodantin) and modified-release (Macrobid). For the management of acute UTIs, immediate-release nitrofurantoin is prescribed as 50mg tablets taken 4 times daily for 3 days whereas modified-release nitrofurantoin is prescribed as 100mg tablets taken 2 times daily for 3 days [262]. For prophylactic usage, modified-release nitrofurantoin is prescribed as 100mg tablets taken once daily at night [262]. In clinical usage, both formulations of nitrofurantoin performed equally in the management of uncomplicated and complicated cystitis. However, the modified-release formulation performed significantly better than the immediate-release formulation in the management of pregnancy-associated cystitis [263].

Most of the available pharmacokinetic data on nitrofurantoin was derived from decades-old studies employing what are now perceived as outdated laboratory and analytical methods [260,264]. Furthermore, a number of studies were conducted using animal models that may not accurately reflect the metabolic profile and urinary habits of actual human patients [253,265,266]. Interestingly, a recent study that generated valuable information on the urinary concentration-time curve of nitrofurantoin in healthy female volunteers showed considerable interindividual variability in maximum concentration and half-life (Figure 19) [267].

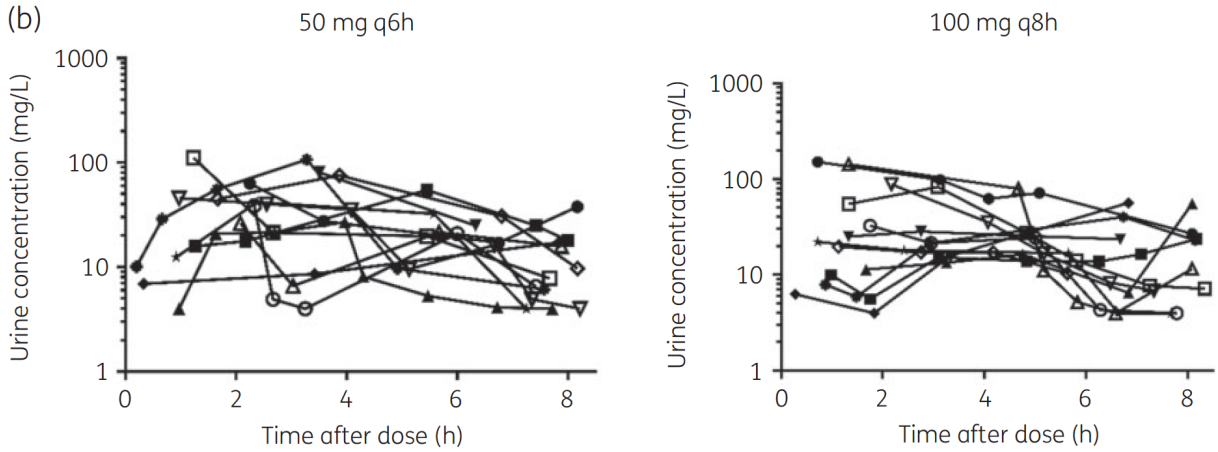



**Figure 19 Urinary nitrofurantoin concentration over time in healthy female volunteers (n=12) under 2 different dosing regimens (50mg 4x daily and 100mg 3x daily) with the immediate-release formulation.** The former regimen (50 mg 4x daily) is more commonly used in the UK. Each line represents one of the twelve volunteers. Urinary nitrofurantoin concentration shows a considerable fluctuation after antibiotic consumption and high interindividual variability.

### 1.7.2.2 Pharmacodynamics of nitrofurantoin

Nitrofurantoin is a bactericidal antibiotic, which has the unusual characteristic of displaying different killing behaviours depending on the targeted bacterial species. Against *E. coli* and *K. pneumoniae*, nitrofurantoin displays time-dependent killing, but against *E. cloacae*, killing is concentration-dependent [268]. For *E. coli* and *K. pneumoniae*, increasing nitrofurantoin concentration beyond the threshold required for bactericidal activity does not increase killing efficiency [269]. This concentration threshold is determined by the innate resistance of the bacteria against nitrofurantoin (1-2X MIC for *E. coli*) and urinary pH (lower urinary pH will decrease the threshold) [270]. Once the concentration threshold has been reached, killing efficiency is determined by the length of exposure at and/or above the threshold [269].

The rate of killing remains constant regardless of concentration as long as the concentration threshold is reached [269]. Therefore, effective clearance of *E. coli* with nitrofurantoin is dependent on maintaining a high enough urinary nitrofurantoin

concentration for extended periods [269]. The clinical implication is that the efficacy of nitrofurantoin at bacterial clearance is highly dependent on both patient compliance and renal function.

#### 1.7.2.3 Mechanism of resistance

The low incidence of nitrofurantoin resistance was originally attributed to the ability of the drug to simultaneously interfere with multiple metabolic mechanisms especially when reduced by bacterial nitroreductases [271]. However, bacterial nitroreductases are not essential proteins, which allows for susceptible bacterial isolates to acquire nitrofurantoin resistance via inactivation of corresponding genes [271]. In *E. coli*, the deletion of both *nfsA* and *nfsB* is implicated in the acquisition of clinical nitrofurantoin resistance. The genes, *nfsA* and *nfsB,* encode oxygen insensitive nitroreductases, with *nfsA* accounting for 70% of nitroreductase activity in *E. coli* and *nfsB* accounting for the remaining 30% [272].

The acquisition of nitrofurantoin resistance follows a two-step evolutionary pathway. This involves the deletion/inactivation of *nfsA* followed by *nfsB* [273]. However, *in-vitro* experiments have demonstrated Δ*nfsA* single mutants show higher levels of nitrofurantoin resistance compared to their parental isolates [273]. In support, genetic surveillance studies of nitrofurantoin resistant clinical isolates have occasionally uncovere mutants with *nfsA* but not *nfsB* mutations [274]. It should be noted however, that most *nfsA* single mutants are not highly resistant against nitrofurantoin (MIC < 64 μg/ml) [273,274]. Nevertheless, these findings support the view that the inactivation/deletion of *nfsA* is the key first-step mutation in the acquisition of nitrofurantoin resistance.

Recent studies have uncovered other candidate genes (*oqxAB*, *ribE*, CTX-M-14) that may contribute towards nitrofurantoin resistance in *E. coli*. The plasmid-borne operon *oqxAB* is commonly associated with the IS26 insertion element and encodes a multidrug efflux pump [275]. Overexpression of *oqxAB* increases resistance against multiple antimicrobial agents including chloramphenicol, nitrofurantoin, quinolones, quinoxalines, and tigecycline [275] although it only bestows sub-clinical levels of nitrofurantoin resistance (MIC < 64 μg/ml) [274,276]. However, it can synergise well with existing *nfsA*/*nfsB* mutations to greatly enhance resistance (2-16 fold increase) [276]. Despite these concerns, its relevance in the UK clinical environment is questionable as *oqxAB* expression is associated with regions where there is unregulated and

rampant antibiotic usage [275], with the unsurprising majority of studies originating from southern mainland China [275].

The gene *ribE* encodes lumazine synthase, which is essential for the production of riboflavin [277]. The inability of Gram-negative bacteria to utilise environmental riboflavin means that these bacteria are entirely dependent on endogenous riboflavin synthesis to meet their metabolic needs [278]. Although riboflavin is not biologically active by itself, it is the precursor of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD) [279]. These are essential components of cellular metabolism as they are the cofactors of numerous flavoproteins, including NfsA and NfsB [273].

The association between *ribE* and nitrofurantoin resistance was established in a study which isolated a *ribE* mutant generated via *in-vitro* selection of NitR mutants in an anaerobic environment [280]. This mutant had a partial deletion in *ribE* which involved the active site of the enzyme and compared to its parental isolate, the MIC of the mutant was two-fold higher (32 vs 16 μg/ml) [280]. Moreover, complementing the mutant with a recombinant vector encoding wild-type *ribE* restored nitrofurantoin susceptibility, thereby confirming the role of *ribE* in mediating nitrofurantoin resistance [281].

However, *ribE* may not be clinically relevant with regards to nitrofurantoin resistance due to the complete absence of *ribE* mutants amongst NitR clinical isolates [276,280]. Additionally, the inactivation of *ribE* is associated with a reduction in fitness rendering it a poor evolutionary choice for highly competitive clinical isolates [280]. Furthermore, *ribE* mutants were only isolated *in-vitro* when selection was carried out in an anaerobic environment [281], which does not accurately model the aerobic environment of the lower urinary tract (4ppm dissolved oxygen) [282]. Nitrofurantoin is only ever used in a clinical setting to manage pathologies associated with the urinary tract due to its unique pharmacokinetics [267]. Therefore, *ribE* mutants are in essence a laboratory curiosity with potentially low clinical implications.

The enzyme CTX-M-14 belongs to a group of extended-spectrum β-lactamases that is rapidly spreading among *Enterobacteria* due to widespread carriage on mobile genetic elements [283]. Although it is a β-lactamase, specific mutations including T55A, A273P and R277C enable it to hydrolyse nitrofurantoin with surprising efficiency when overexpressed [284]. However, unlike *oqxAB*, this mutated variant of CTX-M-14 can confer high-level nitrofurantoin resistance (MIC ≥512 μg/ml) by itself without

requiring any additional mutations in *nfsA*/*nfsB* [284]. This novel β-lactamase was first detected in several urinary *E. coli* isolates from a Welsh hospital in 2020 [284]. The extrachromosomal nature of this gene also means that it can easily spread via horizontal gene transfer especially when there is sufficient selective pressure [284].

### *1.7.3 Ciprofloxacin resistance*

Ciprofloxacin is a fluoroquinolone that is commonly used to treat pyelonephritis. This antibiotic exhibits its antimicrobial properties by binding to GyrA and ParC, preventing re-ligation of broken DNA strands during DNA replication [285]. This leads to an accumulation of double-stranded breaks, which disrupt cell growth and eventually cause apoptosis [286]. Clinical resistance to ciprofloxacin (MIC > 0.5 μg/ml) in *E. coli* is associated with a two-step mutation in GyrA (S83L-D87N point mutations) that prevents the binding of ciprofloxacin to GyrA [287]. It should be noted that the first-step mutation (GyrA S83L) does grant sub-clinical resistance to ciprofloxacin. Higher levels of resistance is associated with an additional point mutation in ParC (S80I or E84K) that is dependent on the presence of an existing GyrA mutation [288,289].

### *1.7.4 Cefalexin and β-lactam resistance*

Cefalexin is a β-lactam antibiotic belonging to family of first-generation cephalosporins [290]. It is mostly effective against Gram-positive cocci and some Gram-negative bacteria including *E. coli, P. mirabilis, K. pneumoniae* [290]. All β-lactam antibiotics interfere with bacterial cell wall synthesis by irreversibly binding to penicillin-binding proteins (PBPs) via their β-lactam rings [290]. Among Gram-positive bacteria, β-lactam resistance is primarily driven by mutations in the PBPs, which reduce the binding affinity of β-lactams [291]. In UPEC and most Gram-negative bacteria, the predominant mechanism of resistance is through the acquisition of genes which encode the β-lactamase enzyme [291]. This enzyme inactivates β-lactam by hydrolysing its β-lactam ring, which removes its ability to bind to PBPs [291]. Mutations affecting the binding pocket of this enzyme can either (i) broaden its specificity by “opening up” the enzyme active site, allowing it to accommodate bulkier β-lactams which were designed to be more resistant against hydrolysis or (ii) alter the kinetic properties of the enzyme, resulting in faster hydrolysis or (iii) make them more resistant against β-lactamase inhibitors, e.g. clavulanic acid [291]. These mutated β-lactamases are usually termed extended-spectrum β-lactamases (ESBL).

High-level β-lactam resistance in UPEC can develop via the acquisition of ESBL genes, increased expression of β-lactamase due to promoter changes, and reduced β-lactam access to the periplasmic space from reduced porin expression and increased expression of efflux pumps [291]. Despite an increasing prevalence of ESBL-producing UPEC, the clinical efficacy of β-lactams in managing uncomplicated UTI cases is less affected by this resistance mechanism due to the high drug concentrations in urine [292].

## *1.8 Bioinformatics*

The process of transforming raw genomic data into useful information requires a general understanding and appreciation of the history and techniques employed in each step of this workflow: sequencing, sequence assembly, and data mining. Sequencing is the process of determining the DNA sequence of an organism's genome. However, current sequencing techniques are unable to read whole genomes in one go, instead they generate shorter reads in the range of 100 to 30,000 bases depending on the sequencing platform used. Sequence assembly is a necessary step to reconstruct the whole genome by aligning and merging the generated DNA reads. Depending on the provided DNA reads, sequence assembly can either generate a contiguous genome or long DNA fragments called contigs and scaffolds.

Both are suitable for data mining, but a contiguous genome can provide more information on loci structure, extrachromosomal elements, and repeat sequences. Data mining usually involves analysing the assembled sequences for genes of interest and a key part of this process is genome annotation where the sequences are compared against a reference database to identify the locations of genes and all possible coding regions. An annotated genome can be used in various analyses which have generated key findings in this thesis. Examples include the identification of antimicrobial genes to predict the resistance profile of a bacterial isolate, phylogenetic analysis to determine the relatedness of different isolates, and identification of extrachromosomal elements that frequently carry pathogenicity islands, i.e. plasmids, phages, transposons, insertion elements.

### *1.8.1 A brief history of genome sequencing*

#### 1.8.1.1 First-generation sequencing

Sanger sequencing (chain termination method) makes use of chain-termination PCR, capillary gel electrophoresis and fluorescent detection to determine the identity of a target DNA strand. Chain-termination PCR is similar to standard PCR except that modified dNTPs called dideoxyribonucleotides (ddNTPs) are included in the PCR reaction mixture in addition to standard dNTPs. These ddNTPs are fluorescently labelled, with a different fluorescent marker used for each of the four species (ddATP, ddTTP, ddCTP, ddGTP) so that they can be differentiated. The inclusion of a ddNTP into a DNA strand will terminate its synthesis as unlike dNTPs, ddNTPs lack the 3' hydroxyl group that is required for the extension of DNA strand (Figure 20). These ddNTPs are mixed into a PCR reaction at a fraction of the concentration of standard dNTPs to generate DNA strands of every possible length up to the full length of the target sequence. The DNA strands are sorted based on increasing fragment size using capillary gel electrophoresis. As the DNA fragments migrate through the gel, a fluorescent detector reads the fluorescently labelled ddNTPs on the 3' end of each DNA fragment to determine the identity of the ddNTP. Since the DNA fragments are sorted based on fragment size from smallest to largest, the 5' to 3' sequence of the target DNA sequence can be determined based on the chronological order by which the ddNTPs are identified (Figure 20).

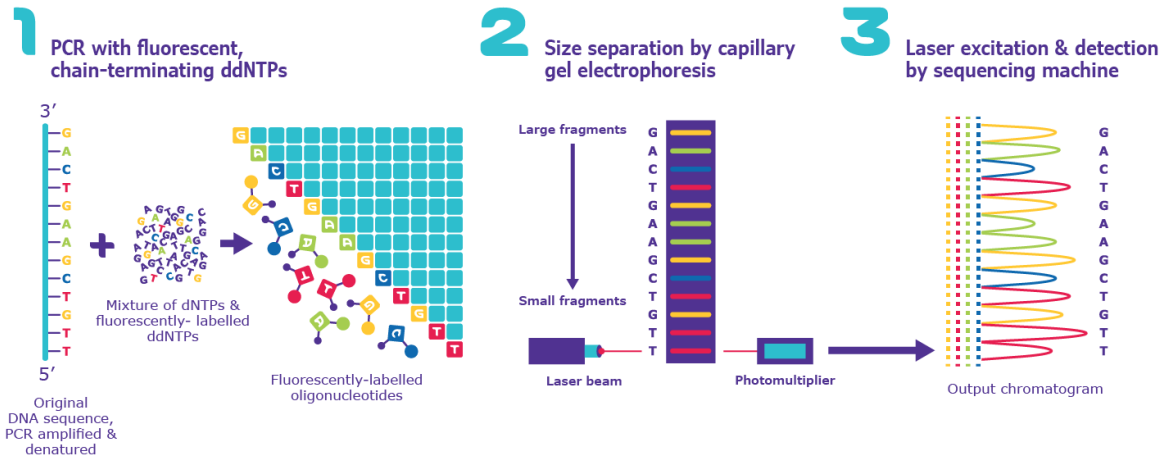



**Figure 20 A brief summary of how modern-day Sanger sequencing works.** Reproduced from a Merck-owned website [293]

Sanger sequencing is limited to sequencing DNA strands that are no more than one kilobase (1 kb) in length. To sequence longer strands, "shotgun sequencing" is performed in which the DNA strand is randomly broken up into smaller fragments (of

less than 1kb in length), cloned into a high copy number plasmid and transformed into *E. coli*. The cloned region containing the DNA fragment is amplified using flanking PCR primers and sequenced, generating a “sequence read” for each fragment. These reads are analysed using a computer algorithm which detects overlapping regions between the reads in order to stitch them together and generate a contiguous DNA sequence. As a result, the low throughput and high operational costs of Sanger sequencing limited its use for whole-genome sequencing.

#### 1.8.1.2 Second-generation sequencing

Illumina sequencing (sequencing by synthesis, SBS) has proven to be the most successful implementation of second-generation sequencing. As with “shotgun sequencing”, the DNA strand is randomly broken up into smaller fragments. Unlike “shotgun sequencing” the DNA fragments are ligated with oligonucleotide adapters so that they can be amplified *in-vitro* on a glass substrate (flow-cell) using a technique called Bridge PCR. After several rounds of amplification, this generates a cluster of ~1000 identical DNA fragments which are in close proximity, generating a “clonal cluster”. Clonal amplification ensures accurate sequencing by increasing the intensity of the fluorescence signal from each clonal cluster such that the signal will always be greater than the noise threshold.

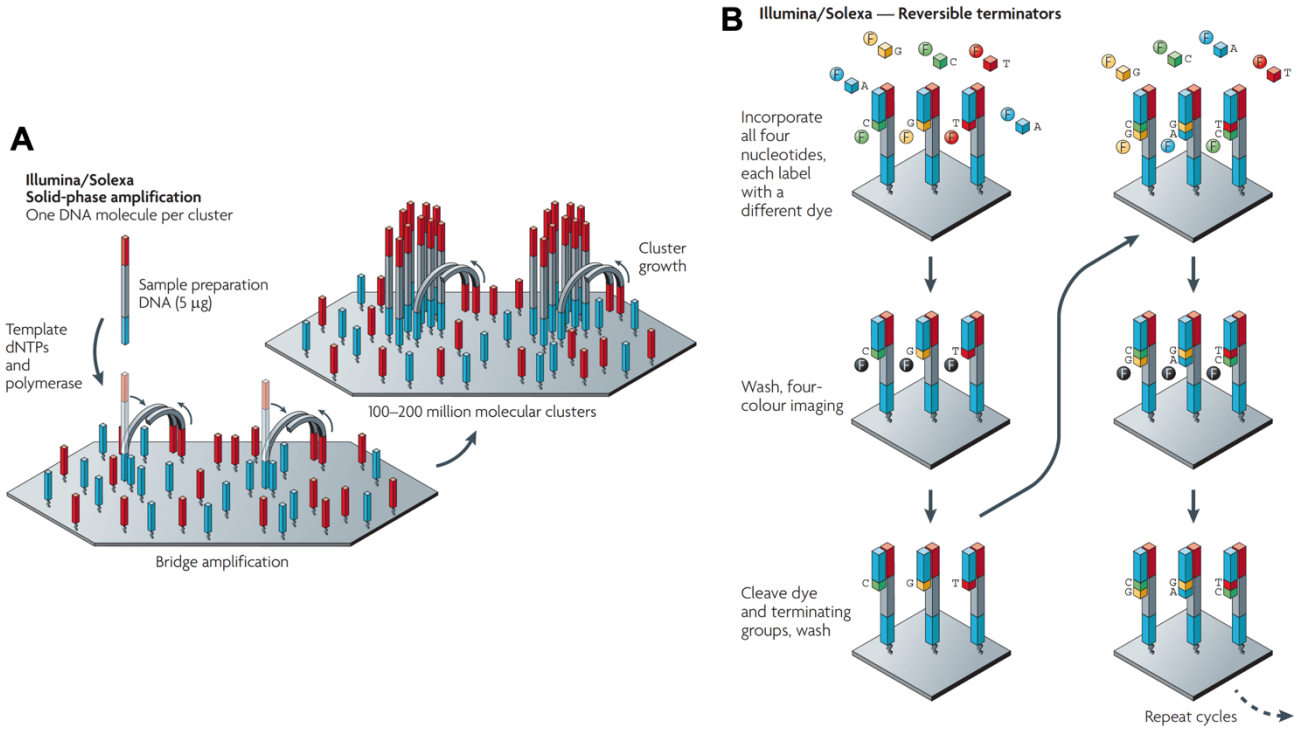



**Figure 21 Summary of the amplification (A) via bridge PCR and (B) sequencing steps in Illumina sequencing. (A)** In bridge PCR, forward and reverse primers which sequences are complementary to the oligonucleotide adapters are attached to glass surface by a flexible linker. The ligated DNA fragments are then hybridised on the forward and reverse primers attached to the glass surface. The free end of the ligated DNA fragment can interact with nearby primers on the glass surface, forming a bridge structure. PCR amplification is used to generate a second strand from the immobilised primers, any unbound DNA fragments will be washed away. **(B)** After cluster generation, the flow-cell is washed, and sequencing primers are added which hybridise to the oligonucleotide adapters.

Sequencing proceeds in cycles using a modified DNA polymerase and four nucleotides corresponding to the four DNA bases. Similar to Sanger sequencing, the nucleotides are labelled with a fluorescent reporter at the 3' end. Unlike Sanger sequencing, ddNTPs are not needed since these nucleotides only allow the synthesis of a single base during each cycle. Once the nucleotide has been incorporated, the flow-cell is washed to remove any unbound nucleotides and lasers of different wavelengths are used to excite the clonal clusters. This releases bioluminescence signals from the fluorescent reporters, which can be identified using a CCD/CMOS sensor, which translates the bioluminescence signals into digital signals that can be interpreted by a computer. For sequencing to continue, the fluorescent group at the 3' end is chemically cleaved and the flow-cell is washed before fresh nucleotides are added. The cycle is repeated, generating paired end reads of up to 300bp from each DNA fragment.

The main advantage of second-generation sequencing over first-generation sequencing is its high throughput and scalability, making it a very a cost-effective

technique for sequencing whole genomes. This democratised whole genome sequencing, making it affordable for most research groups to sequence organisms of interests, as evidenced by the explosive growth of publicly available genomes. Unfortunately, the short-read lengths of second-generation sequencing (maximum: 300bp) means that it is unable to resolve repetitive regions within genomes which can span thousands of base-pairs. Furthermore, the assembly of short read sequences has presented several computational challenges, many of which remain unresolved [294].

#### 1.8.1.3 Third-generation sequencing

Third-generation sequencing or long-read sequencing attempts to address the main limitations of second-generation sequencing: (i) short-read lengths and (ii) the need for PCR amplification which is prone to biases and errors. There are two widely used methods for third-generation sequencing: PacBio sequencing (Single-molecule real-time sequencing, SMRT) and Oxford Nanopore MinION sequencing (nanopore DNA strand sequencing).

For PacBio sequencing, fragments are generated from a template DNA, which are processed and ligated to hairpin adapters at each end. This generates a circular DNA molecule, called a “SMRTbell”, that has a constant single-stranded DNA (ssDNA) region at each end with the double-stranded DNA (dsDNA) template in the middle. SMRTbell templates undergo a size-selection protocol to remove templates, which are too large or too small to ensure efficient sequencing. Primers and a φ29 DNA polymerase are attached to the ssDNA regions of the remaining SMRTbell templates. The generated complex is added to the zero-mode waveguide (ZMW) SMRT cell where sequencing takes place. A ZMW SMRT cell is a specialised flow cell that consists of thousands of individual picolitre wells with transparent bottoms. The polymerase part of the construct is chemically fixed to the bottom of one of these wells and fluorescently labelled nucleotides are added to the flow cell. As the nucleotide is incorporated by the polymerase into an elongating DNA strand, the nucleotide will cause the polymerase to temporarily pause so that fluorescent signature can be captured by a camera that is located at the bottom of the ZMW SMRT cell. As with Sanger and Illumina sequencing, the fluorescent patterns can be interpreted to generate a DNA sequence for each well.

For the Oxford Nanopore MinION sequencing, fragmented DNA (8-10kb in length) is passed through a protein pore, which has an electric current flowing through it. A secondary motor protein translocates the DNA across the pore, which has the effect of modulating the current flowing through the pore. Several nucleotide bases of the DNA sequence can fit inside the pore at any one time and as the DNA strand is translocated, this generates a pattern of change in the electric current, known as "squiggles", that is characteristic of the particular DNA sequence inside the pore. These squiggles can be used to predict the DNA sequence.

Compared to second-generation sequencing, third-generation sequencing suffers from a much higher error rate (~15% vs ~0.5%), which limits its application and necessitates the use of second-generation sequencing data to correct any errors. Many modern genome assemblers can exploit both short (second generation) and long-read (third generation) sequence data to create a consensus genome. This combines the best of both worlds: the accuracy of second-generation sequencing and the long sequence lengths of third-generation sequencing, provided that one is able and willing to bear the additional costs.

### *1.8.2 Genome assembly*

To date, there is no sequencing technology that is able to accurately and uninterruptedly read the entire DNA of a typical bacterial genome (5-6 Mbp) [295]. Present sequencing approaches are limited to read lengths of hundreds of base pairs on Illumina platforms [296], thousands to tens of thousands for PacBio [297], and the occasional hundreds of thousands on the rapidly evolving Oxford Nanopore platform [298]. As discussed earlier, the latter two are particular prone to sequencing errors especially at longer read lengths [299]. The limitation of current sequencing techniques has necessitated the creation of computer algorithms that can automatically, accurately, and efficiently stitch together the thousands to millions of DNA sequences generated from these platforms to reconstruct a genome [300].

#### 1.8.2.1 *De novo* assembly

*De novo* assembly is the process of reconstructing a genome from short DNA sequences without the aid of a reference genome. The quality of a *de novo* assembly is primarily determined by coverage, sequencing depth and sequencing quality. The high coverage, depth, and quality of second-generation sequencing makes the generated data well suited for *de novo* assembly. However, sequencing biases and

errors are still present even in high quality reads, which can impact the quality of the assembled genome. With regards to Illumina reads, there is an inherent sequencing bias favouring GC-balanced regions compared to GC-poor regions, resulting in uneven sequencing depth across the genome especially in repeat rich regions. Sequencing errors are almost always substitution errors occurring at a rate of 0.02-0.05%.

Fortunately, sequencing errors can be analysed and removed using a combination of tools such as FastQC (quality assessment) and Trimmomatic/Fastx (trimming). FastQC is a commonly used tool that can generate plots and statistics showing the overall quality of the generated reads. The short reads generated by the Illumina sequencing platform are accompanied by quality control data that provide the probability that a given base in a DNA sequence is correct, the metric is called the Phred quality score (Q score). These data can be used to filter and trim low-quality reads to remove nucleotides that are likely to be erroneous. An important additional step in the processing of raw short reads is the screening and removal of adapter sequences. Although most sequencers nowadays automatically removes adapter sequences, there are occasions when the sequencing reaction continues past the DNA insert and into the adapter at the 3' end, e.g. short DNA fragments. Adapter sequences in the 3' end may escape the adapter screening step on the sequencing machine and if not removed, can impair the accuracy of *de novo* assembly as the artificial nature of these sequences means that their sequence composition will almost always be very different from the sequenced organism.

There are a variety of assemblers available, but almost all of them utilise one of two algorithms: de Bruijn graphs and overlap layout consensus (OLC). An example of a de Bruijn graphs assembler is Unicycler, which is an assembly pipeline that makes use of the SPAdes assembler. This is a computationally efficient strategy as sequences are mathematically abstracted but requires reads with minimal sequencing errors as it can only generate contigs by using exact matches between the sequences [301,302]. The length of these exact matches is known as the k-mer length, longer k-mers facilitate accurate assembly of high coverage regions whereas short k-mers facilitate the assembly of low coverage regions [301]. Coverage is the number of short reads associated with a specific region, high coverage regions have more short reads associated with them, meaning that there are more overlaps between the short reads and these overlaps are longer [301].

Using shorter k-mers increases the probability for misassemblies as specificity is traded for sensitivity. Using longer k-mers increases assembly accuracy, but increases the number of gaps within the assembly as longer k-mers will prevent the assembly of low coverage regions [301]. K-mer length is therefore a compromise between sensitivity and accuracy. Unicycler automatically chooses an optimal k-mer length by constructing multiple assemblies at different k-mer sizes and using an assembly graph to determine quality of each assembly [302]. The k-mer size that generates an assembly graph with the longest contigs and the least number of gaps will be chosen to generate the final assembly [302].

MaSuRCA relies on a combination of de Bruijn graphs and overlap-layout consensus (OLC) to generate assemblies [303]. Compared to de Bruijn graphs, OLC is more computationally demanding as it computes all pairwise overlaps between the millions of short reads, using sequence similarity to determine overlapping regions [301]. However, the fact that OLC uses sequence similarity instead of exact matches means that it is more robust against sequencing errors, allowing it to better cope with low-quality reads [304]. OLC is particularly useful for the assembly of third-generation sequencing reads as the algorithm benefits to a greater extent from longer read lengths and overlaps. Most current de Bruijn graph software can only work with a k-mer size and therefore overlap length of up to 127bp. Furthermore, the computational efficiency of de Bruijn graph decreases as overlap length increases, negating the main benefit of this algorithm for longer reads. Lastly, OLC is more forgiving of sequencing errors, which are especially prevalent in third-generation sequencing data, something that de Bruijn graph struggles with.

There is no one-size-fits-all with regards to *de novo* assembly as some assembly programmes work best with different datasets. For example, SPAdes/Unicycler works best with smaller amounts of data and so is well suited for bacterial projects. On the other hand, assembly programmes like MaSuRCA are designed to handle large amounts of data well, with the only downside of requiring more computational power (especially RAM) with larger amounts of data. Some assemblers can even work with both long-read and short-read data to generate a hybrid assembly. This technique entails using long-read data to resolve the gaps between the contigs generated from short-read data. Short-read data alone cannot be used to resolve those gaps as they are usually highly repetitive genomic regions and determining their boundaries is impossible due to how identical the reads are. It is akin to assembling a featureless

jigsaw puzzle where all of the pieces are identical to each other and there is no picture that can be referred to. However, with long-read data, the boundaries of such genomic regions can be easily determined as their read lengths are often greater than the length of such regions.

#### 1.8.2.2 Reference-based assembly

Another approach to genome assembly is to use a reference genome to inform the joining of read sequences. This is a computationally less demanding approach compared to *de novo* assembly as the sequenced reads do not need to be compared against each other but only against the reference genome. However, this approach cannot be used to discover novel genes or genes that are not present in the reference genome, limiting its use for genetic surveillance especially with regards to genes associated with antibiotic resistance or pathogenicity [305].

In practice, reference-based assembly is rarely performed directly on short-read sequences. Instead, the short-read sequences are first assembled into contigs using *de novo* assembly. These contigs are then mapped onto a reference genome that is closely related to the sequenced organism to generate a more contiguous genome. Reference-based assembly can be further enhanced by using multiple closely related reference genomes instead of just a single reference, this is called a synteny-based approach [306]. This approach can reduce the bias caused by rearrangements specific to a single reference genome.

The main advantage of reference-based assembly is that it does not require the generation of new sequencing data to complete a genome, instead relying on existing publicly available reference genomes. Despite the rapid progress that has been made for third-generation sequencing, it is still relatively costly compared to second-generation sequencing, limiting its accessibility. The primary drawback is that some species/genera may have more structural rearrangements compared to others, necessitating a larger number of reference genomes. For some species/genera that are not well studied, the number of available reference genomes may be insufficient or completely absent, even when the search is expanded to include more distant relatives. However, this is becoming less of an issue with more high-quality assemblies being published with each passing month.

#### 1.8.2.3 Gauging the quality of the assembly

A commonly used metric to evaluate the quality of an assembly is the N50 value, which is the median length of all contigs in an assembly. It is a measure of contiguity, with higher values indicating lower levels of fragmentation. However, N50 cannot be used to evaluate whether the contigs were accurately/correctly assembled. Some aggressive assemblers may produce longer contigs and scaffolds compared to more conservative assemblers, but are more likely to assemble genomic regions in the wrong order and orientation. Therefore, it is usually prudent to use different assemblers and compare their output using an assembly evaluation tool such as QUAST. Another useful tool is BUSCO which reviews the genes that should be present in a genome of the investigated taxonomic lineage type and then reports the number of complete and fragmented genes in the assembly. Choosing the assembly with the highest percentage of complete genes is a priority if the genome project aims to involve data mining.

### *1.8.3 Genome data mining*

#### 1.8.3.1 Annotation

A raw genomic sequence, even if properly assembled is of little use to most biologists. Genome annotation has to be performed to attach biologically relevant information to genome sequences. This is achieved by analysing the sequence structure and composition of the genome, and comparing the information with a known reference from a closely related species or genus. Most genome annotation packages/software are usually focused on the accurate identification of protein coding genes. Identification of other genetic elements such as plasmids, phages, transposable elements, non-coding RNA and CRISPR elements are usually the focus of more specialised analyses.

To correctly determine the location and structure of protein coding genes, there are two approaches that can be employed: intrinsic (*ab-initio*) or extrinsic. The intrinsic approach involves using trained algorithms to identify genomic regions that are likely to contain genes. For this approach, it is vitally important that a good training set, a collection of structurally well annotated genes, is used to train the algorithm. Since each genome is different, the algorithm has to be rebuilt and retrained for each new species, making it very labour intensive. However, a major advantage of this approach is that it is able to predict fast evolving, novel, and species-specific genes.

On the other hand, the extrinsic approach is more versatile and less demanding. There are already extensive and well-curated databases of nucleotide and polypeptide sequences (RefSeq, UniProt, NCBI non-redundant protein) that can be used as references to search for protein-coding genes within a genome. It is straightforward to annotate genomes using this approach: a sequence similarity search algorithm (such as BLAST+) is used to compare the nucleotide sequences in the genome of interest against a collection of reference databases. Sequences in the genome of interest that are similar enough to any existing entries in the reference database will be annotated with the most similar entry. However, it is important that the similarity thresholds are adjusted so that the annotation pipeline is sensitive enough to detect less conserved genes and yet specific enough to minimise false positives (misannotated regions).

An example of an effective annotation pipeline is Prokka, which combines both intrinsic and extrinsic approaches. Prokka employs multiple search strategies and curated databases in order to accurately and comprehensively annotate bacterial genomes. Prodigal (intrinsic) is initially used to predict the location of protein-coding genes within contigs. The sequences corresponding to these locations are then compared to multiple curated databases (using BLAST+ and hmmer) in a hierarchal manner to annotate these locations (extrinsic). The default e-value threshold of $10^{-6}$ is used for BLAST+ annotation [307].

#### 1.8.3.2 *In silico* AMR analysis

For clinical research, there is a strong interest to use genomic data to assess the carriage of genes that are associated with antimicrobial resistance (AMR) in pathogens. Although most available genome annotation packages/pipeline do annotate these genes, their output requires further processing to translate them into meaningful data with regards to antimicrobial resistance [308]. There are chromosomal genes, which if appropriately mutated, can bestow a degree of resistance against certain antibiotics, as in the case of *gyrA* and ciprofloxacin resistance. Although most annotation packages do annotate these genes, they do not perform any additional analyses on them to determine whether or not they might bestow resistance. Furthermore, the databases used in most annotation packages are very generalised and as such may lack the necessary information to accurately identify subsets of certain antimicrobial resistance genes, such as those involved in β-lactam resistance.

However, there are specialised annotation packages which focus on the *in silico* identification of these genes. One such example is RGI-CARD, which uses a similar approach to Prokka, but with a curated database of AMR genes that is frequently updated. Most other available *in silico* AMR annotation packages, e.g. AMRFinder, ResFinder, are merely curated nucleotide databases that can be easily integrated into an existing bioinformatic pipeline. However, the presence of a gene does not always guarantee its expression, which is why data generated from *in silico* AMR analysis cannot be used alone and needs to be accompanied with actual phenotypical data generated from culture-based assays. This shortfall is one reason why *in silico* AMR analyses have yet to replace current culture-based methods in the clinical setting.

#### 1.8.3.3 *In silico* genotyping and phylogenetic analyses

Historically, bacterial isolates were genotyped using multilocus enzyme electrophoresis (MLEE) which is based on the principle that mutations in genes encoding enzymes involved in bacterial metabolism can be detected by differences in the electrophoretic migration patterns of these enzymes [309]. Many bacterial enzymes are polymorphic and thus MLEE can easily discriminate between bacterial clones which that can be visually identified by their distinctive electrophoretic enzyme profiles. This means that all strains exhibit a unique enzyme electrophoretic profile, which can be used to identify an isolate. However, there are substantial variations in MLEE results between different laboratories, which made it difficult to compare findings and thus establish the identity of isolates when using multiple electrophoretic profile databases [310].

To address the challenges of MLEE, multilocus sequence typing (MLST) was developed, which entails the sequencing of six or seven well-conserved, housekeeping genes. The allelic variation at each gene is catalogued and assigned a unique allele number. The combination of allele number specific to an isolate is abstracted into a unique allelic profile, which is often referred to as a sequence type (Figure 22). It should be noted that most publications do not explicitly provide any rationale as to why certain housekeeping genes were chosen in their MLST scheme or why the number of genes used in most MLST schemes is frequently six or seven. In the first paper that described a MLST scheme, the genes in that scheme were chosen because the combination of those genes generated findings that were very similar to an existing MLEE scheme [310]. Nevertheless, the *E. coli* Achtman MLST

scheme that was proposed in 2006 and is still currently used for *E. coli* genotyping appear to be congruent with the findings generated from other genotyping techniques, agreeing with even a recently published findings of a whole genome evolutionary analysis of that species [311].

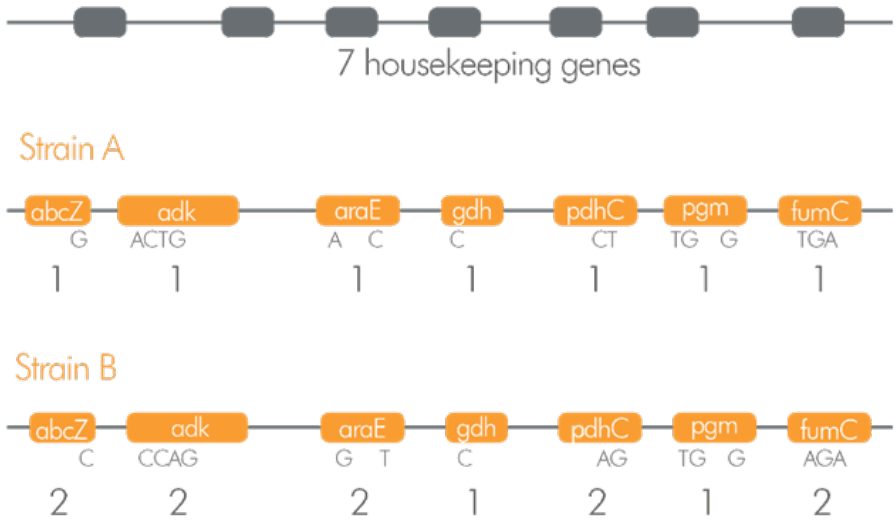



**Figure 22 How MLST works.** MLST analysis make use of housekeeping genes which should be universally present among a certain species, due to their highly conserved nature. The unit of comparison in MLST are the alleles of these genes. Each allelic change is counted as a single genetic event, regardless of the number of nucleotide polymorphisms involved. Each unique allele sequence is assigned an arbitrary and unique allele number. The combination of allele numbers specific to a strain is abstracted into unique (and arbitrary) allelic profile (1-1-1-1-1-1-1 for strain A) or sequence type (ST 1). [312]

The only drawback of MLST is its tediousness as it requires the PCR amplification and sequencing of seven different genes for each tested isolate. However, with whole-genome sequencing becoming more accessible, it is now possible to perform *in silico* MLST once the data has been assembled and annotated. This is rather straightforward as all that needs to be done is to use a sequence similarity search algorithm (BLAST) to extract the MLST genes from the assembled genome. Once done, the sequences can be uploaded to an online MLST database (pubMLST) that can automatically assign allele and ST numbers to the sequence data. There are even ready-to-use analytical pipelines that can work directly with raw sequencing data. This not only makes the task easier, but also allows for the inclusion of more genes in the scheme, thereby increasing the phylogenetic resolution, and allowing more closely related isolates to be identified. Furthermore, instead of assigning arbitrary sequence type numbers to isolates, the genomic data of the sequenced isolates can be used to generate phylogenetic trees, which not only allow data to be compared, but also provide more evolutionary data.

## 1.9 Research aims

Although most acute UTI infections can be easily treated with antibiotics, there exists a substantial population of patients who present with long-term bacterial colonisation of the urinary tract. Such patients react very differently to this colonisation with some remaining asymptomatic but others exhibiting severe and frequent UTI symptoms. This picture is further complicated by patients who frequently transition between asymptomatic and symptomatic states. At present there is little understanding of the host-microbial interactions underpinning these clinical conditions.

Hence further understanding of the host-microbial interactions underlying the pathogenesis of asymptomatic bacteriuria and recurrent UTIs is essential to effectively manage UTI sufferers. To explore these interactions further, data and isolates from two clinical studies (AnTIC & BUTI) were analysed to determine the effects of antibiotic prophylaxis (AnTIC) and host-specific factors (BUTI) on long-term bacterial colonisation [14,50]. In AnTIC, antibiotic prophylaxis has been effective at reducing the frequency of symptomatic UTIs among intermittent catheterised patients with a history of rUTIs but this was not always associated with bacterial clearance. In BUTI, a subset of patients with a common TLR1 single nucleotide polymorphism exhibited stable yet clinically asymptomatic bacterial colonisation. Investigating these two cases where asymptomatic bacterial colonisation was either promoted by antibiotic prophylaxis or a host-specific factor should not only help further our understanding on the host-microbial interactions underlying UTI but also lay the foundation for alternative antibiotic-free approaches of managing UTIs by modulating instead of completely eliminating microbial colonisation.

Additionally, the discovery of *E. coli* isolates which developed nitrofurantoin resistance *in-vivo* during AnTIC provided an opportunity to investigate nitrofurantoin resistance using clinical isolates from which mutants can be generated and compared against. Nitrofurantoin is of particular importance in UTI management as the low incidence of nitrofurantoin resistance amongst *E. coli,* the predominant causative organism of UTIs, makes it highly effective at treating most UTI cases. However, little is still known about the factors driving the low incidence of nitrofurantoin resistance despite its long clinical usage. Due to the recent uptick in nitrofurantoin usage within the UK and the general concern of increasing antibiotic resistance, which was the main side-effect of antibiotic prophylaxis usage in AnTIC,

investigating the evolutionary dynamics underlying the low incidence of nitrofurantoin resistance should generate findings that would ideally inform proper stewardship of this antibiotic, ensuring its viability in the decades to come.

The main aims of this project were to:

1. Investigate the impact of antibiotic prophylaxis on bacterial colonisation (Chapter 3).
2. Investigate the evolutionary dynamics underlying the low incidence of nitrofurantoin resistance in *E. coli* isolated from rUTI patients (Chapter 4).
3. Investigate the impact of a host-specific factor (TLR1 1805T SNP) on bacterial colonisation (Chapter 5).

# Chapter 2 Methods

## 2.1 Growth and storage of bacteria

All bacterial strains were catalogued and stored at -80°C in Luria-Bertani (LB) liquid media with 10% dimethyl sulfoxide (DMSO). Frozen bacterial strains were activated by overnight incubation on LB agar plates. LB agar cultures were stored at 4°C for up to 2 weeks before being discarded. Bacterial suspensions were prepared by inoculating 5 ml (3 ml for large-scale experiments) of liquid media with a colony forming unit (CFU) from the LB agar cultures. Bacterial suspensions were grown overnight in a SANYO orbital incubator shaker set to 160 rpm. For strains with plasmids, appropriate antibiotics at following concentrations (Table 8) were added to the agar or liquid media to ensure plasmid retention.

**Table 8 Antibiotic concentrations**

| Antibiotic | Stock concentration (200X) | Final concentration |
|---|---|---|
| Ampicillin (Amp) | 20 mg/ml | 100 µg/ml |
| Chloramphenicol (Cm) | 2.5 mg/ml | 12.5 µg/ml |
| Kanamycin (Kan) | 10 mg/ml | 50 µg/ml |
| Tetracycline (Tet) | 2.5 mg/ml | 12.5 µg/ml |
| Gentamycin (Gent) | 4 mg/ml | 20 µg/ml |
| Spectinomycin (Spec) | 10 mg/ml | 50 µg/ml |

## 2.2 Wet-lab genetics

### 2.2.1 Primer storage and usage

Primers were sourced from Integrated DNA Technologies (IDT) Inc. and purified using standard desalting. Primers were resuspended in PCR grade water (Merck) to create a stock solution with a final concentration of 200 µM and stored at -20°C.

For PCR reactions, the stock primers were diluted 1:10 to a working primer solution of 20 µM. For a routine 50 µl PCR reaction, 2.5 µl of this working primer solution was added to achieve a final primer concentration of 1µM for each primer. Both stock and working solutions were stored at -20°C.

### 2.2.2 Genomic DNA extraction and purification

Bacterial genomic DNA was isolated using the Sigma-Aldrich GenElute Bacterial Genomic DNA Kit via centrifugation as per manufacturer's instructions with the

exception that genomic DNA was eluted using PCR grade water. The concentration and quality of the genomic DNA was determined using a Nanodrop 2000 spectrophotometer. Genomic DNA was used for next-generation sequencing, mutagenesis, and PCR reactions.

### *2.2.3 DNA extraction via heat treatment*

Bacterial suspensions were generated by overnight incubation of inoculated liquid media. For each sample, 20 µl of bacterial suspension was mixed with 180µl of PCR water in a 1.5ml microcentrifuge tube. The tubes were centrifuged at 14 000 x *g* for 5 minutes, the supernatant removed using vacuum suction, and the bacterial pellet resuspended in 100 µl of PCR water by vortexing. The tubes were then incubated at 100°C for 10 minutes in a dry block incubator and immediately transferred into an ice bath where they remained until used. Fresh DNA preparations were used when performing PCR assays.

### *2.2.4 Plasmid extraction*

Plasmids were isolated using the New England Biolabs (NEB) Monarch Plasmid Miniprep Kit as per manufacturer's instructions with the exception that plasmids were eluted using PCR grade water. The concentration and quality of the eluted plasmid was determined using a Nanodrop 2000 spectrophotometer and agarose gel electrophoresis.

### *2.2.5 Amplification of DNA using Polymerase Chain Reaction (PCR)*

Both Q5 and Taq polymerases were sourced from NEB and came with their own respective buffer solutions. Stock solutions of deoxynucleotides (dNTPs) were also sourced from NEB, which were mixed and diluted to a working concentration of 8 mM. All reagents were stored at -20°C. PCR reactions were prepared in 0.2 ml PCR tubes (Table 9), which were centrifuged at 13 000 x *g* prior to being placed in a Biometra T3000 thermocycler. The thermocycler was programmed following the recommended guidelines provided by NEB, with the annealing temperature calculated using the company's online tool (https://tmcalculator.neb.com/).

Completed PCR reactions were held at 4°C before being transferred to -20°C storage. The quality and quantity of the PCR products were routinely assessed using agarose gel electrophoresis and a Nanodrop 2000 spectrophotometer.

Table 9 PCR reaction reagents.

| Q5 PCR (50 µl reaction) | | | Taq PCR (50 µl reaction) | | |
|---|---|---|---|---|---|
| **Item** | **Final conc.** | **Vol. (µl)** | **Item** | **Final conc.** | **Vol. (µl)** |
| Water | | 27.5 | Water | | 32.75 |
| 5X Q5 buffer | 1x | 10 | 10X Taq buffer | 1x | 5 |
| DNTP | 200 µM | 5 | DNTP | 200 µM | 5 |
| Forward Primer | 1 µM | 2.5 | Forward Primer | 1 µM | 2.5 |
| Reverse Primer | 1 µM | 2.5 | Reverse Primer | 1 µM | 2.5 |
| Template DNA | 1 ng/µl gDNA or 0.1 ng/µl plasmid DNA | 2 | Template DNA | 1 ng/µl gDNA or 0.1 ng/µl plasmid DNA | 2 |
| Polymerase | 20 units/ml | 0.5 | Polymerase | 25 units/ml | 0.25 |

### *2.2.6 Agarose gel electrophoresis*

PCR products and plasmids were routinely assessed using 0.8% Tris-Acetate EDTA (TAE, 0.4 M Tris-Acetate, 1 mM EDTA, Fisher-Scientific) Agarose (0.8% w/w) gels. Agarose concentration was adjusted based on the expected size of the PCR products. NBS Biologicals SafeView Nucleic Acid Stain was added to the agarose solution as per manufacturer's instructions. DNA ladders (1 kb and 100 bp) and gel loading buffer solutions were sourced from NEB.

Electrophoresis routinely used 1x TAE buffer (0.4 M Tris-Acetate, 1 mM EDTA) and were conducted using a Labnet Power Station 300 set at 100V (constant voltage) for 45 minutes. Agarose gels were photographed in a Syngene G:BOX Chemi XRQ outfitted with an inbuilt UV filter and a Syngene Transilluminator that emitted UV light at a wavelength of 302 nm.

### *2.2.7 PCR clean-up*

PCR products that were used for sequencing and mutagenesis were treated with the DpnI restriction enzyme and then purified using the Monarch PCR & DNA Cleanup Kit from NEB as per manufacturer's instructions with the exception that purified PCR products were eluted using PCR grade water. The concentration and quality of the eluted plasmid DNA was determined using a Nanodrop 2000 spectrophotometer.

### 2.2.8 Sanger Sequencing

PCR products that required sequencing were sent to Source BioScience. Quality control and error correction of raw sequence data was done using 4Peaks and Benchling.

### 2.2.9 High Resolution Melting Analysis for SNP genotyping

SNP genotyping of RT4 urothelial cells was carried out using the Thermofisher Scientific SYBR Green PCR Master Mix kit (Table 10) on a Roche Lightcycler 480 in melting curve analysis mode as per manufacturer's recommendations.

**Table 10 qPCR reaction reagents**

| Item | Final conc. | Vol. (µl) |
|---|---|---|
| Water | | 7.2 |
| 2X qPCR Mix | 1X | 10 |
| Forward Primer | 300 nM | 0.3 |
| Reverse Primer | 300 nM | 0.3 |
| SYBR Green | 1 µM | 0.2 |
| Template DNA | Varies | 2 |

### 2.2.10 Multiplex PCR for E. coli phylotyping

*E. coli* phylotyping was based on the 2013 Clermont *E. coli* phylotyping scheme [313], primer sequences are shown in Table 11. Quadruplex Primer Master Mix (Table 11A) and the Group E Primer Master Mix (Table 11) were prepared from the primers listed in Table 11. PCR reactions were performed (Table 12 & Table 13) and assessed on 2% TAE-agarose gels (0.4 M Tris-Acetate, 1 mM EDTA, 2% agarose w/w) (Figure 23A,B) to determine the phylotype of the isolates.

**Table 11 Clermont phylotyping primers**

| Primer No. | Primer Name | Sequence | Volume (µl) |
|---|---|---|---|
| **1482** | chuA.1b (F1) | ATGGTACCGGACGAACCAAC | 10 |
| **1483** | chuA.2 (R1) | TGCCGCCAGTACCAAAGACA | 10 |
| **1484** | yjaA.1b (F2) | CAAACGTGAAGTGTCAGGAG | 10 |
| **1485** | yjaA.2b (R2) | AATGCGTTCCTCAACCTGTG | 10 |
| **1486** | TspE4C2.1b (F3) | CACTATTCGTAAGGTCATCC | 10 |
| **1487** | TspE4C2.2b (R3) | AGTTTATCGCTGCGGGTCGC | 10 |
| **1488** | AceK.f (F4) | AACGCTATTCGCCAGCTTGC | 20 |
| **1489** | ArpA1.r (R4) | TCTCCCCATACCGTACGCTA | 20 |
| **1490** | ArpAgpE.f (F5) | GATTCCATCTTGTCAAAATATGCC | 10 |
| **1491** | ArpAgpE.r (R5) | GAAAAGAAAAAGAATTCCCAAGAG | 10 |

Primers highlighted in green were used to create the quadruplex primer master mix in which an appropriate volume of the highlighted primers was mixed together. Primers highlighted in yellow were used to create the Group E Primer Master Mix in which an appropriate volume of the highlighted primers was mixed with 80 µl of PCR grade water. The forward (F) and reverse primers (R) for each primer pair (the number following F/R) are shown next to the primer name in parentheses.

**Table 12 Clermont PCR reaction**

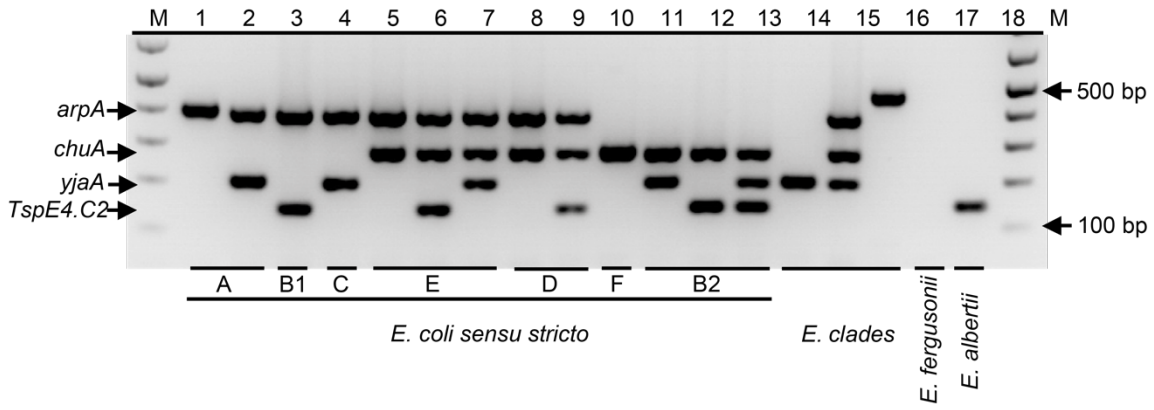
| Water | Buffer | DNTP (2mM) | Master Mix | Template DNA | Taq Polymerase | Total Volume |
|---|---|---|---|---|---|---|
| 11.6 µl | 2 µl 10X Taq buffer | 2 µl (200µM) | 1 µl [a] | 3 µl bacterial lysate [b] | 0.4 µl (40 units per ml) | 20 µl |

[a] 1µM for Primers 1482-1487, 2 µM for Primers 1488-1489, 0.6 µM for Primers 1490-1491. [b] Alternatively, genomic DNA can be used with a final concentration of 10 ng/µl.

**Table 13 Clermont PCR program**

| | Temperature | Time | Number of cycles |
|---|---|---|---|
| **Initial denaturation** | 94°C | 4 minutes | 1 |
| **Denaturation** | 94°C | 5 seconds | 30 |
| **Annealing** | 59°C (Quadruplex)<br>57°C (Group E) | 20 seconds | |
| **Final extension** | 72°C | 5 minutes | 1 |
| **Hold** | 4°C | Indefinite | |

**A**




**B**

| Quadruplex genotype | | | | | |
|---|---|---|---|---|---|
| *arpA* (400 bp) | *chuA* (288 bp) | *yjaA (211 bp)* | *TspE4.C2 (152 bp)* | Phylo-group | Next step |
| + | - | - | - | A | |
| + | - | - | + | B1 | |
| - | + | - | - | F | |
| - | + | + | - | B2 | |
| - | + | + | + | B2 | |
| - | + | - | + | B2 | Could be confirmed by testing ibeA gene[a] |
| + | - | + | - | A or C | Screen using C-specific primers. If C+ then C, else A |
| + | + | - | - | D or E | Screen using E-specific primers. If E+ then E, else D |
| + | + | - | + | D or E | Screen using E-specific primers. If E+ then E, else D |
| + | + | + | - | E or clade I | Screen using E-specific primers. If E- then clade I, confirm using cryptic clade primers[b] |
| - | - | + | - | Clade I or II | Confirm using cryptic clade primers[b] |
| - | - | - | - | Clade III, IV or V | Confirm using cryptic clade primers[b] |
| - | (476)[c] | - | + | Unknown | Perform MLST |
| - | - | + | + | Unknown | Perform MLST |
| + | - | + | + | Unknown | Perform MLST |
| + | + | + | + | Unknown | Perform MLST |
| - | - | - | - | Unknown | Confirm *Escherichia* identification using uidA or gadA/B[d], if positive screen using cryptic clade primers[b] and/or perform MLST |

**a.** Gordon and colleagues (2008).
**b.** Clermont and colleagues (2011b).
**c.** The quadruplex PCR reaction will result in strains belonging to cryptic clade III, IV or V yielding a 476 bp PCR product. If this outcome eventuates then such strains should be screened using the cryptic clade detection primers (Clermont
**d.** McDaniels and colleagues (1996).

**Figure 23 (A) Agarose gel and (B) PCR profiles of the different *E. coli* phylotypes using the Clermont phylotyping scheme.** [313]. Figure reproduced from Clermont et al 2013.

### *2.2.11 RNA extraction*

RNA was extracted using the Promega SV Total RNA Isolate System via centrifugation as per manufacturer's instructions.

### *2.2.12 Two-step Reverse Transcriptase PCR (RT-PCR)*

Reverse Transcriptase PCR (RT-PCR) reactions were prepared in 0.2 ml PCR tubes. The first step of RT-PCR involves synthesising cDNA from the RNA template. 0.5 µg of RNA template was mixed with the RT-PCR annealing mix containing random hexamers (Table 14), incubated at 65°C for 5 minutes in a thermocycler, and placed on ice for at least 2 minutes before the addition of RT-PCR cDNA synthesis mix (Table 15). Tubes were centrifuged before being transferred into a thermocycler (42°C for 2 hours, 70°C for 30 minutes). Tubes were immediately transferred to -

20°C storage once the thermocycler run was completed. The synthesised cDNA was then PCR amplified following the protocol in Section 2.2.5.

**Table 14 RT-PCR annealing mix**

| Item | Final conc. | Vol. (µl) |
|---|---|---|
| RNA | 50 ng/µl | Varies |
| Random hexamers | 2.5 µM | 0.5 |
| Water | 300 nM | Up to a final volume of 10 µl |

**Table 15 RT-PCR cDNA synthesis mix**

| Item | Final conc. | Vol. (µl) |
|---|---|---|
| PCR water | | 2 |
| 5X MMLV buffer | 1X | 5 |
| DNTP | 400 µM | 5 |
| RNasin | 2 units/µl | 2.5 |
| MMLV RT enzyme | 4 units/µl | 0.5 |
| cDNA solution | Varies | 10 |

## *2.3 Transformation techniques*

### *2.3.1 Heat shock*

*Escherichia coli* NEB C2987 (NEB 5-alpha Competent *E. coli)* cells were used for heat-shock transformations. Tubes containing NEB C2987 cells stored at -80°C were thawed on ice for 10 minutes, and 50 ng of plasmid DNA was added to each tube. The cells and DNA were mixed by flicking the tube a few times. The tubes were placed on ice for 30 minutes before being heat-shocked at 42°C for 30 seconds in a pre-heated water bath. Immediately after heat-shock, tubes were placed on ice for 5 minutes and 950 µl of room temperature **S**uper **O**ptimal broth with **C**atabolite repression (SOC) Outgrowth medium added to each tube. Tubes were incubated at 37°C for 30 minutes in a SANYO orbital incubator set at 160 rpm. After this incubation period, the cultures were diluted $1:10^5$ and $1:10^6$ in SOC and 100 µl of each dilution was spread onto the appropriate antibiotic selection plate before being incubated overnight at 37°C.

### *2.3.2 Electroporation*

A bacterial colony was inoculated in liquid media, incubated overnight in a SANYO orbital incubator set at the appropriate temperature and 160 rpm, and the final $OD_{600}$ of the culture was measured using a Pharmacia Biotech Ultrospec 1000

spectrophotometer. A 45 ml liquid culture with a starting $OD_{600}$=0.02 was prepared from the overnight culture.

If the culture was used for lambda-red or CRISPR mutagenesis, it was incubated until it reached an $OD_{600}$= 0.08–0.12. At this point, an appropriate volume of 20% w/v arabinose solution (200X, solvent: MilliQ water) was added to the culture such that the final concentration of arabinose was 0.1% v/v before being incubated again. Once the culture reached an $OD_{600}$= 0.6-0.8, it was pelleted at 3960 x *g* using a Sigma 2-16P benchtop centrifuge installed with a 11180 rotor for 20 minutes at 4°C. The culture was washed twice with sterile, ice-cold MilliQ water using the same centrifuge settings. The final pellet was then resuspended in 300µl of sterile, ice-cold MilliQ water with vortexing.

Each electroporation reaction required 50 µl of the bacterial suspension mixed with an appropriate amount of PCR or plasmid DNA (Table 16). Negative controls were included to determine transformation efficiency (Table 16). The mixture was electroporated in a Flowgen electroporation cuvette (0.1 cm gap) using a Bio-Rad Micropulser Electroporator. The inbuilt “Ec1” program was used for all electroporation reactions (1.8kV, single pulse, no time constant).

**Table 16 Electroporation protocols for plasmid transformation (highlighted yellow), lambda-red mutagenesis (highlighted red), and CRISPR mutagenesis (highlighted green)**

| Protocol | Bacterial suspension | Plasmid DNA | Donor DNA | pTRG DNA |
|---|---|---|---|---|
| Plasmid transformation | 50 µl | 50ng | - | - |
| Plasmid transformation negative control | 50 µl | - | - | - |
| Lambda-red mutagenesis | 50 µl | - | 400ng | - |
| Lambda-red mutagenesis negative control | 50 µl | - | - | - |
| CRISPR mutagenesis | 50 µl | - | 400ng | 50ng |
| CRISPR mutagenesis pTRG negative control | 50 µl | - | - | 50ng |
| CRISPR mutagenesis negative control | 50 µl | - | - | - |

Immediately after electroporation, the cells were allowed to recover in LB liquid media for 60 minutes in a SANYO orbital incubator set at the appropriate temperature and 160 rpm (Table 17). After incubation, the transformed cells were immediately plated on the appropriate LB-antibiotic plates and incubated overnight at

the appropriate temperature. The volume and dilutions used for plating differed depending on the type of mutagenesis or transformation that was being performed (Table 17).

**Table 17 Culture volume and dilutions for plating transformed cells**

| Protocol | Dilutions | Incubation conditions |
|---|---|---|
| Plasmid transformation | 200 µl at 1:1, 1:10, 1:20 dilutions | 30°C, LB-Kanamycin plates for pKD46-Kan or pCas9 selection |
| Lambda-red mutagenesis | 500 µl and 200 µl at 1:1, and 200 µl at 1:10 dilutions | 37°C, LB-Chloramphenicol plates for *cat* gene selection |
| CRISPR mutagenesis | 200 µl at 1:10, 1:20, 1:50 dilutions | 30°C, LB-Kanamycin-Spectinomycin plates for pTRG and pCas9 selection |

## *2.4 Mutagenesis*

### *2.4.1 Gibson assembly*

#### 2.4.1.1 Constructing *nfsA* and *nfsB* gene regions with specific deletions

Gibson assembly was used to design and generate PCR products with specific gene deletions in the *nfsA* and *nfsB* regions, to be used as donor DNA in CRIPSR mutagenesis. Gibson-specific primers were manually designed and tested in Benchling using the W3110 (NC_007779.1) and 1646 BASE assembled genomes (Table 18). The “inserts” used in the Gibson assembly were generated via PCR using genomic DNA from the 1646 BASE isolate as the template. The plasmid/vector backbone was generated via PCR using the pBluekSII plasmid (AmpR) as the template. The plasmid backbone and PCR products were subjected to agarose gel electrophoresis, DpnI treatment, PCR clean-up, and DNA quantification prior to Gibson assembly.

**Table 18 Gibson assembly primers.**

| Primer No. | Primer Name | Sequence |
|---|---|---|
| **1574** | pBKSR_ybjCF | CTCGAGGTCGACGGTATCGATAAGCagatttatgcctgtaatgatcacgc |
| **1575** | delnfsA_R | catcgacgtggcagttttagcgcgtGCCACAAATAAGTTCAATGGTTGG |
| **1576** | delnfA_rmKR | gtaatgtgattgcgactaaccaccGCCACAAATAAGTTCAATGGTTGG |
| **1577** | delnfsA_F | CCAACCATTGAACTTATTTGTGGCacgcgctaaaactgccacgtcgatg |
| **1578** | delrimKR | gtaatgtgattgcgactaaccaccGGCAATTTTCACCCTGCACCTCTC |
| **1579** | delnfA_rmKF | GCCACAAATAAGTTCAATGGTTGGggtggttagtcgcaatcacattac |
| **1580** | delrimKF | GAGAGGTGCAGGGTGAAAATTGCCggtggttagtcgcaatcacattac |
| **1581** | pBKSF_rimKR | GGATCCCCCGGGCTGCAGGAATTCGatatggggcagatgcagagcctgac |
| **1582** | pBKSR_ybdFF | CTCGAGGTCGACGGTATCGATAAGCcgtcgattatcgccgaacttcgcgg |
| **1583** | delnfsBR | GCCCGGCAAGAGAGAATTACACTTCgatatccataaagactccatgtg |
| **1584** | delnfsBF | CACATGGAGTCTTTATGGATATCgaagtgtaattctctcttgccgggc |
| **1585** | pBSKF_ybdGR | GGATCCCCCGGGCTGCAGGAATTCGcagcggtaactgagatgccgccggg |
| **1586** | ybjCF_pBSKR | gcgtgatcattacaggcataaatctGCTTATCGATACCGTCGACCTCGAG |
| **1587** | ybdFF_pBSKR | ccgcgaagttcggcgataatcgacgGCTTATCGATACCGTCGACCTCGAG |
| **1588** | rimkR_PBSKF | gtcaggctctgcatctgccccatatCGAATTCCTGCAGCCCGGGGGATCC |
| **1589** | ybdGR_pBKSF | cccggcggcatctcagttaccgctgCGAATTCCTGCAGCCCGGGGGATCC |
| **1590** | ybdG+411R | cagcggtaactgagatgccgccggg |
| **1591** | ybjC-144F | agatttatgcctgtaatgatcacgc |
| **1592** | rimK+1100R | atatggggcagatgcagagcctgac |
| **1593** | ybdF-63F | cgtcgattatcgccgaacttcgcgg |

Using pBluekSII plasmid as the template, Primers 1586 & 1588 (highlighted yellow) generated the vector backbone for *nfsA* insert fragments whereas Primers 1587 & 1589 (highlighted yellow) generated the vector backbone for *nfsB* insert fragments. Modified *nfsA* insert fragments were generated with the following primer combinations using 1646 BASE genomic DNA as the template: Δ*nfsA* (Primers 1574 & 1575, Primers 1577 & 1581), Δ*nfsA-rimk* (Primers 1574 & 1576, Primers 1579 & 1581), Δ*rimk* (Primers 1574 & 1578, Primers 1580 & 1581). Modified *nfsB* insert fragments were generated with the following primer combinations using 1646 BASE genomic DNA as the template: Δ*nfsB* (Primers 1582 & 1583, Primers 1584 & 1585). To generate the donor DNA for CRIPSR mutagenesis, the insert sequence within a Gibson construct was amplified using the following primers for *nfsA* (Primers 1591 & 1592) and *nfsB* (Primers 1590 & 1593) insert sequences.

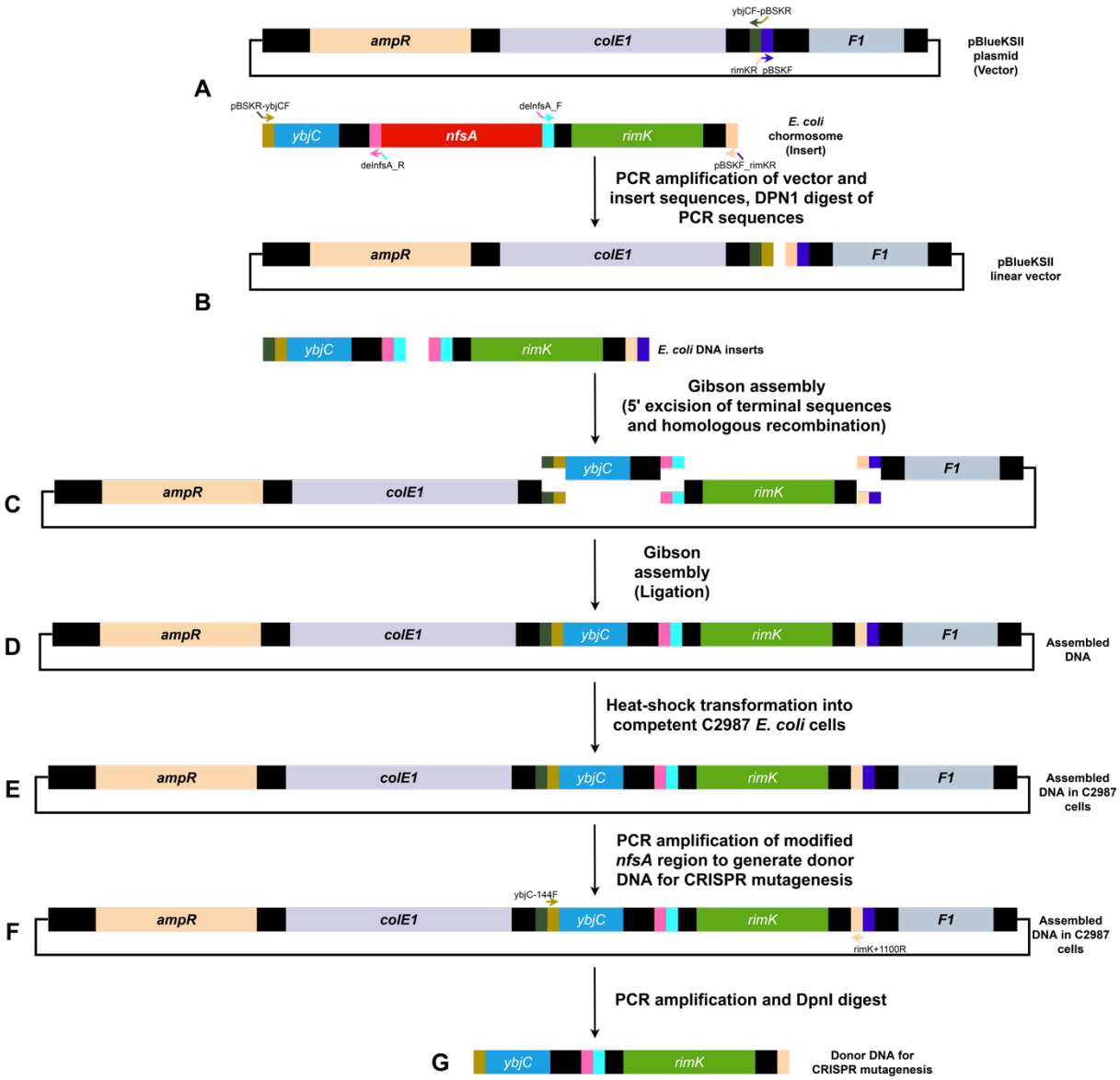



**Figure 24 Creating donor DNA with specific deletions in the *nfsA* and *nfsB* regions with Gibson assembly.** **(A)** Using the Δ*nfsA* donor DNA as a case example, DNA insert fragments were first generated from 1646 BASE genomic DNA. **(A)** These DNA fragments were generated by using primers that only amplified the genetic regions adjacent to *nfsA* but not *nfsA* itself (pBSKR-ybjCF & delnfsA_R, delnfsA_F & pBSKF_rimKR) (Table 18). **(A,B)** These primers possessed overhangs which added homologous regions to the 3' end of the DNA insert fragment that was upstream to *nfsA* and to the 5' end of the DNA insert fragment that was downstream to *nfsA*. **(C)** During Gibson assembly, these two DNA insert fragments will ligate with each other forming a DNA insert fragment that essentially a modified *nfsA* region (Δ*nfsA*). **(C)** The 5' and 3' ends of this DNA insert fragment also possessed homologous regions that corresponded to the 3' and 5' ends of a linearised vector fragment. **(A,B)** This linearised vector fragment was generated from PCR amplification of pBluescript kSII plasmid (pBlueKSII). **(D)** The DNA insert fragment will then be ligated with the linearised vector fragment to generate a plasmid with the modified *nfsA* region. **(E)** C2987 cells (NEB) were transformed with this plasmid via heat-shock. **(F)** The plasmid was then extracted from the transformed C2987 cells and sequenced to verify that the modified *nfsA* region has been properly assembled. **(F,G)** To generate the donor DNA template for CRISPR mutagenesis, the modified *nfsA* region within the plasmid was PCR amplified using primers that bind to end of the insert sequence (ybjC-144F & rimK+1100R) (Table 18). The PCR product was then purified before being used for CRISPR mutagenesis.

Gibson assembly was carried performed as per manufacturer's instructions using the NEB Gibson Assembly Master Mix. The assembled plasmids were chemically

transformed into NEB 5-alpha Competent C2987 *E. coli* cells by heat shock transformation (Section 2.3.1). The transformed cells were plated out on LB-Ampicillin plates and incubated overnight at 37°C. The plasmids of the transformed cells were extracted and sequenced to verify the modified *nfsA/nfsB* genetic regions contained within them.

#### 2.4.1.2 Constructing a kanamycin resistant version of pKD46 (pKD46-Kan)

The Tn5 neomycin phosphotransferase gene was PCR amplified from a pKD4 plasmid and used as the "insert" sequence. The plasmid/vector backbone was PCR amplified from a pKD46 plasmid with the ampicillin resistance gene (*ampC*) excluded. Gibson assembly and chemical transformation were carried out with LB-Kanamycin plates (Section 2.4.1).

### *2.4.2 Lambda-red/CRISPR two-step mutagenesis*

A typical CRISPR mutagenesis involves several key components as outlined in Table 19.

**Table 19 Components of CRISPR mutagenesis**

| | |
|---|---|
| **pCas9** | Synthesises the Cas9 endonuclease (in the presence of arabinose) that creates a double-stranded break (DSB) at a specific DNA locus when directed by single guide RNA (sgRNA). This incentivises the cell to use the donor DNA to replace the broken DNA locus via homologous recombination.<br>Contains *KanR* that encodes for kanamycin resistance (selection marker) |
| **pTargetF (pTRG)** | Synthesises sgRNA that targets a specific genetic region<br>Contains *aadA* that encodes for spectinomycin resistance (selection marker) |
| **Donor DNA** | Functions as the genome editing template that the cell will use to repair the double-stranded break created by Cas9 |

Instead of using a pTRG-expressed single guide RNA (sgRNA) that targeted *nfsA/nfsB*, a pTRG-expressed sgRNA that targeted a chloramphenicol resistance gene (*cat*) was used. This sgRNA was developed by Sim et al [314] and was known be highly specific and efficient. However, this technique necessitated a two-step mutagenesis protocol (Figure 25). The first step used lambda-red mediated recombination to replace *nfsA or nfsB* with *cat*, providing a unique genomic target for CRISPR mutagenesis (Figure 25). The second step used CRISPR mutagenesis to replace *cat* with the donor DNA that contained a modified version of the *nfsA/nfsB* region.

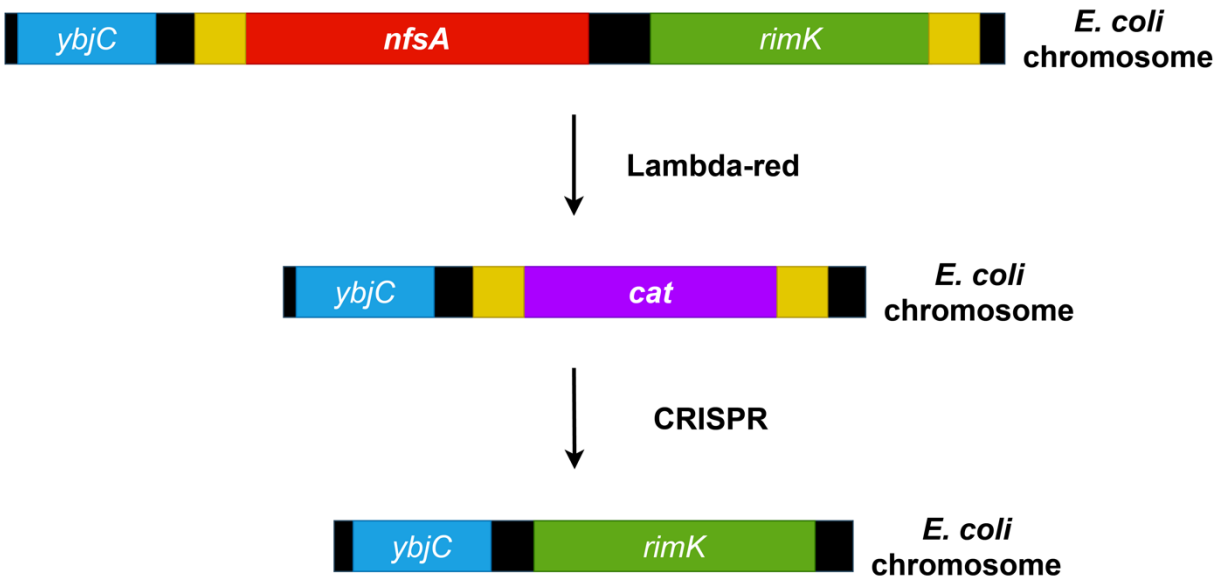



**Figure 25 Two-step mutagenesis overview.** Lambda-red recombination was first used to replace the region of interest (in this case *nfsA-rimK*) with *cat*. This provides a unique genomic target for CRISPR mutagenesis as the sgRNA in pTRG specifically targets this gene. The final outcome of CRISPR mutagenesis is the replacement of *cat* with an appropriate donor DNA (in this case r*imK*), producing a mutant that lacks *nfsA*

The main advantage of this method was that it combined the high specificity of lambda-red recombination with the high efficiency of CRISPR mutagenesis. Furthermore, using a single sgRNA construct eliminated the need to optimise multiple sgRNA constructs. However, this method required two rounds of mutagenesis, which could have increased the prevalence of secondary site mutations.

### 2.4.2.1 Lambda-red mediated recombination

Lambda-red mutagenesis was performed on 1646 BASE and W3110 isolates to replace the *nfsA* and *nfsB* regions with the chloramphenicol resistance (*cat*) gene (Figure 26). The primers used to generate the donor DNA for lambda-red mutagenesis are listed in Table 20. Prior to lambda-red mutagenesis, 1646 BASE and W3110 isolates were transformed via electroporation with the pKD46-Kan helper plasmid. pKD46-Kan is a modified pKD46 plasmid that had its ampicillin resistance (*bla*) gene replaced with a kanamycin resistance (*kan*) gene (Section 2.4.1.2). pKD46-Kan is a temperature-sensitive (curing temperature >37°C) helper plasmid that expresses the lambda-red system. To prevent unwanted recombination, the expression of the lambda-red proteins is tightly regulated by an arabinose-inducible pBAD promoter.

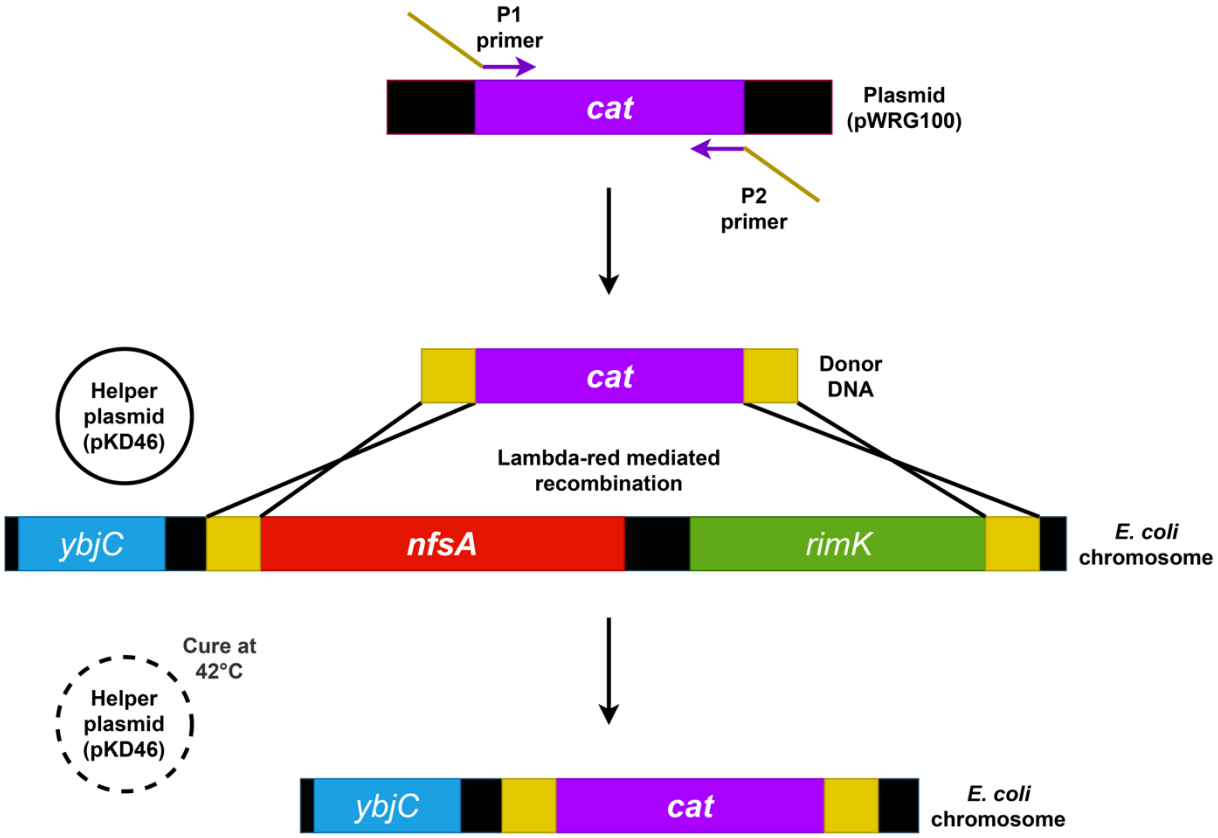



**Figure 26 Lambda-red mutagenesis where the *nfsA-rimK* region was replaced with a chloramphenicol resistant gene (*cat*).** The cat gene from the pWRG100 plasmid was PCR amplified using primers that possessed 50nt overhangs (Table 21B). The 5' end of the PCR product has a 50nt sequence that is homologous to the genetic region that is immediately upstream of *nfsA*. The 3' end has a 50nt sequence that is homologous to the genetic region that is immediately downstream of *rimK*. The host cell is transformed with a pKD46 helper plasmid, which carries the components of the lambda-red system. The lambda-red system mediates the recombination of the PCR product (donor DNA) with the *nfsA-rimK* region of the cell, replacing the *nfsA-rimK* region with *cat.* The *cat* gene functions as a selection gene and a sgRNA target for CRISPR mutagenesis.

Firstly, pKD46-Kan was electroporated into suitable isolates following the electroporation transformation protocol described in Section 2.3.2. A single transformed colony selected using kanamycin was inoculated in LB-Kanamycin liquid media and incubated overnight at 30°C and 160 rpm in a SANYO orbital incubator. The cells were prepared, electroporated with the appropriate donor DNA (Table 21), and selected following the protocols described in Section 2.3.2. Two mutants were generated using lambda-red mutagenesis: (i) *ΔnfsA-rimK::cat* and (ii) *ΔnfsB::cat*. Lambda-red mutagenesis was also successfully performed with the pCas9 plasmid as it carries the lambda-red system.

Table 20 Primers that were used to generate the donor DNA using the pWRG100 plasmid as the template.

| Primer No. | Primer Name | Sequence |
|---|---|---|
| 1570 | nfsA-26_pWRG_F | ttcagaaagagaaaaagataatgacgccaaccattgaacttatttgtggcCTAGACTATATTACCCTGTT |
| 1571 | nfsB-41_pWRG_F | taatctgctggcacgcaaaattactttcacatggagtctttatggatatcCTAGACTATATTACCCTGTT |
| 1572 | rimK+944_pWRG_R | aaaaagcgcaggcaaaaccatgatcagtaatgtgattgcgactaaccaccCGCCTTACGCCCCGCCCTGC |
| 1573 | nfsB+695_pWRG_R | tctgaggggaaatagccgggcagatgcccggcaagagagaattacacttcCGCCTTACGCCCCGCCCTGC |

Table 21 Lambda-red donor DNA

| Donor DNA | Targeted upstream region | Targeted downstream region |
|---|---|---|
| nfsA-pWRG100 | -76bp to -26bp upstream from the start of *nfsA* | +1100bp to +1150bp downstream from the start of *rimK* |
| nfsB-pWRG100 | -91bp to -41bp upstream from the start of *nfsB* | +695bp to +745bp downstream from the start of *nfsB* |

### 2.4.2.2 CRISPR mutagenesis

CRISPR mutagenesis was carried out following the protocol developed by Jiang et al [315]. The pCas9 plasmid was electroporated (Section 2.3.2) into isolates that had undergone successful lambda-red mediated recombination (Section 2.4.2.1). A colony was inoculated in LB-Kanamycin liquid media and incubated overnight in a SANYO orbital incubator set at 30°C and 160 rpm. Cells were prepared, electroporated with pTRG and donor DNA, and selection carried out on LB-Kanamycin-Spectinomycin plates at 30°C (Section 2.3.2).

Five colonies from each CRISPR reaction had the *nfsA* and *nfsB* regions sequenced to confirm the integration of the modified *nfsA/nfsB* region present in the donor DNA. The pTRG plasmid was cured by incubating the cells overnight at 30°C in LB liquid media with 0.5mM of isopropyl-β-D-thiogalactopyranoside (IPTG). The pCas9 plasmid was cured by incubating the cells overnight at 42°C on LB plates. Plasmid curing was confirmed by plating treated colonies on spectinomycin (pTRG) or kanamycin (pCas9) plates and incubating them at 30°C, with the absence of growth indicating successful plasmid curing. Multiple rounds of plasmid curing were occasionally required for some isolates to ensure complete plasmid removal from the population.

Double mutants with modified *nfsA* and *nfsB* regions were generated by curing pTRG from the *nfsA* single mutants, leaving pCas9 for lambda-red mediated recombination. Using pCas9 instead of pKD46-Kan for lambda-red mediated recombination saved time as it negated the need to cure pCas9 and electroporate pKD46. The *nfsB* region of the *nfsA* single mutants was replaced with the *cat* gene via lambda-red, which was replaced with a modified *nfsB* region via CRISPR. Plasmid curing was performed to remove the pTRG and pCas9 plasmids.

## 2.5 Cell culture techniques

### 2.5.1 RT4 cell culture maintenance

RT4 cell culture media was formulated as shown in Table 22.

**Table 22 RT4 cell culture media composition**

| | |
|---|---|
| **RPMI 1640 medium (Sigma-Aldrich R8758)** | 500 ml |
| **Foetal Calf Serum** | 50 ml |
| **L-Glutamine 200mM** | 5 ml |
| **HEPES buffer 1M** | 12.5 ml |
| **Penicillin-Streptomycin (10000ug/mL)** | 5 ml |

Cells were trypsinised every week with 250 000 cells being seeded into a new Corning 75 ml cell culture flask with 25 ml of fresh, pre-warmed RT4 cell culture media. Cell population was quantified using an automated cell counter (Nexcelom Bioscience Cellometer Auto T4). The media was changed 3 days after trypsinisation. Procedures were performed in a Class 2 laminar flow hood and cells were incubated at 37°C, 5% $CO_2$ in a SANYO incubator.

### 2.5.2 RT4 cell culture storage

RT4 cells were stored frozen in liquid nitrogen, trypsinised cells were quantified using an automated cell counter, pelleted using a centrifuge set at 150 x g for 3 minutes, and resuspended in an appropriate volume of freezing medium (RPMI 1640 with 10% foetal calf serum and 10% DMSO) to achieve a final cell density of $3*10^6$ cells/ml. 1 ml of the cell suspension was pipetted into labelled cryoprotective ampoules. The ampoules were labelled and then inserted into a freezing module which was transferred to a liquid nitrogen storage facility.

### 2.5.3 RT4 cell culture challenges

During routine trypsinisation, 100 000 cells were seeded into each well of a Corning 12-well culture plate filled with 1 ml of RT4 cell culture media. Cells were incubated in a SANYO incubator at 37°C, 5% $CO_2$ for 5-7 days before being used for challenges. The media was changed twice during this period, (i) 3 days after trypsinisation and (ii) on the day when the challenge was performed. Challenges were performed by pipetting an appropriate volume of agonist into each well and incubating the cells for 24 hours. On the following day, the media in each well was transferred to 1.5 ml microcentrifuge tubes before either analysed using ELISA or stored at -20°C for future analysis.

## 2.6 Protein purification

### 2.6.1 Flagella purification

A bacterial colony was inoculated in liquid media, incubated overnight in a SANYO orbital incubator set at the appropriate temperature and 160 rpm, and the final $OD_{600}$ of the culture was measured using a Pharmacia Biotech Ultrospec 1000 spectrophotometer. A 1 L liquid culture with a starting $OD_{600}$ = 0.02 was prepared from the overnight culture and incubated in the orbital incubator until the $OD_{600}$ = 0.6-0.8.

The culture was centrifuged at 9820 x *g* for 10 minutes, the supernatant discarded, and the cell pellet resuspended in 30 ml of ice-cold phosphate-buffered saline (PBS) solution. A blending stick set at 13500 rpm was used to shear the cells in the suspension for four minutes. The sheared cell suspension was centrifuged at 3960 x g for 30 minutes.

The supernatant was transferred into ultracentrifuge tubes, centrifuged at 100 000 x *g* for 1 hour, the supernatant discarded, and the pellet was resuspended in 1 ml of ice-cold PBS by repeated pipetting. The suspension was mixed with 29 ml of ice-cold PBS, centrifuged at 3960 x *g* for 15 minutes, transferred into ultracentrifuge tubes, and centrifuged at 100 000 x *g* for 1 hour. The supernatant was discarded and the pellet resuspended in 1ml of ice-cold PBS by repeated pipetting. The protein concentration of the mixture was quantified using a nanodrop before it was stored at -20°C.

### 2.6.2 SDS-PAGE

Polyacrylamide gels were prepared by pipetting the separating gel mixture (Table 23) into a casting tray until it was two-thirds full, at which stacking gel mixture was added. The gel was left undisturbed for an hour to polymerise. Samples were mixed 1:1 with 2X sample buffer (Table 23) in microcentrifuge tubes and heated at 100°C for 10 minutes. After heating, the tubes were immediately placed on ice until loaded onto the SDS-PAGE gel. Sigma-Aldrich mPAGE Color Protein Standard was used as the molecular weight marker.

The gel was electrophoresed at 25 mA (constant amperage) for 25 minutes or until the protein bands had migrated to the edge of the stacking gel. After this, the gel was electrophoresed at 40 mA (constant amperage) for 50 minutes or until the protein bands were close to the end of the gel. Once electrophoresis was completed, the gel was recovered, submerged in staining solution (Table 23), and placed on an oscillating rocker. After 30 minutes, the staining solution was discarded and replaced with a destaining solution (Table 23). After 30 minutes, the destaining solution was discarded and replaced with MilliQ water. The gel was incubated on an oscillating rocker overnight in MilliQ water before being photographed using an EPSON scanner.

Table 23 SDS-PAGE reagents.

| Reagent | Ingredients |
|---|---|
| Tricine Gel Buffer (200ml)<br>Titrate to pH 8.45 using ~ 45ml of 5M HCl and then filter it. | 56.32 g Tris Base<br>26.52 g Tris HCl<br>6 ml 10% SDS or 0.6 g SDS powder |
| 10x Anode Buffer (1L) | 212.8 g Tris Base<br>38.4 g Tris HCl |
| 10x Cathode Buffer (1L) | 121.1 g Tris Base<br>179.2 g Tricine<br>100 ml 10% SDSs |
| 2X Sample Buffer | 1 ml 1.25M Tris-HCl<br>4 ml 50% Glycerol<br>4 ml 10% sodium dodecyl sulphate (SDS)<br>1 ml 100%/14.3M b-Mercaptoethanol (or DTT)<br>500 mg of Bromophenol Blue (aka Coomassie Blue or Brilliant Blue G/R) |
| Separation gel (10%) | 3 ml separating acrylamide<br>5 ml tricine gel buffer<br>5 ml 50% glycerol<br>2 ml MilliQ water<br>75 µl 10% ammonium persulfate (APS)<br>7.5 µl 1,2-Bis(dimethylamino)ethane (TEMED) |
| Stacking gel (4%) | 250 µl separating acrylamide<br>755 µl tricine gel buffer<br>2.1 ml MilliQ water<br>25 µl APS<br>2.5 µl TEMED |
| Staining Solution (1L) | 500 ml 100% Ethanol<br>50 ml Glacial Acetic Acid<br>~450 ml MilliQ water<br>100 – 200 mg Brilliant Blue R/G |
| Destaining Solution (1L) | 500 ml 100% Ethanol<br>50 ml Glacial Acetic Acid<br>~450 ml MilliQ water |

## *2.7 Enzyme Linked Immunosorbent Assay (ELISA)*

IL8 concentration was quantified using an Invitrogen IL8 Human Uncoated ELISA kit following the manufacturer's instructions. Nunc MaxiSorp 96-well flat-bottom plates were used for all ELISA analyses. Plates were read using a BMG Labtech FLUOstar

OMEGA microplate reader at wavelengths of 450nm and 570nm. The 570nm values was subtracted from the 450nm values.

## 2.8 Minimum Inhibitory Concentration (MIC) Assay

MIC assays were performed using sterile Greiner Bio-One clear 96-well flat-bottom plates. MIC plates were prepared by adding 200 µl of an antibiotic solution at 2X the upper concentration to column 1 using a multi-channel pipette. 100 µl of Muller Hinton broth (MHB) was added to columns 2-12. 100 µl of the antibiotic solution in column 1 was transferred to column 2; this was repeated until column 11, where 100 µl of the mixture was discarded. This produced a decreasing antibiotic concentration gradient from left to right, with column 12 being antibiotic free.

Bacterial isolates were grown overnight in MHB. Subcultures were generated by diluting overnight cultures 1: $10^4$ with MHB in sterile glass tubes. For each isolate, 100 µl of subculture was pipetted into every well of a single row. The plates were covered with Breathe-Easy sealing membrane following manufacturer's instructions before being incubated statically for 16 hours at 37°C in a SANYO incubator . After incubation, the membrane was removed and plates were read using a BMG Labtech FLUOstar OPTIMA microplate reader set up for $OD_{600}$ measurement.

## 2.9 Growth curve assays and analyses

### 2.9.1 Microplate growth curve assay

Growth curve assays were performed using sterile Greiner Bio-One clear 96-well flat-bottom plates. Three independent colonies from each isolate to be tested were grown overnight each in 3 ml of MHB. The $OD_{600}$ of 10 random cultures was measured and averaged to generate the mean $OD_{600}$ value for the entire batch of isolates. This mean $OD_{600}$ value was used to determine the volume of overnight culture required to generate 2 ml subcultures (MHB) with a starting $OD_{600}$ = 0.02 for each isolate/colony.

A 200 µl volume of each subculture was pipetted into each well of a 96-well plate. Three wells acted as controls and each contained only 200 µl of MHB. The plate was covered with Breathe-Easy sealing membrane following the manufacturer's instructions before being inserted in a BMG Labtech FLUOstar OPTIMA microplate reader. The microplate reader had been configured to incubate the plate at 37°C, orbital shaking = 250 rpm, with $OD_{600}$ readings taken every 6.8 minutes for a period of 10.2 hours (90 timepoints). This whole process was repeated with MHB containing

0, 2, 4, 8, 16, 32, 64, and 128 µg/ml of nitrofurantoin. A fresh nitrofurantoin stock solution was prepared and used for each experiment.

### *2.9.2 Analysing growth curve data*

Growth curve data was imported into Microsoft Excel for reformatting and analysis. Reformatted data was also exported into R Studio so that the growth curve of each well could be visualised for quality-control purposes. Isolates that exhibited an aberrant growth curve pattern were excluded from the dataset.

The maximum growth rate of each isolate was determined by exploiting the maximum slope (gradient) across the entire 10-hour growth period after the isolate had reached an $OD_{600}$ = 0.2. $OD_{600}$ readings used for slope calculations were converted to natural log. The slope was calculated across the entire growth curve by using a moving "frame" of 8 timepoints. For example, the slope for t = 0 minute was derived from the $OD_{600}$ readings t = 0 minute (1$^{st}$ timepoint) and t = 47.6 minute (8$^{th}$ timepoint). This was repeated for t = 6.8 minute using $OD_{600}$ readings that corresponded to t = 6.8 minute (2$^{nd}$ timepoint) and t = 54.4 minute (9$^{th}$ timepoint). This pattern continued until t = 564.4 minute (83$^{rd}$ timepoint) since this is the last timepoint that fitted the moving "frame". The final outcome of all of these slope calculations was a slope dataset from which the maximum slope could be determined [316]. The doubling time of each isolate was calculated from the maximum slope using the following equation: $\frac{\ln 2}{maximum\ slope}$.

## *2.10 Growth competition assays*

Overnight bacterial cultures were prepared, and following $OD_{600}$ measurements were diluted to achieve a starting density of 25 CFU/ml in 6 ml of MHB media. The dilution was performed in two steps: the first being a 1:1000 dilution of the overnight culture in 1 ml of MHB media, followed by a second dilution in which an appropriate volume of the 1:1000 diluted culture was pipetted into 10 ml of MHB media to achieve a final $OD_{600}$ of $2.8*10^{-7}$. A 100 µl volume of the second dilution was pipetted into either 5.9 ml of MHB media (monoculture controls) or 5.8 ml of MHB media (where two different bacterial isolates were being tested). The competition cultures were incubated in a SANYO orbital shaker at 37°C and 160 rpm for 16 hours. Additionally, for each bacterial isolate, 100 µl of the second dilution was pipetted onto a CPSE plate, which

was incubated at 37°C for 16 hours. CPSE plates functioned to check the starting bacterial density for each diluted culture.

After 16 hours, the competition cultures were removed from the orbital shaker. Each culture was diluted ($1:10^5$ or $1:10^6$) and 100 µl of the diluted culture was plated onto a CPSE plate. The inoculated CPSE plates were incubated overnight at 37°C, photographed and the generated digital images were processed with an image processing package (Fiji, ImageJ distribution) to manually count the colonies.

## 2.11 Computer software and bioinformatic pipelines

### 2.11.1 Next-generation sequencing (NGS), assembly and annotation of bacterial isolates

Genomic DNA extracted from bacterial isolates was sequenced in an Illumina NextSeq500 at the Genomic Core Facility, Newcastle University. NextSeq DNA libraries were prepared using the Nextera Flex library prep and sequenced using a 2 x 150 bp High Output Kit, generating average fragment sizes of around 400 bp.

Raw genomic data were quality controlled using Trimmomatic to remove Illumina adapter sequences, low quality bases (phred quality < 20) and reads <50bp. Quality-controlled NGS data was assembled using Unicycler 0.4.8 with default settings. Assembled NGS data was annotated using Prokka v1.14.5, using its inbuilt Escherichia genus BLAST database.

### 2.11.2 In silico Multi Locus Sequencing Typing (MLST) of E. coli

The Achtman MLST scheme was used to type the sequenced AnTIC *E. coli* isolates. The genes used in this MLST scheme (*adk, fumC, gyrB, icd, mdh, purA, recA*) were extracted from the genome assembly of each isolate using BLAST and aligned using mafft. A maximum-likelihood phylogenetic tree was constructed from the MLST sequences of the AnTIC isolates and reference *E. coli* isolates from McNally et al. 2013 [317].

### 2.11.3 Core genome Multi Locus Sequence Typing (cgMLST) of E. coli

Core genome MLST (cgMLST) was performed on AnTIC *E. coli* isolates using chewBBACA, an open-source software suite that can create and evaluate cgMLST schemes [318]. The cgMLST scheme generated by chewBBACA for the AnTIC dataset consisted of 404 loci that were found in all of the assemblies. The allele calling data

was translated into a distance matrix table using phyloviz. The distance matrix table was then converted into a dendrogram in RStudio using the ape package.

## 2.12 Ethical approval

Permission to use clinical isolate and data from the AnTIC and BUTI clinical studies was derived from prior ethical approvals (Ethics: 19/NS/0024, IRAS Project ID: 243903, Ref: 2586/2016)

## 2.13 Data presentation and statistics

### 2.13.1 Data presentation

Draw.io (https://app.diagrams.net/) was used to generate flowcharts, gene maps, and various other figures. Statistical figures were generated using Microsoft Excel. These graphical objects, as well as those from other published works were further processed with Adobe Illustrator, saved as pdf files, and exported as png files at a resolution of 600 dots per inches (dpi) for use in this thesis.

### 2.13.2 Statistics

Fisher's exact test analyses were performed in Microsoft Excel using the Real Statistics Resource Pack add-on (https://www.real-statistics.com). T-test and growth curve analyses were performed in Microsoft Excel using the inbuilt functions, "t.test" and "slope" respectively. Quality control of growth curve data was performed by plotting the raw growth curve data in RStudio using inbuilt functions.

# Chapter 3 The interplay between antibiotics and bacterial colonisation in CISC patients

## *3.1 Introduction to the AnTIC dataset*

AnTIC was an open label randomised controlled trial that assessed the efficacy of antibiotic prophylaxis against non-prophylaxis in reducing the incidence of symptomatic UTIs within clean intermittent self-catheterised (CISC) patients over a period of 12 months [14]. For CISC patients, the catheter is only present in the lower urinary tract during micturition, unlike patients with an indwelling catheter where the catheter is present at all times [319].

Four hundred and four patients with a recent history of symptomatic recurrent UTIs were recruited into the trial and randomly assigned to either the non-prophylaxis (n=201 patients) or prophylaxis arm (n=203 patients). Patients in the prophylaxis arm were prescribed once-a-day antibiotic doses which were taken daily during the study period (Figure 27). The prophylaxis arm consisted of three subgroups, each of which were given different prophylactic antibiotics: cefalexin (250 mg, n=33 patients), nitrofurantoin (50 mg, n=71 patients), and trimethoprim (100 mg, n=92 patients). There were 7 patients in the prophylaxis arm who did not have any records of their assigned prophylactic antibiotic. These patients and their associated data (20 urine samples) were excluded from the investigations reported in this chapter.

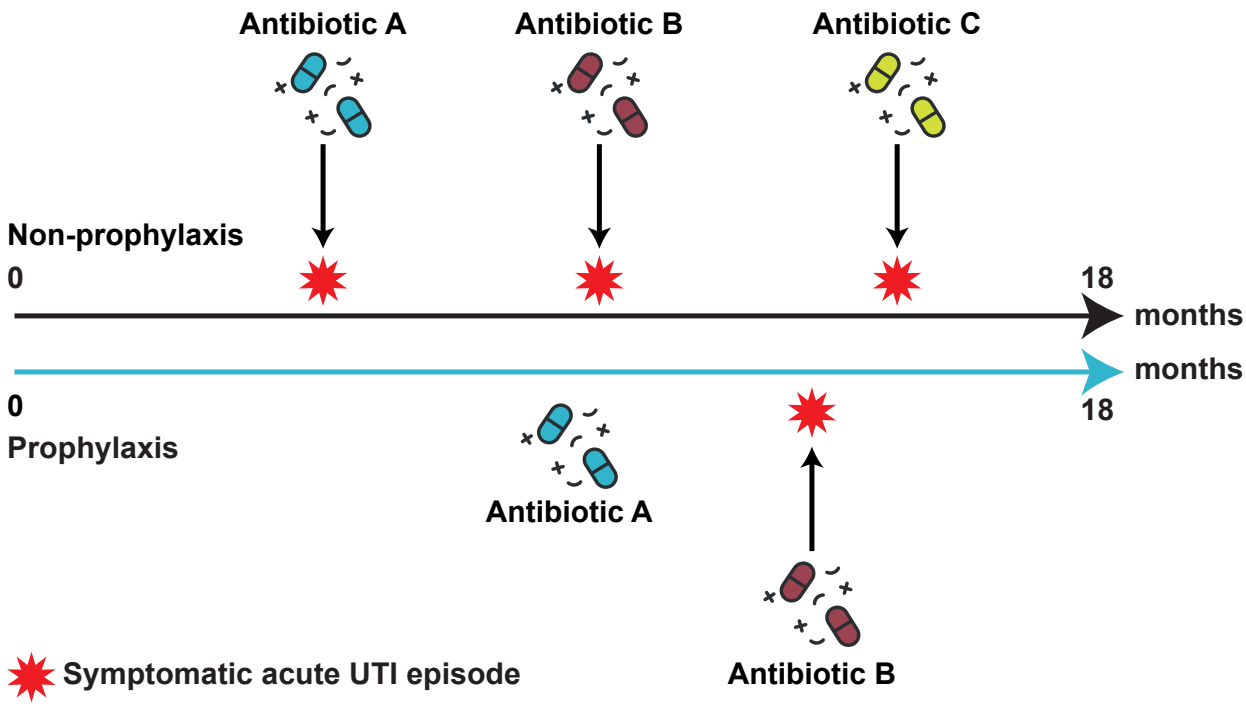



**Figure 27 AnTIC study arms.** Patients in the non-prophylaxis arm were only prescribed antibiotics as when needed, usually during an episode of symptomatic UTI. Patients in the prophylaxis arm were assigned a specific antibiotic which they were expected to take on a daily basis during the study period. In the event that a patient in the prophylaxis arm experienced an episode of symptomatic UTI, the prophylactic treatment will be temporarily halted, and a different antibiotic will be prescribed to that patient on a short-term basis. Once the symptomatic acute UTI episode has been resolved, the prophylactic treatment will be resumed.

The trial design (Figure 28) meant that each patient submitted a urine sample at five different timepoints (0, 3, 6, 9, 12 months) during the study period. When a patient experienced a symptomatic UTI episode, which necessitated antibiotics, the patient was required to submit a urine sample before commencing antibiotic treatment) Perianal swabs were also submitted at three different time-points (0, 6, 12 months). Zero (0)-month samples were regarded as the baseline (BASE) sample for each patient. At the end of 12 months, patients were given the option of participating in the study for a further 6 months. Patients who accepted this extension submitted both a urine and a perianal sample, which were taken at 18 months and reflected asymptomatic infection periods.

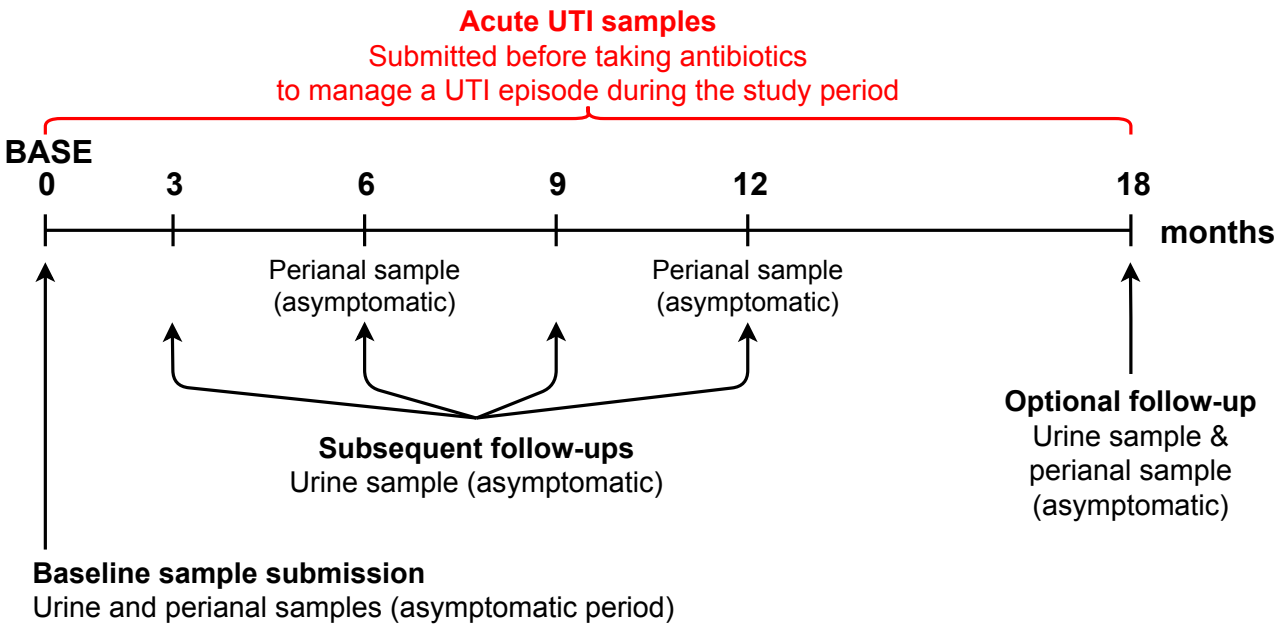



**Figure 28 AnTIC sample submission timeline.** Participants were expected to submit a urine sample taken during an asymptomatic period every 3 months, starting on the 1st day of their participation (Month 0, 3, 6, 9, 12). Should the participant experience an episode of symptomatic UTI, the participant was expected to submit a urine sample before starting antibiotic treatment. Participants were also expected to submit a perianal sample taken during an asymptomatic period every 6 months, starting on the 1st day of their participation. (Moth 0, 6, 12). Participants who accepted a 6-month extension were expected to submit a urine and a perianal sample taken during an asymptomatic period at the end of that extension (Month 18)

The main finding of the AnTIC study was that CISC patients participating in the prophylaxis arm of the study experienced one less incidence of symptomatic UTI compared to the non-prophylaxis patients (a 48% reduction in UTI frequency) [14]. However, this came at the cost of an increased frequency of resistance against the antibiotic used for prophylaxis. While the authors proposed reasons for the reduction in UTIs, e.g. host tolerance to microbial colonisers, the trial did not address the mechanism(s) by which antibiotic prophylaxis achieved this beneficial effect. In this chapter, curated AnTIC data was further exploited to investigate the impact of antibiotic prophylaxis on host-microbial interactions with the focus on microbial colonisation.

## *3.2 Revisiting the AnTIC dataset*

Each patient in the AnTIC study was assigned a unique patient ID which was associated with the following information: (i) patient age, sex, and underlying health conditions, (ii) assigned treatment arm, (iii) urine sample records, (iv) acute antibiotic prescriptions, and (v) trial outcome data (Figure 29). Due to the longitudinal nature of the study, each patient ID was expected to link to urine sample records associated with it. However, the number of urine samples associated with each patient was

variable and depended upon (i) the length of the patient's participation in the study, (ii) the number of symptomatic UTI episodes experienced by each patient and treated with antibiotics, and (iii) patient adherence to sample submission protocols [320].

Each urine sample record contained the following information: (i) submission date, (ii) whether it was either a routine asymptomatic sample or a sample taken during a symptomatic UTI episode, (iii) the dominant bacterial species isolated from the sample (if positive), and (iv) the antimicrobial resistance (AMR) profile of the dominant bacterial species (Figure 29). Positive urine samples were defined as samples that had no more than two bacterial species with each species having a population density of at least $10^4$ CFU/ml [320]. Samples that did not meet this requirement were classified as negative samples. To ensure standardisation, all submitted urine samples were analysed by central diagnostic laboratories including one based in the Freeman Hospital, Newcastle upon Tyne [320]. Any *E. coli* isolates identified in the urine samples were stored at -80°C and the genomes of 47 isolates were sequenced. This AnTIC database of sequenced genomes was used alongside patient data to investigate and explain potential interactions between antibiotic therapy and microbial colonisation of the lower urinary tract.

For some patients, there were occasions during the study when they were prescribed short-course antibiotics by their respective general practitioners (GPs) following consultation about a putative symptomatic acute UTI. These GPs were not involved in the study, but had full autonomy with regards to UTI diagnosis and antibiotic prescription [320]. Information, antibiotic type, start and end date of treatment, relating to these short-course antibiotic therapies was obtained from patients or relevant clinical records whenever possible before being recorded on the AnTIC database (Figure 29).

Using the AnTIC database and a curated selection of clinical isolates, the aims of the following studies were to explore potential interactions between antibiotic therapy and microbial colonisation of the lower urinary tract in CISC patients.

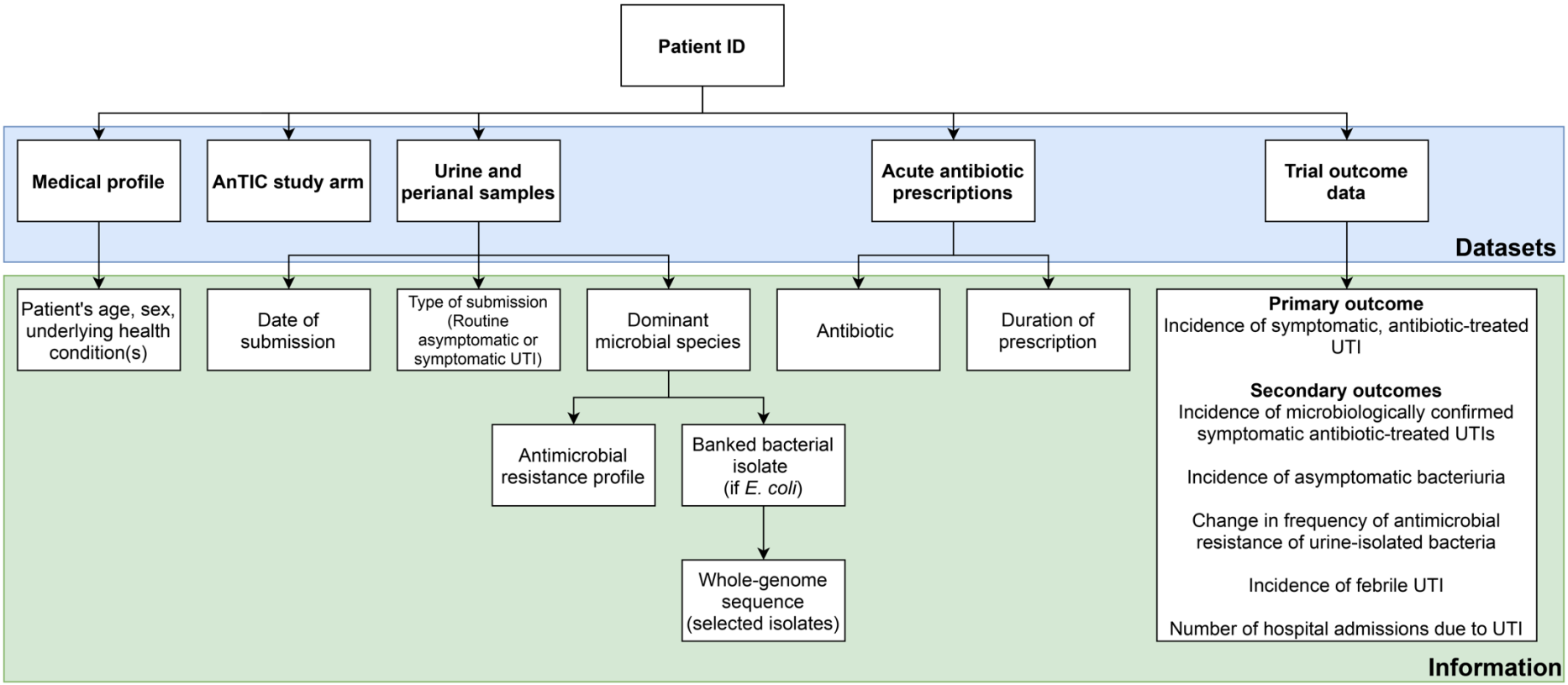



**Figure 29 AnTIC dataset summary.** Each patient was assigned a unique patient ID to anonymise the patient's real-world identity. Each patient ID is associated with a number of databases (Medical profile, AnTIC study arm, Urine and perianal samples, Acute antibiotic prescriptions, Trial outcome data) (Blue box). The "medical profile" dataset contains information regarding a patient's age, sex, and any underlying health conditions. "AnTIC study arm" dataset contains information about which study arm the patient was assigned to as well the assigned antibiotic prophylaxis for patients assigned to the prophylaxis study arm. The "urine and perianal samples" dataset contains information for each of the submitted urine/perianal sample. Each sample in that dataset had the following information associated with it: date of submission, type of submission (whether it was a routine asymptomatic sample or a symptomatic UTI sample taken before starting antibiotic treatment), and the dominant microbial species isolated from that sample (if present). If a microbial species is present in a sample, the antimicrobial resistance profile of that species will be reported. If the microbial species was *E. coli*, it was then stored at -80°C for future research. The "acute antibiotic prescriptions" dataset contains information regarding which short-course antibiotic was prescribed for each patient and the duration of the short-course prescriptions. "Trial outcome data" datset contains the following information for each patient ID: (i) the incidence of symptomatic, antibiotic-treated UTI (primary outcome), (ii) incidence of microbiologically confirmed UTI episodes that were treated with antibiotics, (iii) incidence of asymptomatic bacteriuria, (iv) change in the frequency of antimicrobial resistance of urine-isolated bacteria, (v) incidence of febrile UTI, and (vi) number of hospital admissions associated with UTI

### *3.2.1 Does antibiotic prophylaxis reduce UTI incidence by reducing the overall incidence of bacterial colonisation?*

There is a suggestion from the literature that antibiotic prophylaxis reduces UTI incidence by promoting a sterile environment in the lower urinary tract [58,321]. To test this viewpoint, patient urine samples were stratified and analysed according to the study arms of non-prophylaxis and prophylaxis. Urine samples associated with the prophylaxis arm were further stratified into three sub-arms- cefalexin, nitrofurantoin, or trimethoprim – based on the assigned prophylactic antibiotic treatments of the patients associated with these urine samples. The incidence of bacterial colonisation for each arm was calculated by dividing the total number of positive urine sample by the total number of urine samples (positive and negative) associated with each arm/sub-arm. Positive urine samples were associated with bacterial colonisation whereas negative urine samples were presumed to be associated with a sterile urinary environment.

The incidence of bacterial colonisation for the prophylaxis arm was significantly reduced when compared to the non-prophylaxis arm (34.4% vs 48.2%, $p<0.001$) (Figure 30). Comparison of data from the three prophylaxis sub-arms with the non-prophylaxis arm produced comparable results. However, when the prophylaxis sub-arms were compared to each other, the nitrofurantoin sub-arm had a significantly lower incidence of bacterial colonisation compared to the trimethoprim sub-arm (27.3% vs 39.5%, $p<0.001$).

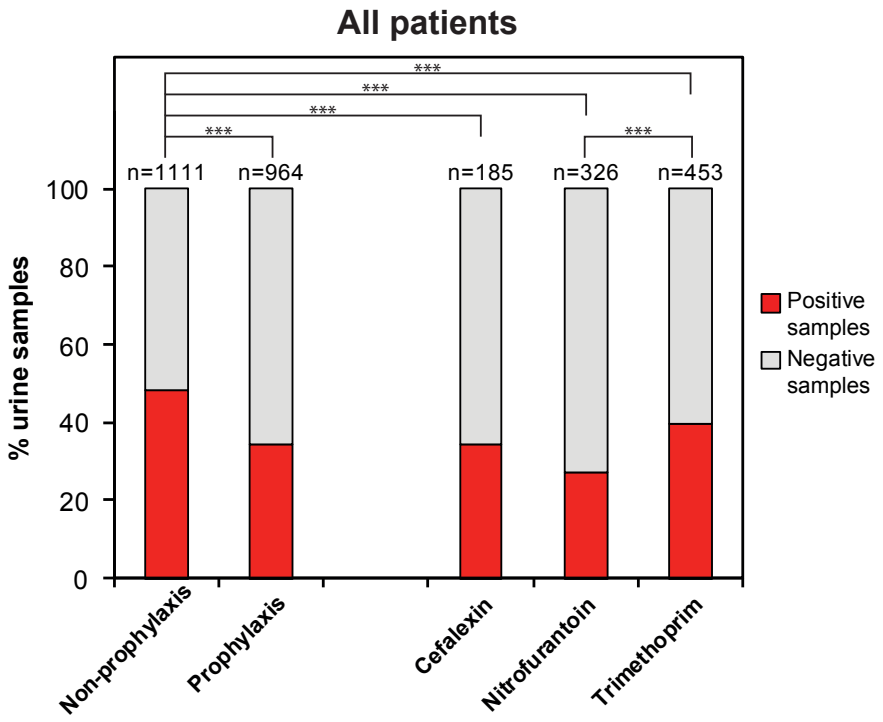



**Figure 30 Percentage of positive and negative urine samples across study arms.** The number of positive urine samples was significantly fewer in the prophylaxis when compared against the non-prophylaxis arm. All three prophylaxis sub-arms had significantly fewer positive urine samples compared to the non-prophylaxis arm. The nitrofurantoin prophylaxis sub-arm had significantly fewer positive urine samples compared to the trimethoprim sub-arm. ***= p<0.001

The difference in the total number of urine samples between the non-prophylaxis (n=1111) and prophylaxis (n=964) arms may have introduced a bias, so was examined using a Fisher's exact test. However, there was no significant difference in the total number of urine samples between the different components (Table 24).

**Table 24 A summary of the Fisher-test analysis that was performed to determine whether there was any significant difference in the number of urine sample between the different trial components in AnTIC.**

| Comparisons | Fisher-test significance |
|---|---|
| Non-prophylaxis vs prophylaxis | 0.297 |
| Non-prophylaxis vs cefalexin prophylaxis | 1.000 |
| Non-prophylaxis vs nitrofurantoin prophylaxis | 0.240 |
| Non-prophylaxis vs trimethoprim prophylaxis | 0.402 |
| Cefalexin prophylaxis vs nitrofurantoin prophylaxis | 0.432 |
| Cefalexin prophylaxis vs trimethoprim prophylaxis | 0.590 |
| Nitrofurantoin prophylaxis vs trimethoprim prophylaxis | 0.727 |

Comparisons were made between the different study arms (non-prophylaxis, prophylaxis) and sub-arms (cefalexin, nitrofurantoin, trimethoprim prophylaxis). None of the comparisons were found to be statistically significant ($p<0.05$), meaning that there was no significant difference in the number of urine samples between the various trial components.

In summary, the significantly lower incidence of bacterial colonisation observed in the prophylaxis arm samples suggested that antibiotic prophylaxis reduced UTI incidence by promoting a sterile urinary environment. However, the incidence of bacterial colonisation within the prophylaxis arm still remained at a substantial level (34.4% positive samples). This raised the question of whether there were other mechanisms by which antibiotic prophylaxis was functioning to reduce UTI incidence.

### *3.2.2 Does antibiotic prophylaxis reduce UTI incidence by selectively reducing the incidence of certain bacterial species/genera?*

It has been suggested that antibiotic prophylaxis reduces UTI incidence by selectively inhibiting the colonisation of certain microbial species that are associated with symptomatic UTI, primarily *E. coli* [322]. To investigate this hypothesis, the dominant coloniser present in every positive urine sample was allocated to one of four groups based on their genus: *Escherichia, Enterococcus, Klebsiella*, and “others” (Table 25). The AnTIC dataset made this possible because all positive urine samples were subjected to urine culture analysis and the species of the dominant pathogen was identified and recorded (Figure 29). The first three groups represent bacterial genera whose member(s) are considered to be major uropathogens, making up ~86% of uncomplicated UTI cases and ~84% of complicated UTI cases [166]. The fourth group (“Others”, Table 25) contains bacterial species/genera that are considered to be minor uropathogens, making up ~14% of uncomplicated UTI cases and

~16% of complicated UTI cases [166]. This group will now be referred to as “minor uropathogens”.

**Table 25 The four major bacterial groups and their constituent bacterial species.**

| Group | Bacterial species |
|---|---|
| *Escherichia* | *E. coli* |
| *Enterococcus* | *E. faecalis, E. faecium, and other Enterococcus species* |
| *Klebsiella* | *K. oxytoca, K. pneumoniae, and other Klebsiella species* |
| Others<br>“Minor uropathogens” | *Citrobacter genus, Enterobacter genus, M. morganii, Proteus genus, Pseudomonas genus, S. liquefaciens, S. marcescens, S. epidermidis, S. agalactiae, C. albicans, C. aurimucosum* |

Antibiotic prophylaxis significantly affected colonisation of the urinary tract by *Escherichia, Enterococcus* and the minor uropathogens. *Klebsiella* colonisation was unaffected by antibiotic prophylaxis as there was no significant difference in the incidence of *Klebsiella* between the treatment arms (Figure 31c). Antibiotic prophylaxis was associated with significantly lower *Escherichia* colonisation (47.3% vs 62.7%, $p<0.001$, Figure 31a) and significantly higher colonisation by minor uropathogens (29.2% vs 9,7%, $p<0.001$, Figure 31d). Comparison of each of the prophylaxis sub-arms against non-prophylaxis yielded similar observations for both *Escherichia* and the minor uropathogens. Of the three antibiotics that were used for prophylaxis, cefalexin was associated with the greatest decrease in *Escherichia* and the greatest increase in minor uropathogens, followed by nitrofurantoin and trimethoprim.

There was no difference in the incidence of *Enterococcus* between the non-prophylaxis and prophylaxis arms. However, cross-comparisons between the non-prophylaxis and prophylaxis sub-arms revealed a significant decrease in *Enterococcus* colonisation in the nitrofurantoin sub-arm (1.1 vs 8.0%, $p<0.01$, Figure 31b). Nitrofurantoin prophylaxis resulted in the greatest decrease in *Enterococcus* followed by trimethoprim and cefalexin prophylaxis.

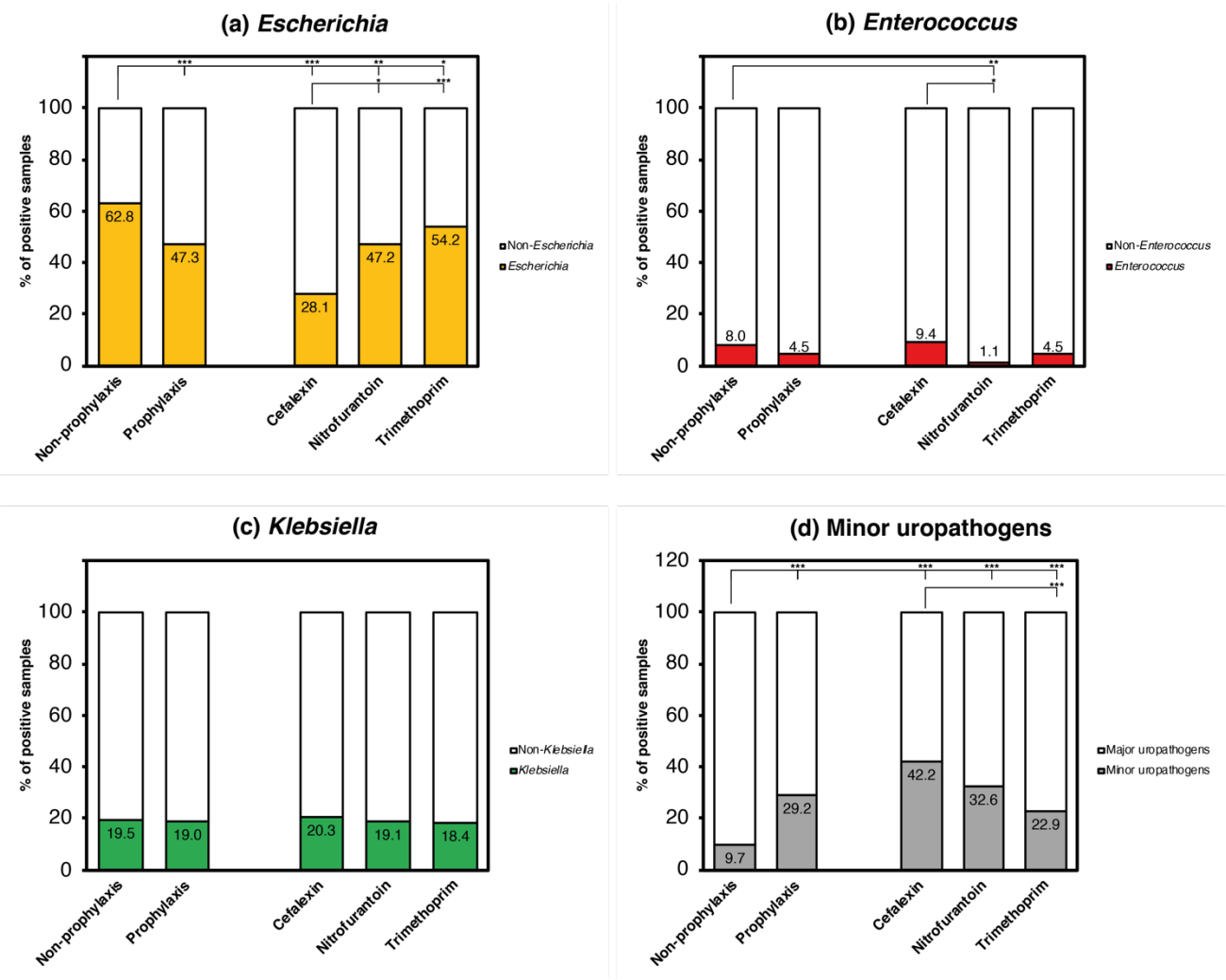



**Figure 31 Incidence of (a) *Escherichia*, (b) *Enterococcus*, (c) *Klebsiella,* and (d) Minor uropathogens across study arms.** The numbers in each coloured bar represents the percentage of positive samples in which the stated bacterial group was identified. * p<0.05. ** p<0.01, *** p<0.001

Antibiotic prophylaxis was effective at reducing *Escherichia* colonisation, but this was accompanied by increased colonisation with minor uropathogens. In contrast, *Klebsiella* and *Enterococcus* were largely unaffected by antibiotic prophylaxis, with *Enterococcus* only being affected by a specific antibiotic (nitrofurantoin). From these findings, it could be inferred that antibiotic prophylaxis may reduce UTI incidence by altering the urinary microbiome such that *Escherichia* colonisation is inhibited while colonisation by minor uropathogens is promoted.

Analysis of the isolates' AMR profile data (Figure 29) provided some insights that could explain this outcome. Minor uropathogens were more likely to be resistant against cefalexin (50% vs 11.8%) and nitrofurantoin (46.9% vs 8.6%) when compared to *Escherichia* (Figure 32). Hence, it was no surprise that antibiotic prophylaxis in AnTIC selectively inhibited *Escherichia* while favouring minor uropathogens.

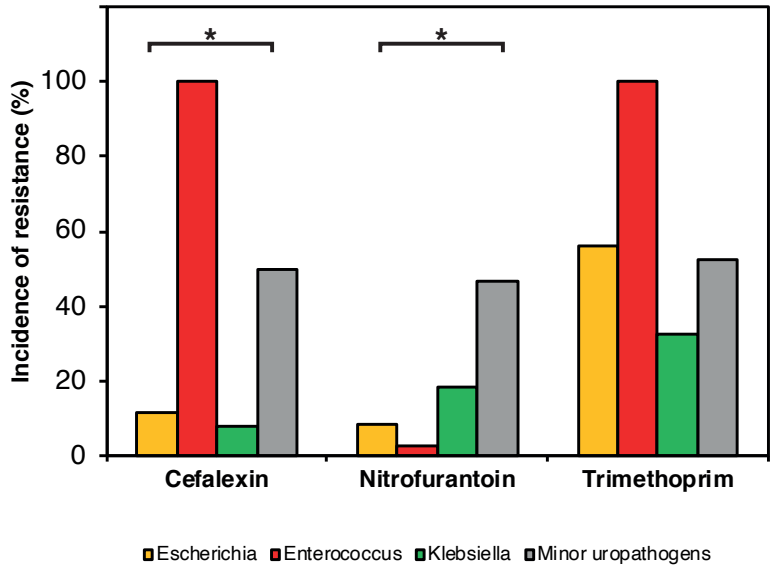



**Figure 32 Incidence of resistance against the prophylactic antibiotics used in AnTIC by bacterial genera**

These analyses argue that reducing bacterial load (Section 3.2.1) and altering microflora composition together are valid mechanisms by which antibiotic prophylaxis is able to reduce the incidence of UTIs. These investigations revealed potential mechanisms by which antibiotic prophylaxis reduced the incidence of UTI in AnTIC patients. However, the key focus was on the overall impact of antibiotic prophylaxis across the entire study without taking individual patients into account. Hence the next step was to explore microbial colonisation of individual patients within each study arm.

### *3.2.3 What was the overall impact of antibiotic prophylaxis on microbial colonisation within individual patients?*

The longitudinal nature of the AnTIC study meant that each patient donated a number of urine samples at different time points throughout the study (Figure 28). These urine samples were therefore examined to determine the patterns of microbial colonisation in each patient. A key criterion for patient inclusion in this analysis was the donation of at least two consecutive urine samples that were either (i) negative or (ii) had the same microbial group in both. The reasons underpinning these criteria are (i) that a longitudinal analysis requires at least two samples collected at different timepoints and (ii) to be confident that a bacterial group was present in a patient across a specific time period, two consecutive urine samples with the same bacterial group were needed to determine the beginning and the end of the colonisation period for that bacterial group.

Analyses of the associated data from these urine samples resulted in each patient being characterised by a “dominant coloniser”. This term described the bacterial group showing the longest period of continued colonisation, with case examples shown in Figure 33. Following these analyses, the overall impact of antibiotic prophylaxis on microbial colonisation was investigated by comparing the percentage of patients with a given “dominant coloniser” between the non-prophylaxis and prophylaxis arms. A key study limitation however, was the assumption that there were no changes in microbial colonisation during the periods between urine sample collections.

Of 404 patients, only 329 patients met the inclusion criteria described above and visually defined in Figure 33. Although each patient was expected to have at least five urine sample, there were instances when patients dropped out of the study early (Table 26) and/or where samples were unavailable. Figure 34 provides a summary of the number of patients which were associated with a given number of urine samples, ranging from 0-20. The overall data showed that 343 patients were associated with 5 or more urine donations, comprising of 176 non-prophylaxis and 167 prophylaxis patients (Table 26).

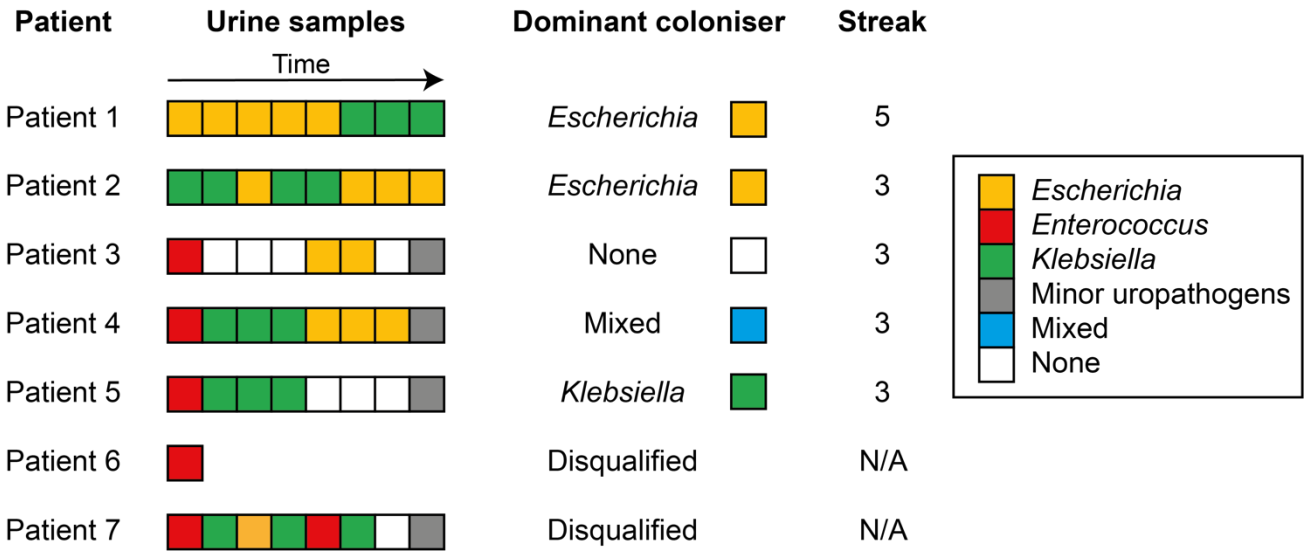



**Figure 33 A summary of the urine analysis used to determine the "dominant coloniser" for each patient.** Each box represents a urine sample, the colour denotes the isolated bacterial species. In the case of patient 1, the dominant coloniser was Escherichia as it had the longest period of consecutive colonisation (5 consecutive samples). For Patient 2, the dominant coloniser was Escherichia as it had the longest period of consecutive colonisation at 3 samples versus Klebsiella's 2 consecutive samples. For patient 3, the dominant coloniser was "none" because "none" had the longest period of consecutive colonisation. In cases where there were more than one dominant colonisers (Patient 4), the dominant coloniser will be "Mixed". If there is a tie between "None" and another bacterial species, as in the case of patient 5, the dominant coloniser will the bacterial species. Patient 6 had only a single urine sample was thus disqualified for not meeting the criterion (donated of at least two consecutive urine samples that were either (i) negative or (ii) had the same microbial group in both). Patient 7 was also disqualified for not meeting the criterion due to lacking two consecutive urine samples with that were either negative or had the same bacterial group. The streak column shows the longest period of consecutive samples associated with the dominant coloniser for each patient.

**Table 26 Breakdown of patient attrition by study arm during the AnTIC study.**

| No. of urine samples (Study duration) | Total patients | Non-prophylaxis patients | Prophylaxis patients |
|---|---|---|---|
| 1 (Start) | 404 | 201 | 203 |
| 2 (3 months) | 378 | 188 | 190 |
| 3 (6 months) | 361 | 180 | 181 |
| 4 (9 months) | 347 | 171 | 176 |
| 5 (12 months) | 343 | 176 | 167 |

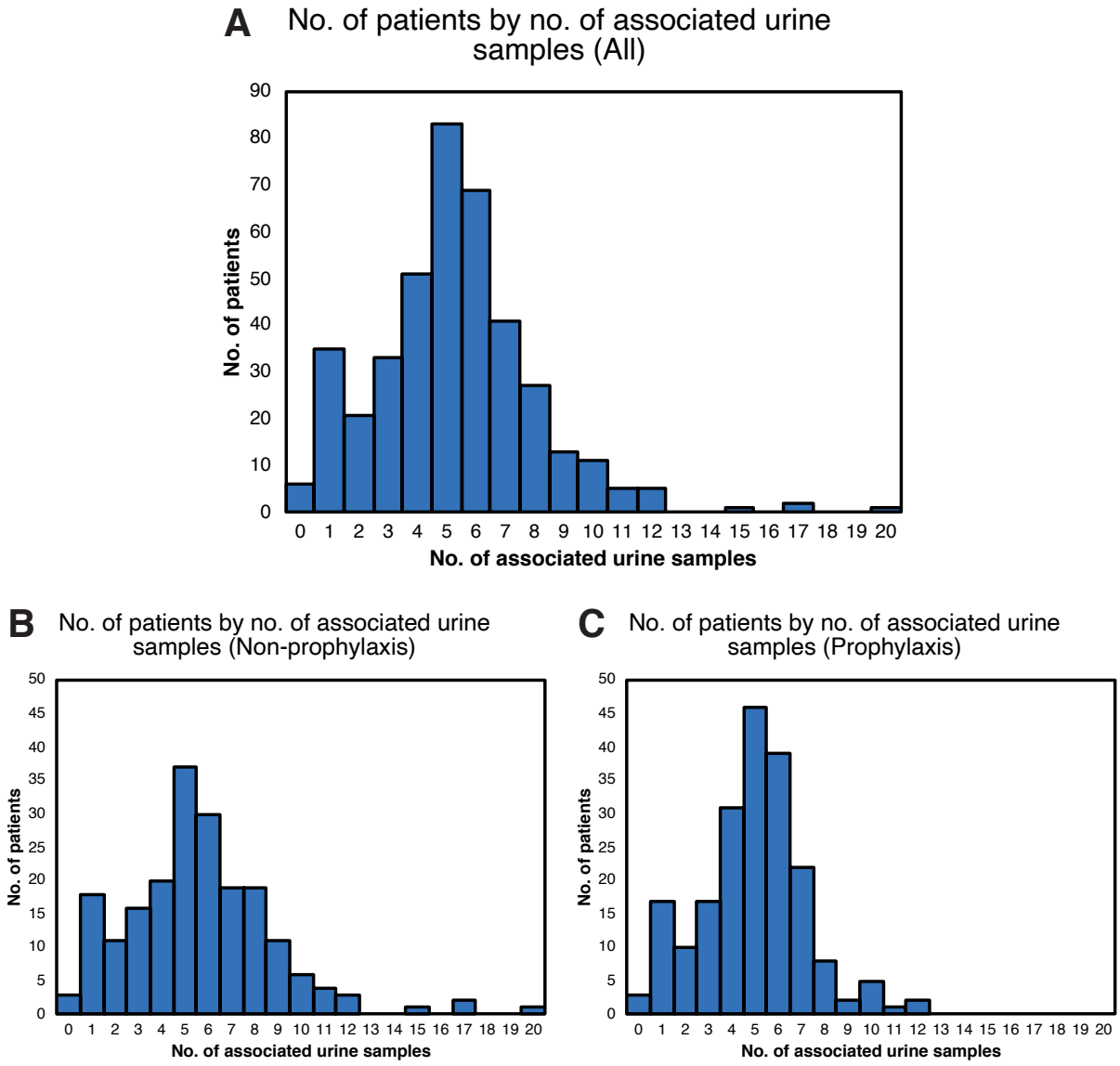



**Figure 34 Histogram showing the number of patients associated with a given number of samples from 0-20 samples.** Patients who experienced symptomatic acute UTI during the study donated additional urine samples, which were taken during the symptomatic period, causing some patients to have more than 5 associated urine samples.

The following sections will discuss the data generated from this urine analyses to determine whether the effects of antibiotic prophylaxis on microbial colonisation remained constant when the level of analysis was shifted from study arms to the individual patients.

In line with earlier findings, it was hypothesised that the prophylaxis arm would have the following differences when compared to the non-prophylaxis arm:

1. Significantly more patients with a dominant coloniser
2. Significantly fewer patients with *Escherichia* as the dominant coloniser
3. Significantly more patients with minor uropathogens as the dominant coloniser

The impact of antibiotic prophylaxis on microbial colonisation was previously investigated in Sections 3.2.1 and 3.2.2 using study arms (non-prophylaxis, prophylaxis) as the unit of analysis. However, the findings from these investigations cannot be applied to the individual patients (which is a lower-level unit of analysis) within the study arms as doing so constitutes an ecological fallacy. An ecological fallacy is a “failure in reasoning that arises when an inference is made about an individual based on aggregate data for a group” (Encyclopaedia Britannica) [323]. Inferences made about microbial colonisation within individual patients cannot be deduced from inferences based on the study arms to which these patients belong. The following investigations resolve the potential issue of ecological fallacy by characterising microbial colonisation for each patient (Figure 33) and then using the generated data to determine the difference in the number of patients with specific “dominant colonisers” between the study arms to gauge the impact of antibiotic prophylaxis on microbial colonisation.

#### 3.2.3.1 Was antibiotic prophylaxis associated with a significant increase in the number of patients with no dominant colonisers?

Patients with no dominant colonisers were defined as microbiologically negative (illustrated by Patient 3; Figure 33). Analyses showed that there was a significant increase in patients with no dominant colonisers in the prophylaxis arm compared to non-prophylaxis (72.6% vs 52.7%, $p<0.001$) (Figure 35). Comparison of the prophylaxis sub-arms against the non-prophylaxis arm revealed that the nitrofurantoin and trimethoprim sub-arms had significantly more patients with no dominant colonisers (Nitrofurantoin: 77.2% vs 52.7%, $p<0.01$, Trimethoprim: 70.3% vs 52.7%, $p<0.05$) (Figure 35). No significant differences were observed when the prophylaxis sub-arms were compared to each other.

These data suggested that antibiotic prophylaxis was associated with a significant increase in the incidence of patients with no dominant colonisers, consistent with earlier findings.

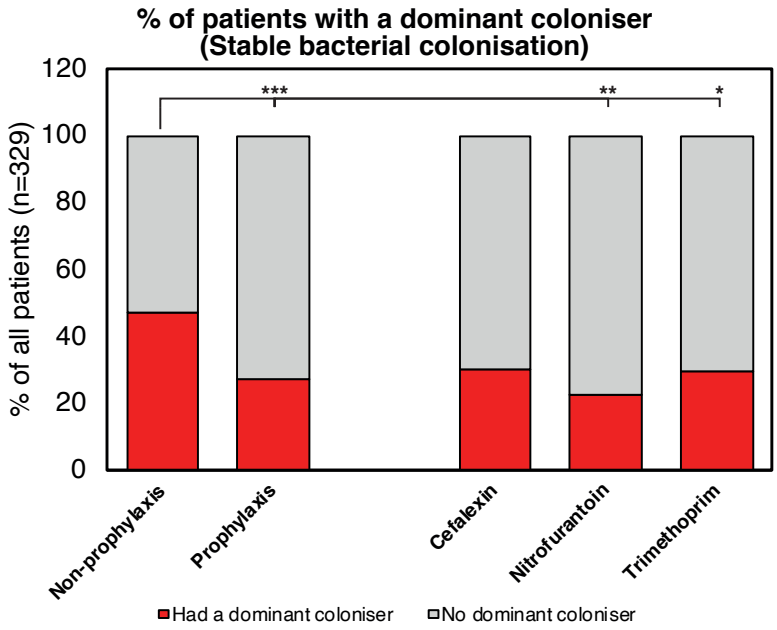



**Figure 35 Percentage of patients with a dominant coloniser (stable bacterial colonisation).** Patients 1,2,4 and 5 from Figure 33 represent case examples of patients that "had a dominant coloniser" (red bar). Patient 3 from Figure 33 represent a case example of patients that had "no dominant coloniser" (grey bar).

### 3.2.3.2 Was antibiotic prophylaxis associated with significant changes in genera-specific microbial colonisation?

In these analyses, only patients that had a dominant coloniser (n=123) were considered. The incidence of patients with *Escherichia* as the dominant coloniser was lower in the prophylaxis arm when compared to the non-prophylaxis arm, with the difference between both arms approaching significance (51.1% vs 69.2%, p=0.054, Figure 36A). Interestingly, comparing the prophylaxis sub-arms to the non-prophylaxis arm revealed that only cefalexin prophylaxis was associated with significantly fewer patients with *Escherichia* as the dominant coloniser (30% vs 69.2%, p=0.03, Figure 36A). No significant differences were observed when the prophylaxis sub-arms were compared to each other (Figure 36A)

When considering minor uropathogens as the dominant coloniser (Figure 36D), the prophylaxis arm data was inverted when compared to *Escherichia* (28.9% prophylaxis vs 3.8% prophylaxis, p<0.001, Figure 36D). Comparison of the prophylaxis sub-arms with the non-prophylaxis arm revealed that all three prophylaxis sub-arms had significantly more patients with minor uropathogens as the dominant coloniser. No significant differences were observed when the prophylaxis sub-arms were compared to each other. Antibiotic prophylaxis did not significantly impact colonisation by *Enterococcus* and *Klebsiella* (Figure 36B,C).

A key finding from all these analyses was that a decrease in stable *Escherichia* colonisation was accompanied by an increase in urinary tract colonisation by minor uropathogens in patients on prophylaxis, consistent with earlier findings. However, a subset of patients in the prophylaxis arm (n=23) displayed stable *Escherichia* colonisation, suggesting that the interaction between antibiotic prophylaxis and *Escherichia* colonisation amongst AnTIC patients could involve more than just the inhibition of *Escherichia* colonisation.

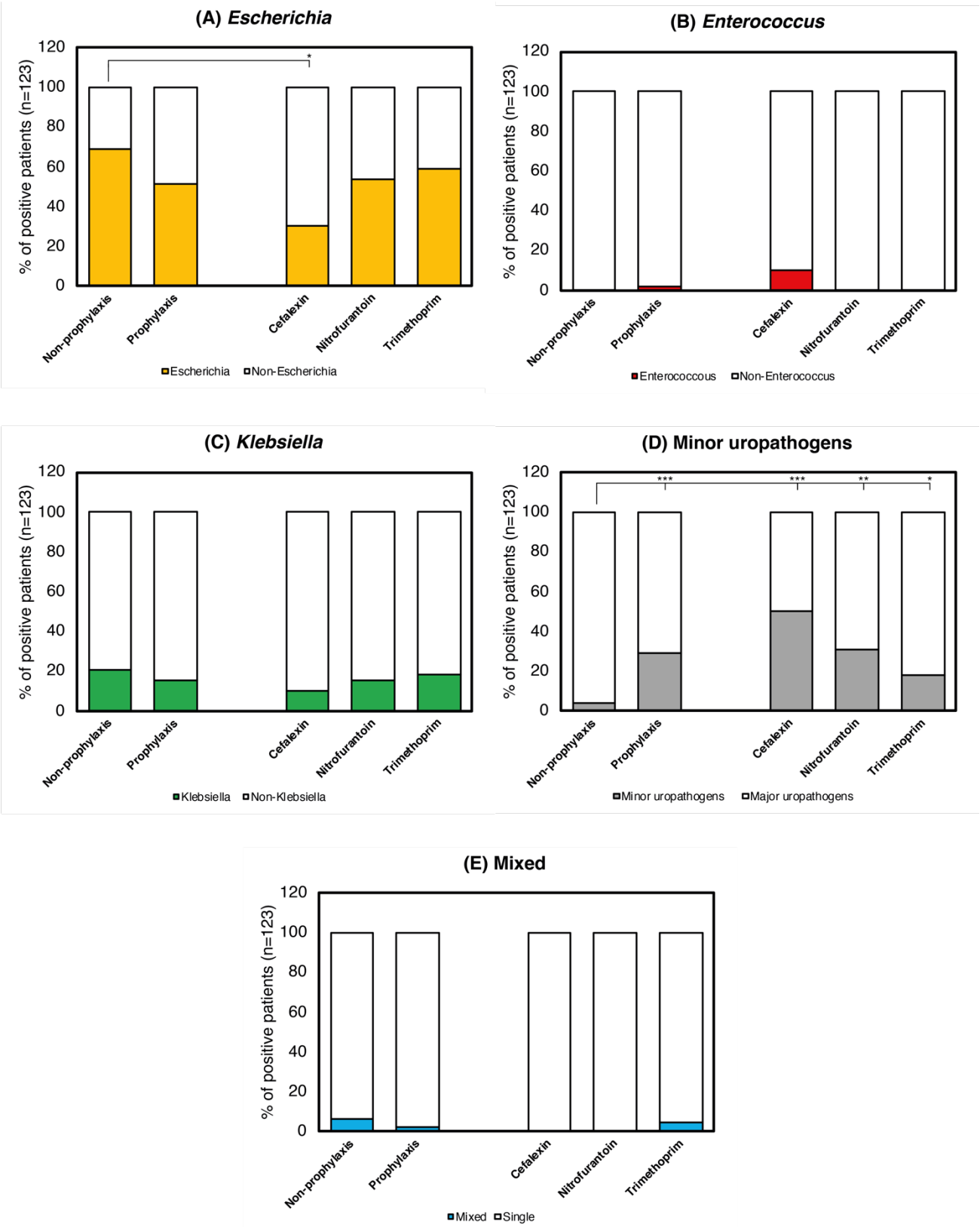



**Figure 36 Breakdown of patients with stable bacterial colonisation (positive patients) by genus (Escherichia, Enterococcus, Klebsiella, Minor uropathogens, or Mixed-species colonisation)**

## 3.3 Genomic analysis of AnTIC E. coli isolates

To further investigate the interplay between antibiotic prophylaxis and *E. coli* colonisation in AnTIC patients, a genomic approach was taken. To do this, the entire AnTIC dataset was filtered to select for patients carrying *E. coli* isolates that displayed evidence of both antibiotic adaptation (via acquisition of multidrug resistance) and persistent colonisation. *E. coli* isolates from these patients were sequenced to determine their phylogeny and genetics.

### 3.3.1 Selection methodology for whole-genome sequencing

The initial AnTIC dataset was filtered for patients that acquired multi-drug resistant (MDR) bacterial isolates during the study (Figure 37). The reasoning for this selection was that a MDR filter would increase the likelihood of discovering isolates that developed evolutionary adaptation to antibiotic therapies. MDR was defined as having resistance to at least three different antibiotic classes [324].

As mentioned earlier, the dominant bacterial isolate in each urine sample was subjected to antimicrobial susceptibility testing, which generated an AMR profile for each isolate (Figure 29). To determine whether a patient acquired MDR bacterial isolates, the AMR profile of the bacterial isolates associated with each patient was longitudinally analysed. If the earlier isolates were MDR- but later became MDR+, the patient was classified as having acquired MDR bacterial isolates (Figure 37C). This selection criterion reduced the number of patients from 404 to 50 (113 *E. coli* isolates) (Figure 37A,B).

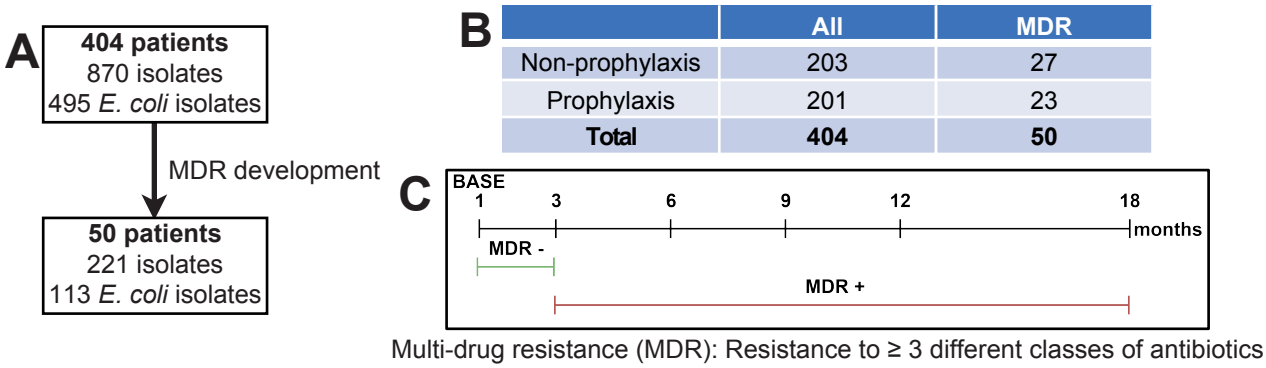



**Figure 37 MDR filter overview.** (A) The MDR filter reduced the original dataset from 397 to 50 patients. (B) The numerical breakdown of the number of patients in each study arm before (All) and after (MDR) the MDR filter was applied. (C) The MDR filter selected patients whose early bacterial isolates were MDR- but later became MDR+.

As only *E. coli* isolates were banked during the AnTIC trial, the microbiological data of the 50 patients was further reviewed to determine which patients were colonised by *E. coli* (Figure 38). *E. coli* colonisation was defined as having a minimum of two urine samples with *E. coli* isolates. This reduced the dataset to 26 patients (105 *E. coli* isolates) (Figure 38B, pink and purple shaded areas). When the prophylaxis subgroups were compared to non-prophylaxis, the incidence of patients with *E. coli* colonisation was found to be significantly lower in both the cefalexin and nitrofurantoin subgroups (1/7 patients vs 16/23 patients, Fisher's exact test: $p<0.05$) (Figure 38A) but not significantly different between trimethoprim prophylaxis and non-prophylaxis.

The microbiological data of the 26 patients was reviewed to determine which patients had persistent *E. coli* colonisation. Persistent *E. coli* colonisation was defined as having a minimum of two consecutive urine samples with *E. coli* isolates (Figure 38B, pink shaded areas). Patients who failed to meet this criterion, but had *E. coli* colonisation as defined in the previous paragraph were classified as having infrequent *E. coli* colonisation (Figure 38B, purple shaded areas). This reduced the dataset to 22 patients with persistent *E. coli* colonisation (96 *E. coli* isolates) (Figure 38B, pink shaded areas).

Analysis of the *E. coli* isolates that were banked and available yielded 45 isolates related to 16 patients with persistent *E. coli* colonisation. When the search was expanded to include isolates from patients with infrequent *E. coli* colonisation, an additional 2 isolates related to a single patient from the non-prophylaxis arm (Figure 38B) were found. Therefore, a total of 47 isolates from 17 patients were sent for whole-genome sequencing (Figure 38, green dots).

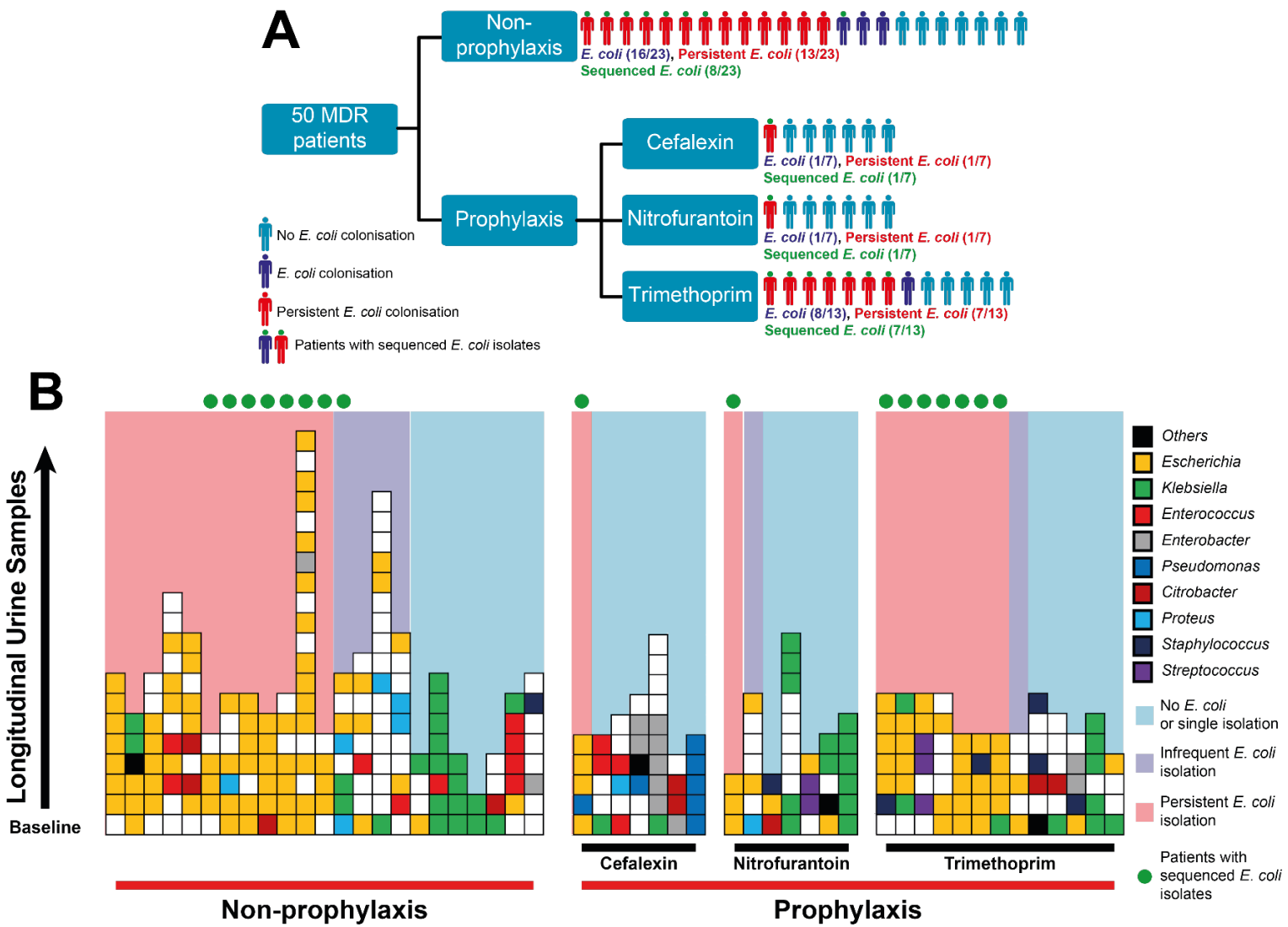



**Figure 38 MDR patients carrying MDR bacteria and selected for further study.** Figure 9A is a breakdown of the 50 patients that shows the number of patients with *E. coli* colonisation (purple shading), persistent *E. coli* colonisation (red shading), and with sequenced *E. coli* isolates (green dots). In Figure 9B, each column represents a single patient and each square within the column represents a urine sample. The urine samples for each patient have been arranged in a chronological order from bottom to top. Some patients may have a greater number of urine samples due to having a higher incidence of acute UTI episodes during the study. The background colour of each column corresponds to whether the patient had no *E. coli* colonisation (blue), infrequent *E. coli* colonisation (purple), or persistent *E. coli* colonisation (pink). Columns with a green dot represent patients with sequenced *E. coli* isolates.

### *3.3.2 Analysing the impact of antibiotic prophylaxis on E. coli colonisation using phylogenetic analysis*

#### 3.3.2.1 Bioinformatic pipelines and quality control of whole-genome sequence data

The raw whole-genome sequence data for each *E. coli* isolate consisted of "fastQ" files that contained millions of short-read sequences and their associated quality data. These short-read sequences were assembled into a draft genome. However, these short-read sequences had to be "quality-controlled" before being assembled to prevent low-quality sequences from being integrated into the draft genome, leading to misassembled or highly fractured genomes.

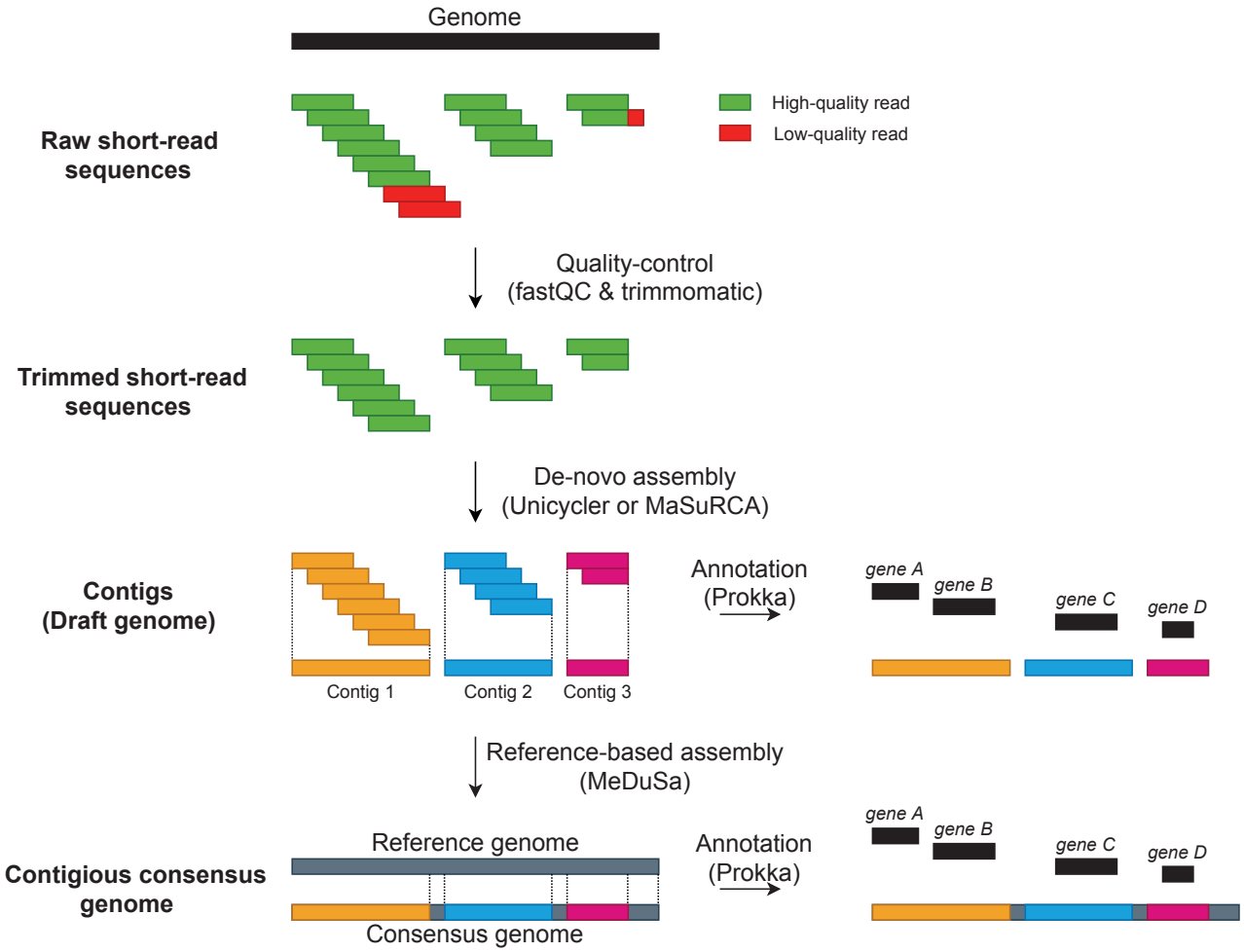



**Figure 39 Bioinformatic pipeline of WGS data.** For each isolate, their short-read sequences were quality-controlled to remove low-quality reads before being assembled into contigs using *de novo* assemblers. Contigs were either annotated using prokka or assembled into a contiguous genome using a reference-based assembler. The contiguous genome was also annotated using prokka and can be used for a wider range of genetic analyses.

Quality control was first performed using “fastQC” to visually assess the overall quality of short-read sequences for each *E. coli* isolate. These short-read sequences were then processed with “trimmomatic”, a package that utilised the quality data associated with each short-read sequence to remove low-quality nucleotide bases [325]. This allowed for some low-quality short-read sequences to be salvaged instead of being entirely discarded (Figure 39).

The quality-controlled short-read sequences were then assembled using two different *d novo* assemblers (Unicycler and MaSuRCA) to generate two assemblies for each sequenced isolate. Unicycler assemblies were used as the primary point of reference as almost all of the sequenced isolates had high-quality short reads. MaSuRCA assemblies were used as a backup option whenever the Unicycler assemblies failed to provide sufficient sequence information for genetic regions of interest. The assemblies were assessed using QUAST, which generated a variety of metrics, of which those pertaining to assembly contiguity were used to gauge assembly quality (Table 27). Contiguity is related to the number and size of contigs, with more contiguous assemblies having fewer contigs of greater lengths (high N50 values, N50: median contig length in an assembly).

**Table 27 Unicycler vs MaSuRCA assembly quality data for each of the 47 isolates.**

| | Unicycler | MaSuRCA | Unicycler | MaSuRCA | Unicycler | MaSuRCA |
|---|---|---|---|---|---|---|
| Assembly | contigs | contigs | N50 | N50 | Total length | Total length |
| 1218_12 | 245 | 210 | 61121 | 61489 | 5209144 | 5412584 |
| 1218_18 | 233 | 201 | 58394 | 58356 | 5191502 | 5363739 |
| 1253_12 | 105 | 102 | 114277 | 128150 | 4692324 | 4772328 |
| 1253_3 | 102 | 110 | 103913 | 112290 | 4635482 | 4709623 |
| 1253_9 | 112 | 148 | 107318 | 85096 | 4683750 | 4807523 |
| 1260_12 | 117 | 123 | 129363 | 114887 | 4896834 | 4978268 |
| 1260_3 | 115 | 123 | 134066 | 148024 | 5000530 | 5085630 |
| 13_3 | 107 | 117 | 99005 | 93271 | 5130249 | 5224060 |
| 13_6 | 95 | 76 | 132093 | 131384 | 4732225 | 4824179 |
| 13_9 | 146 | 160 | 84628 | 66979 | 5151353 | 5202317 |
| 1646_3 | 77 | 84 | 187683 | 156139 | 4739317 | 4820498 |
| 1646_6 | 72 | 59 | 190833 | 193412 | 4727245 | 4749856 |
| 1646_BASE | 75 | 74 | 171655 | 166314 | 4743349 | 4809912 |
| 2015_12 | 210 | 274 | 43413 | 37389 | 4748258 | 4892361 |
| 2015_BASE | 127 | 137 | 101068 | 98428 | 4822968 | 4935362 |
| 2417_6 | 97 | 83 | 192582 | 163169 | 4962680 | 4993866 |
| 2417_BASE | 101 | 97 | 140034 | 133951 | 5007241 | 5050697 |
| 2433_12 | 93 | 63 | 174099 | 208776 | 5091156 | 5359078 |
| 2433_18 | 256 | 303 | 39793 | 38495 | 5013875 | 5202353 |
| 2433_3 | 61 | 48 | 197307 | 198935 | 5048549 | 5064835 |
| 2433_6 | 107 | 137 | 83010 | 108600 | 4799926 | 4960859 |
| 2433_9 | 159 | 125 | 102667 | 134955 | 5170860 | 5278429 |
| 2433_BASE | 49 | 51 | 281497 | 183238 | 4880268 | 4926864 |
| 2434_12 | 130 | 96 | 97023 | 131368 | 4924428 | 4993530 |
| 2434_18 | 131 | 133 | 106192 | 113996 | 4663824 | 4781946 |
| 2439_12 | 171 | 111 | 123748 | 133888 | 5316812 | 5445207 |
| 2439_18 | 133 | 127 | 168767 | 135868 | 5219362 | 5333784 |
| 2439_6 | 280 | 333 | 37413 | 31319 | 5206731 | 5282038 |
| 2439_BASE | 160 | 134 | 83431 | 105108 | 4933828 | 4921002 |
| 2470_3 | 103 | 62 | 341939 | 271215 | 5253728 | 5312117 |
| 2470_6 | 122 | 103 | 122529 | 118803 | 5222721 | 5283137 |
| 2470_BASE | 192 | 228 | 53483 | 45838 | 5068442 | 5215014 |
| 2877_3 | 84 | 59 | 205475 | 213213 | 4827388 | 5036019 |
| 2877_9 | 168 | 135 | 87615 | 112778 | 4901473 | 4974303 |
| 2889_6 | 104 | 119 | 142096 | 162170 | 5012799 | 5068729 |
| 2889_9 | 89 | 139 | 143894 | 110750 | 5016070 | 5076328 |
| 426_18 | 79 | 118 | 133026 | 130314 | 5088873 | 5157360 |
| 426_3 | 126 | 139 | 94431 | 76113 | 5030222 | 5103858 |
| 426_9 | 68 | 90 | 163843 | 167418 | 5117291 | 5147824 |
| 426_BASE | 89 | 81 | 140499 | 158609 | 5082552 | 5234328 |
| 429_6 | 187 | 130 | 124375 | 147304 | 5292413 | 5337297 |
| 429_9 | 215 | 208 | 85828 | 69873 | 5167529 | 5249597 |
| 432_3 | 111 | 87 | 97348 | 162363 | 4976605 | 5105282 |
| 432_6 | 1229 | 369 | 14879 | 58043 | 5909364 | 5935977 |
| 432_BASE | 148 | 122 | 115142 | 114242 | 5003266 | 5075290 |
| 824_18 | 141 | 110 | 136669 | 125970 | 4849145 | 4908614 |
| 824_9 | 213 | 205 | 62777 | 64169 | 5121384 | 5253244 |
| **Average** | **156** | **135** | **123665** | **123031** | **5006071** | **5099086** |

Contigs: Number of contigs ≥ 500bp. N50: Median contig length. Total length: Combined length of all the contigs in the assembly. High quality assemblies have the following characteristics: low number of contigs and high N50 value.

The output from both *de novo* assemblers was a draft genome that consisted of around 50-150 contigs of various lengths. These contigs represented the entire genome of an *E. coli* isolate but in a discontiguous form. The main limitation of using short-read sequences for *de novo* assembly was the inability to assemble genetic regions that are rich in repeat sequences [326]. These sequences often cause premature termination of contig assembly, preventing the contigs from being joined up to each other to form a contiguous genome [326].

The draft genomes generated by Unicycler were annotated using “Prokka” (Section 1.8.3.1) with the default e-value threshold of $10^{-6}$ being used for BLAST+ annotation [307]. Some draft genomes were further processed using a reference-based assembler called MeDuSa before annotation (Figure 39). MeDuSA aligns the contigs to multiple reference genomes to generate a single contiguous consensus genome [327]. Using a reference genome allows for the resolution of repeat regions that interfered with *de novo* assembly. Reference genomes were chosen from available *E. coli* genomes with a high degree of phylogenetic relatedness to the *E. coli* isolates; this was done for each phylogenetic group. Phylogenetic relatedness was determined using a 7-gene Multi Locus Sequence Typing scheme together with reference *E. coli* genomes [317].

#### 3.3.2.2 Phylogenetic analysis using Multi Locus Sequence Typing (MLST) and core-genome MLST (cgMLST)

Initial phylogenetic analysis of the 47 *E. coli* assemblies was performed based on the Achtman MLST scheme, which utilised 7 highly-conserved housekeeping genes as phylogenetic markers [328]. The seven genes in this scheme were extracted from each annotated assembly and then concatenated to generate a MLST dataset. Using the MLST dataset, sequence types (ST) were assigned to each isolate using the online sequence query tool on PubMLST [329]. The assignment of sequence types was used to distinguish between different *E. coli* strains, thus providing more detail of the nature of *E. coli* colonisation in each patient.

These genomic data were used to query the impact of antibiotic prophylaxis on *E. coli* colonisation in each of the 17 AnTIC patients. *E. coli* colonisation in each patient was classified into two categories: unstable and stable. If a patient had different *E. coli* strains, i.e. *E. coli* isolates with different ST numbers throughout the study period, that patient was classified as having unstable *E. coli* colonisation (Figure 40). Examples of such patients include patients 13, 426, 2433 who were part of the non-prophylaxis group and were each characterised by up to 5 different *E. coli* strains throughout the study period (Figure 41).

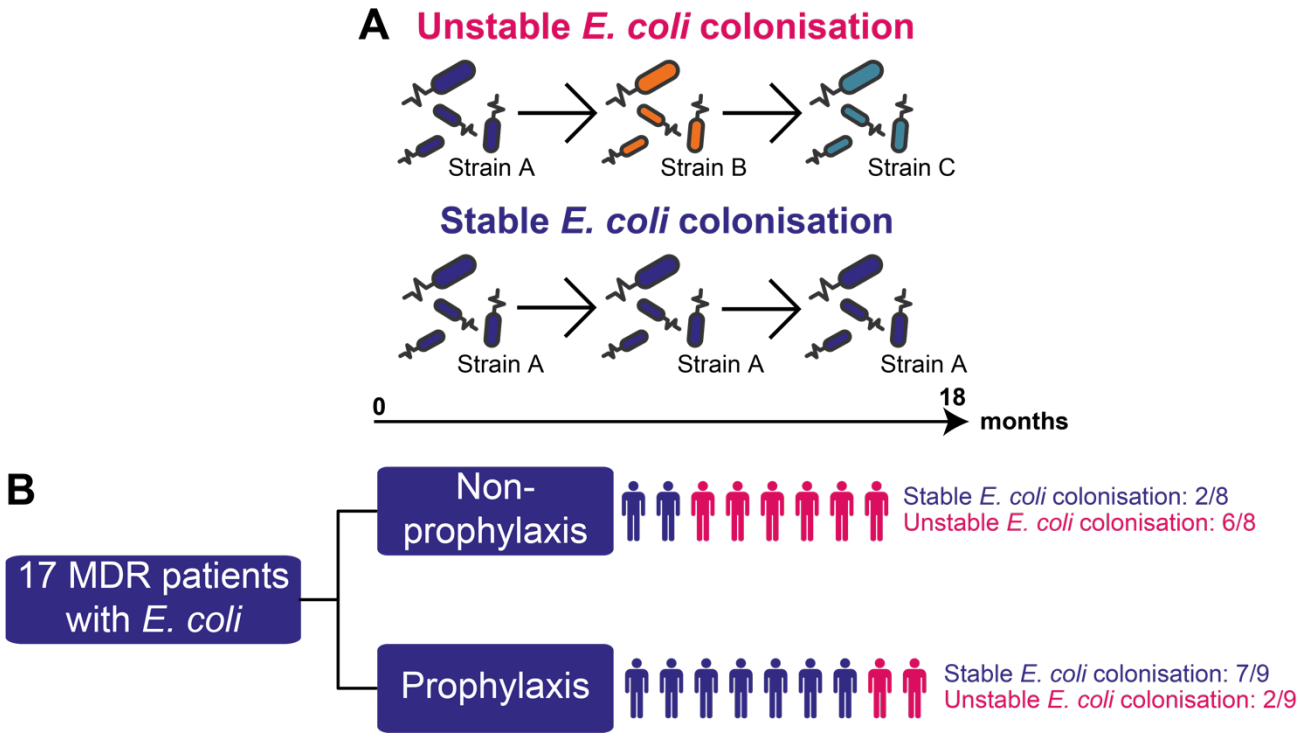



**Figure 40 Breakdown of stable vs unstable *E. coli* colonisation in the 17 patients with sequenced *E. coli* isolates.** (A) Unstable E. coli colonisation was characterised by changes in the colonising E. coli strain during the study period whereas stable E. coli colonisation was characterised by no changes in the colonising E. coli strain during the study period (B)The number of patients with stable *E. coli* colonisation was significantly higher in the prophylaxis arm compared to the non-prophylaxis arm.

Patients who carried the same *E. coli* strain i.e. *E. coli* isolates with the same ST number throughout the study were classified as having stable *E. coli* colonisation (Figure 40). Examples of patients populating this group and all characterised by their prophylaxis antibiotic treatment included patients 429, 2317, and 2470 (Figure 41). As shown by the data and despite being prescribed different prophylactic antibiotic treatments (Green line: Cefalexin, Green line: Nitrofurantoin, Orange line: Trimethoprim, Figure 41), each carried the same strain through their time on the trial. Interestingly, the colonising strains differed between the individual patients. There were exceptions, for example, patients 426 and 1646 populating the non-prophylactic arm were stably colonised with ST131, and patients 2434 and 2439 who were given prophylactic antibiotics (trimethoprim) were characterised by unstable *E. coli* colonisation (Figure 41).

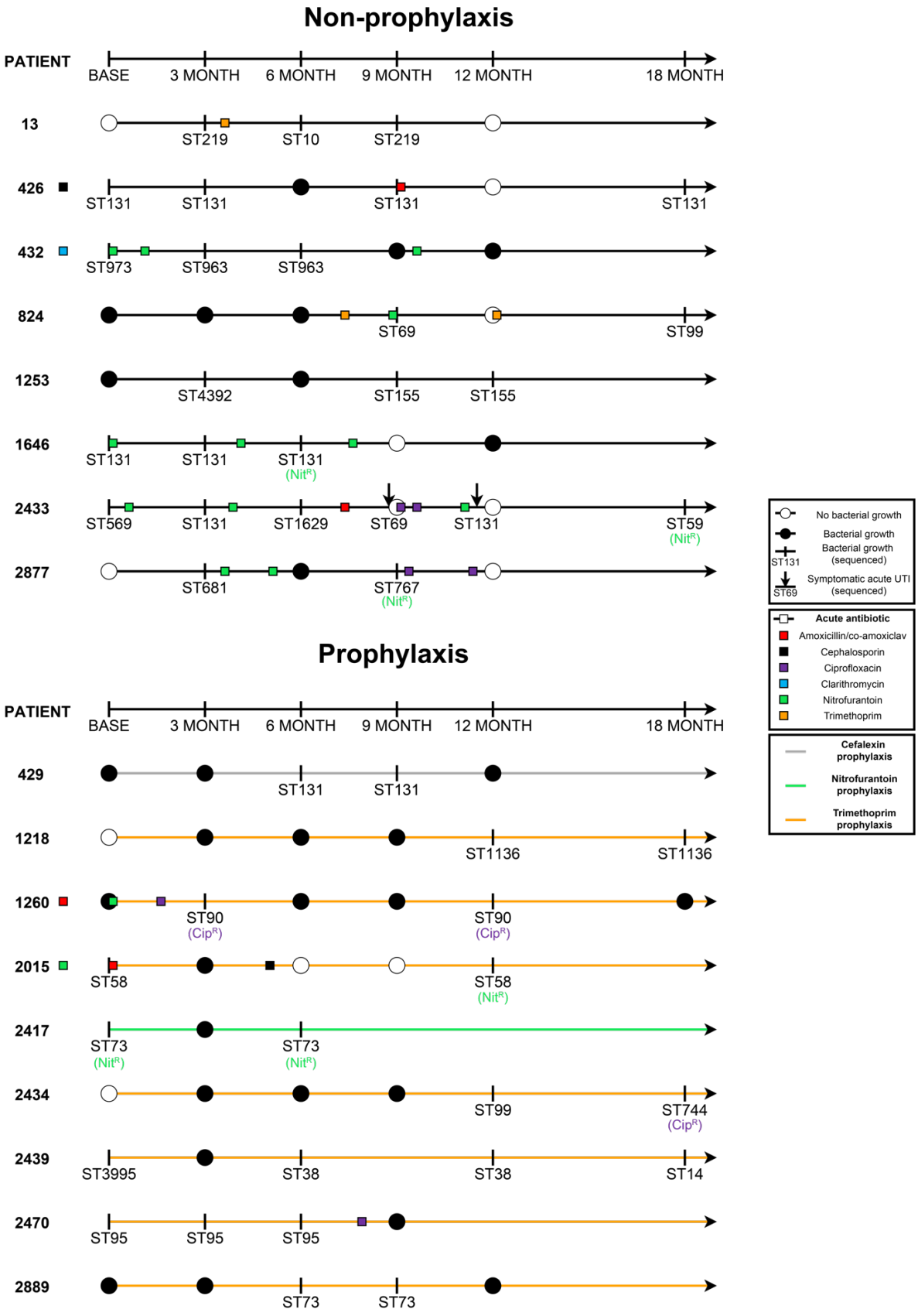



**Figure 41 Colonisation diagram of the 17 patients whose isolates were sequenced.** Each patient is represented by an arrow, refer to the figure legends for more information regarding the symbols. For patients in the prophylaxis arm, the colour of the arrow shaft denotes his/her assigned prophylaxis subgroup, refer to figure legends for more information. A coloured square box next to the patient's ID meant that the patient was prescribed a specific antibiotic for acute UTI management without any accompanying chronological data

Using this classification system, 2 out of 8 patients in the non-prophylaxis arm were found to have stable *E. coli* colonisation versus 7 out of 9 patients in the prophylaxis arm (Figure 40B). Although the total number of patients was small, these data suggested an association between antibiotic prophylaxis and stable *E. coli* colonisation with this association approaching statistical significance (Fisher's exact test p=0.057).

To visualise and further define the differences in *E. coli* colonisation between these two patient groups, a maximum-likelihood phylogenetic tree was constructed from this MLST dataset with a bootstrap value of 250 (Figure 42A). *E. coli* genomes from McNally et al (2013) were included as reference points [317]. As shown in Figure 42A, patients with stable *E. coli* colonisation had isolates that were more closely grouped together compared to patients with unstable *E. coli* colonisation. This suggested that the isolates from patients with stable *E. coli* colonisation were more closely related to each other, and thus more likely to be the same strain.

To further confirm the phylogeny of the *E. coli* isolates, a cgMLST dataset was generated from the *E. coli* assemblies using chewBBACCA. cgMLST analysis was performed using a novel 404-loci schema generated by chewBBACA with the 47 *E. coli* assemblies as reference genomes. The schema was generated by first identifying complete coding sequences (CDS) in each of the 47 *E. coli* genomes using Prodigal. The CDSs were then grouped into loci based on their blastp similarity score. This generated a list of candidate loci and only loci that were present in 100% of the *47 E. coli* genomes were chosen for the final schema. The schema was then used to perform variant calling, a process that translated the nucleotide sequences of the 404 loci within each of the 47 *E. coli* assemblies into a series of allele numbers. This is similar to the approach used to generate the allele numbers in MLST genotyping except that a greater number of loci is involved (404 vs 7) (Section 1.8.3.3 & Figure 22).

Every variant of a gene was assigned an arbitrary and unique allele number. The combination of these allele numbers for all 404 genes in a given isolate makes up the allelic profile of an isolate. The final output of variant calling is a dataset that contains the allelic profiles of the 47 *E. coli* isolates. This dataset was translated into a distance-matrix table using phyloviz. In this distance-matrix table, *E. coli* isolates that shared a greater number of genes with the same allele number were clustered together. This clustering was converted into a distance-based dendrogram in R studio using the "ape" package (Figure 42B).

Using a greater number of genes (404 vs 7) as phylogenetic markers increased phylogenetic resolution, allowing for the differentiation of more closely related isolates. For example, while a seven-gene MLST might fail to differentiate between isolates belonging to a single clonal group due to its limited phylogenetic resolution, a 404-gene cgMLST can differentiate between these isolates. In the case of ST131 isolates, the most abundant clonal group in this dataset, patients with stable *E. coli* colonisation should carry the same ST131 isolate for an extended period.

The cgMLST findings confirmed, for example, that ST131 isolates from the same patient were grouped more closely together compared to ST131 isolates from different patients (Figure 42B). These findings further strengthened the association between antibiotic prophylaxis and stable *E. coli* colonisation.  Therefore, the reduced incidence of UTI associated with antibiotic prophylaxis could be attributed to its ability to stabilise *E. coli* colonisation among a subset of patients.

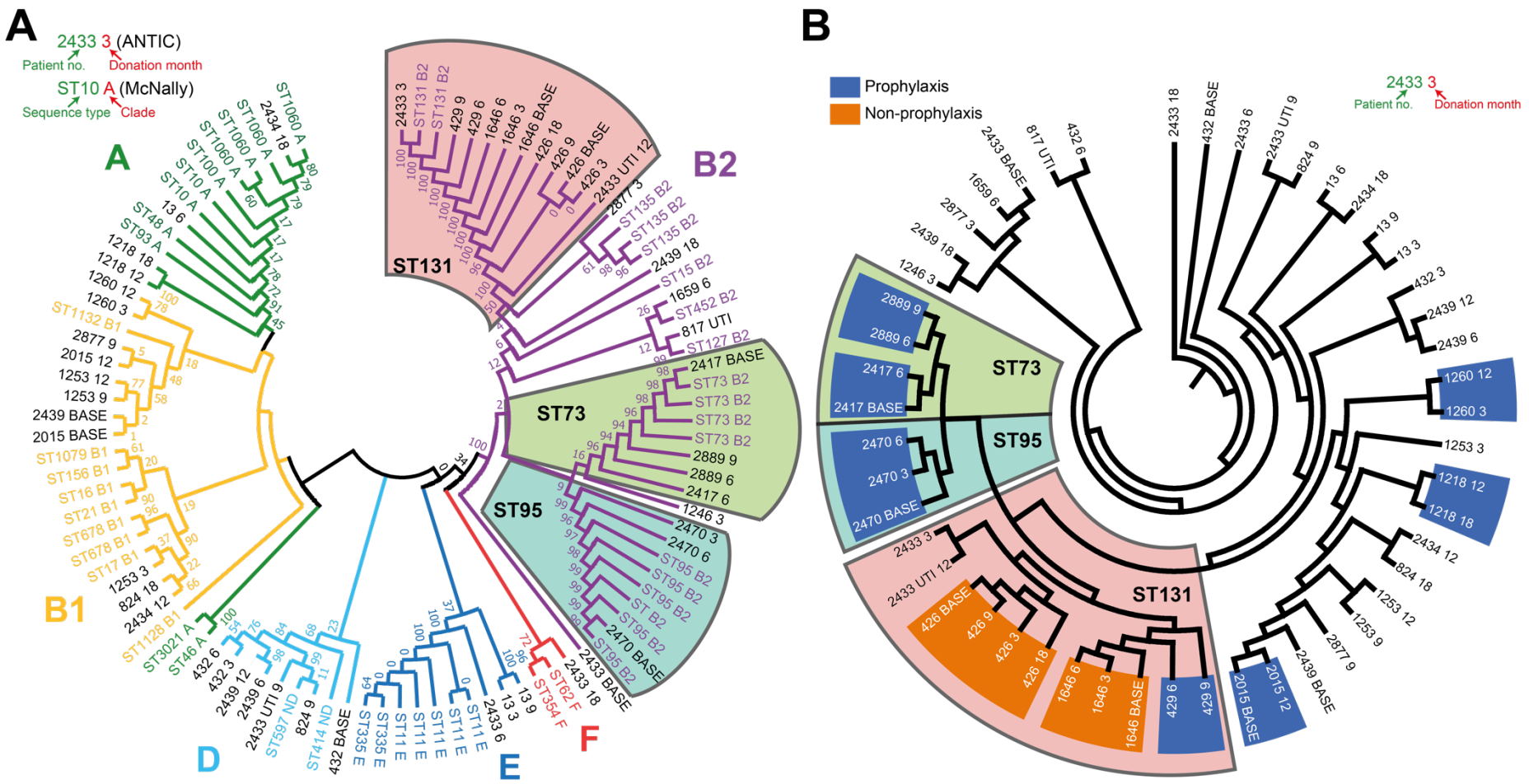



**Figure 42 MLST (A) and cgMLST (B) phylogenetic trees of the 47 sequenced *E. coli* isolates.** Figure A is a maximum-likelihood phylogenetic tree generated from the 7-gene MLST dataset of the 47 sequenced AnTIC *E. coli* isolates and 49 refence *E. coli* genomes from McNally et al. 2013. Clinically relevant ST groups (ST131, ST73, ST95) are demarcated. Figure B is a distance-based phylogenetic tree generated from the 404-gene cgMLST dataset of the 47 sequenced AnTIC *E. coli* isolates. Isolates with a blue (prophylaxis) or orange (non-prophylaxis) background were sourced from patients with stable *E. coli* colonisation. Tree A was generated using the Molecular Evolutionary Genetics Analysis (MEGA) software version 10 from a dataset of aligned nucleotides while tree B was generated using the ape package in R Studio based on a distance-matrix table generated by chewBBACA and phyloviz [330].

## *3.4 Investigating the development of antibiotic resistance amongst the sequenced AnTIC E. coli isolates*

The 47 sequenced *E. coli* isolates were chosen because they displayed evidence of antibiotic adaptation by acquiring multi-drug resistance (MDR). For patients with unstable *E. coli* colonisation, MDR was acquired by strain/species replacement, in which the initial strain was likely replaced by a different strain with a broader repertoire of antibiotic resistance. For patients with stable *E. coli* colonisation, MDR was presumed to be acquired by strain adaptation, in which the initial strain broadened its repertoire of antibiotic resistance either via vertical or horizontal evolution.

The acquisition of antibiotic resistance in *E. coli* is commonly mediated via horizontal evolution. Among the antibiotics that are commonly used in the management of UTIs, resistance against nitrofurantoin and ciprofloxacin is primarily acquired via vertical evolution (chromosomal mutations) whereas resistance against trimethoprim and β-lactam (cefalexin, amoxicillin) is primarily acquired via horizontal evolution (acquisition of exogenous genes via mobile genetic elements) [331,332].

The longitudinal nature of the AnTIC study in which each patient donated multiple urine samples over a 12-18 month period has facilitated the investigation of both horizontal and vertical evolution. For example, patients with stable *E. coli* colonisation will have multiple chronological “snapshots” of the same isolate. By comparing the AMR profiles and genomic data of these isolates, their evolutionary course can be charted in the context of antibiotic adaptation.

Unlike the gut environment, the nutrient-poor nature of the urinary environment severely limits its ability to support a large and diverse microflora population, limiting, but not excluding the opportunities for horizontal evolution. Because of this, it would be expected that vertical evolution would dominate in driving antibiotic adaptation among urinary isolates. For this reason, the focus of the following investigations were on antibiotic resistance acquired via vertical evolution.

### *3.4.1 Nitrofurantoin resistance in AnTIC E. coli isolates*

As shown in Figure 41 and Table 28 (fourth column), nitrofurantoin was used in the prophylaxis arm of the study and prescribed to several patients for management of acute

UTI in both the non-prophylaxis and prophylaxis arms. Five isolates across the MDR dataset were originally specified as being nitrofurantoin resistant ($Nit^R$) (Table 28).

Nitrofurantoin resistance in *E. coli* is associated with an inactivation of nitroreductase activity. Nitroreductases are responsible for converting nitrofurantoin into its biologically active form that exerts its bactericidal activity by damaging DNA, RNA and proteins [259]. There are two main families of nitroreductases in *E. coli*: oxygen-sensitive and oxygen-insensitive [272]. Only the inactivation of oxygen-insensitive nitroreductases is necessary for nitrofurantoin resistance in the urinary tract [273].

This family is comprised of two proteins: NfsA and NfsB; encoded by the *nfsA* and *nfsB* genes respectively [272,273]. The NfsA protein is responsible for 70% of oxygen-insensitive nitroreductase activity in *E. coli*, with NfsB being responsible for the other 30% [272]. Both *nfsA* and *nfsB* need to be inactivated to acquire clinical resistance to nitrofurantoin (MIC ≥ 64 μg/ml) [271,272]. Nitrofurantoin resistance almost always develops in a two-step manner, beginning with the inactivation of *nfsA* followed by the inactivation of *nfsB* [271,273]. Inactivation of *nfsA* alone will yield low-intermediate levels of resistance due to the residual NfsB nitroreductase activity [272,273]. However, inactivation of *nfsB* alone will not increase resistance as NfsA accounts for the majority (70%) of oxygen-insensitive nitroreductase activity in *E. coli* [272].

In this dataset of 17 patients carrying MDR isolates, the emergence of nitrofurantoin resistant isolates corresponded to nitrofurantoin exposure. Figure 41 clearly demonstrates this pattern in patients 1646, 2015, 2417, and 2433. Using patient 1646 as an example, the $Nit^R$ 1646 6 isolate emerged after two separate occasions of nitrofurantoin usage. The first of which occurred shortly after the AnTIC study started (represented by a green box next to BASE, Figure 41) and the second occurred sometime between the 3rd and 6th month of the study (represented by a green box between the markers corresponding to Month 3 and 6, Figure 41).

Nitrofurantoin accounted for 18/44 (40.9%) of all acute antibiotic prescriptions within this dataset (Figure 41), compared to the 273/914 (29.9%) across the entire AnTIC dataset. For comparison, nitrofurantoin accounted for 42.4% of all UTI-related antibiotic prescriptions within the UK primary care setting from 2013-2015 [333]. Amongst the 8 patients who were

prescribed nitrofurantoin during the study (non-prophylaxis and prophylaxis), $Nit^R$ isolates emerged amongst three of them (Patients 1646, 2015, 2417).

The frequency at which $Nit^R$ isolates emerged in response to nitrofurantoin usage suggests that the mutational frequency conferring nitrofurantoin resistance is fairly high among *E. coli* isolates [271]. Interestingly, despite the lower proportional usage of nitrofurantoin in this study, the incidence of nitrofurantoin resistance among *E. coli* isolates across the entire AnTIC dataset is substantially higher compared to the average national incidence rate of NitR urinary *E. coli* isolates within the UK (9.6% vs 2.1%) [334].

#### 3.4.1.1 *In silico* analysis of nitrofurantoin resistance in AnTIC isolates

To further explore nitrofurantoin resistance, the *nfsA* and *nfsB* genes of the 47 *E. coli* isolates were analysed to determine whether the $Nit^R$ phenotype was associated with the presence of mutations within *nfsA* and *nfsB*.

1. Four of the five $Nit^R$ isolates (2433 UTI 18, 2015 12, 2417 BASE, and 2417 9) had mutations in both *nfsA* and *nfsB* (Table 28, highlighted in red).
2. A single $Nit^R$ isolate (2877 9) had a point mutation in *nfsA* but an intact, wild-type *nfsB* gene (Table 28, highlighted in orange).
3. A single $Nit^S$ isolate (1646 6) had mutations in both *nfsA* and *nfsB* but was classified as being nitrofurantoin sensitive ($Nit^S$) (Table 28, highlighted in yellow).

To address the discrepancies between the *in silico* analysis and the AMR profile dataset, nitrofurantoin MIC assays were performed to quantitatively determine the level of nitrofurantoin resistance in isolates that were either classified genetically as being $Nit^R$ or had mutations in both *nfsA* and *nfsB*. The nitrofurantoin MIC assay data validated the $Nit^R$ phenotype of all 4 $Nit^R$ isolates that had mutations in both *nfsA* and *nfsB* (Figure 43). Isolate 2877 9 was found to be $Nit^S$ while 1646 6 was found to be $Nit^R$ (Figure 43). The updated nitrofurantoin resistant phenotypes for both isolates have been included within brackets in Table 28. The emergence of $Nit^R$ isolates always corresponded with nitrofurantoin usage in both acute and prophylactic settings (Table 28).

**Table 28 Nitrofurantoin resistance phenotype and NfsA/B genotype of sequenced AnTIC isolates.**

| Patient | Donation month | Study Arm | Acute Nit. | NitR | NfsA | NfsB | MIC (µg/ml) |
|---|---|---|---|---|---|---|---|
| 13 | 3 | No prop. | No | S | | | |
| | 6 | | | S | | A93V | |
| | 9 | | | S | | | |
| 426 | BASE | No prop. | No | S | | | |
| | 3 | | | S | | | |
| | 9 | | | S | | | |
| | 18 | | | S | | | |
| 432 | BASE | No prop. | Yes | S | | E137Q | |
| | 3 | | | S | | E137Q | |
| | 6 | | | S | | E137Q | |
| 824 | 9 | No prop. | Yes | S | | | |
| | 18 | | | S | L157F | | |
| 1253 | 3 | No prop. | No | S | E28G | | |
| | 9 | | | S | E28G | | |
| | 12 | | | S | E28G | | |
| 1646 | BASE | No prop. | Yes | S | | | 8 |
| | 3 | | | S | | | 8 |
| | 6 | | | S (R) | T37M | complete del. | 128 |
| 2433 | BASE | No prop. | Yes | S | G66R | | |
| | 3 | | | S | | | |
| | 6 | | | S | | | |
| | 9 | | | S | | | |
| | UTI 12 | | | S | A34V | | |
| | UTI 18 | | | R | complete del. | partial del. | 128 |
| 2877 | 3 | No prop. | Yes | S | | H47Y | |
| | 9 | | | R (S) | Q88STOP | | 16 |
| 429 | 6 | Cefalexin prop. | No | S | | | |
| | 9 | | | S | | | |
| 1218 | 12 | Trimethoprim prop. | No | S | | | |
| | 18 | | | S | | | |
| 1260 | 3 | Trimethoprim prop. | Yes | S | | | |
| | 12 | | | S | | | |
| 2015 | BASE | Trimethoprim prop. | Yes | S | partial del. | | 32 |
| | 12 | | | R | partial del. | partial del. | 128 |
| 2417 | BASE | Nitro prop. | No | R | partial del. | partial del. | >128 |
| | 6 | | | R | partial del. | partial del. | >128 |
| 2434 | 12 | Trimethoprim prop. | No | S | | | |
| | 18 | | | S | | A93V | 16 |
| 2439 | BASE | Trimethoprim prop. | No | S | partial del. | | |
| | 6 | | | S | | K122R | |
| | 12 | | | S | | K122R | |
| | 18 | | | S | | | |
| 2470 | BASE | Trimethoprim prop. | No | S | | | |
| | 3 | | | S | | | |
| | 6 | | | S | | | |
| 2889 | 6 | Trimethoprim prop. | No | S | A188V | | |
| | 9 | | | S | A188V | | |

An isolate's name is the combination of the patient's ID (first column) and donation month (second column) separated by a space. For example, isolate 13 3 is Patient 13's *E. coli* isolate that was collected during month 3 of the study. The third column (study arm) shows which cohort the patient was assigned to during the AnTIC trial, no prop. (non-prophylaxis), cefalexin prop. (cefalexin prophylaxis), trimethoprim prop. (trimethoprim prophylaxis), Nitro prop. (nitrofurantoin prophylaxis). The fourth column (NitR) shows whether an isolate was nitrofurantoin sensitive (S) or resistant (R). The fifth (NfsA) and sixth (NfsB) columns shows the mutations that were present in the NfsA or NfsB protein respectively, amino acid changes are shown for point mutations, deletions are either partial deletions (partial del.) or complete deletions (complete del.). The last column (MIC) shows the minimum inhibitory concentration of isolates that underwent MIC assays. Isolates highlighted in green denote nitrofurantoin sensitive ($Nit^S$) isolates. Isolates highlighted in pink denote nitrofurantoin resistant ($Nit^R$) isolates. Isolate highlighted in yellow denote $Nit^R$ isolates that were originally misclassified as being $Nit^S$.

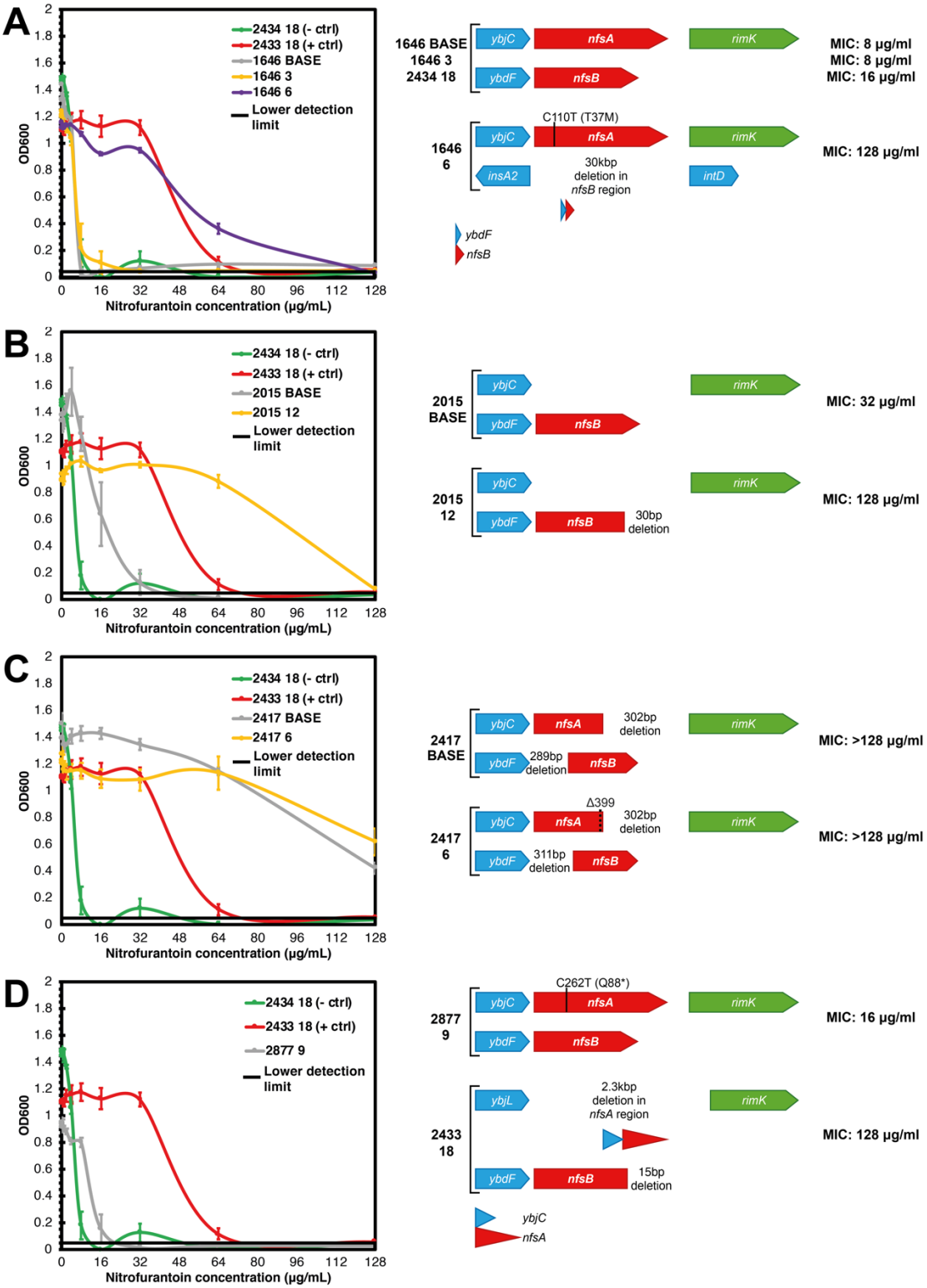



**Figure 43 Nitrofurantoin MIC findings and *nfsA/B* genotypes of selected AnTIC isolates.** 2434 18 was used as a Nit$^S$ control while 2433 18 was used as a Nit$^R$ control, these two isolates are included in A-D as Nit$^S$ and Nit$^R$ references. The MIC data of the tested isolates is visualised using line graphs: (A) Isolates from patient 1646. (B) Isolates from patient 2015. (C) Isolates from patient 2417. (D) Isolates from patient 2877. Gene maps of the *nfsA* and *nfsB* regions of the tested isolates are included to demonstrate the association between nitrofurantoin resistance (MIC ≥ 64 µg/ml) and the presence of inactivating mutations (frequently deletions) in both *nfsA* and *nfsB*. The MIC for each of the tested isolate is shown to the right of the gene maps, the MIC threshold was defined using an $OD_{600}$= 0.05. The antibiotic concentration that resulted in the failure of an isolate to grow beyond this threshold is the MIC of that isolate.

### 3.4.1.2 Development of nitrofurantoin resistance within the same *E. coli* strain

The *E. coli* isolates of Patients 1646 and 2015 respectively developed nitrofurantoin resistance during the AnTIC trial. The baseline isolate of Patient 1646 (1646 BASE) had a MIC of 8 µg/ml, this increased to 128 µg/ml for the 6-month isolate (1646 6). For Patient 2015, the baseline isolate (2015 BASE) had a MIC of 32 µg/ml, which increased to 128 µg/ml for the 12-month isolate (2015 12). Isolate 1646 BASE had wild-type *nfsA* and *nfsB* while 1646 6 had a point mutation in *nfsA* and the complete deletion of *nfsB* (Figure 43A). The deletion of *nfsB* in 1646 6 was part of a larger 30kbp deletion that was confirmed by PCR (Figure 44) and hypothesised to have resulted from the transposition of an insertion element into/out of the *nfsB* region (Figure 45). However, without long-read sequence data, it was impossible to determine the exact transposition events that led to this deletion.

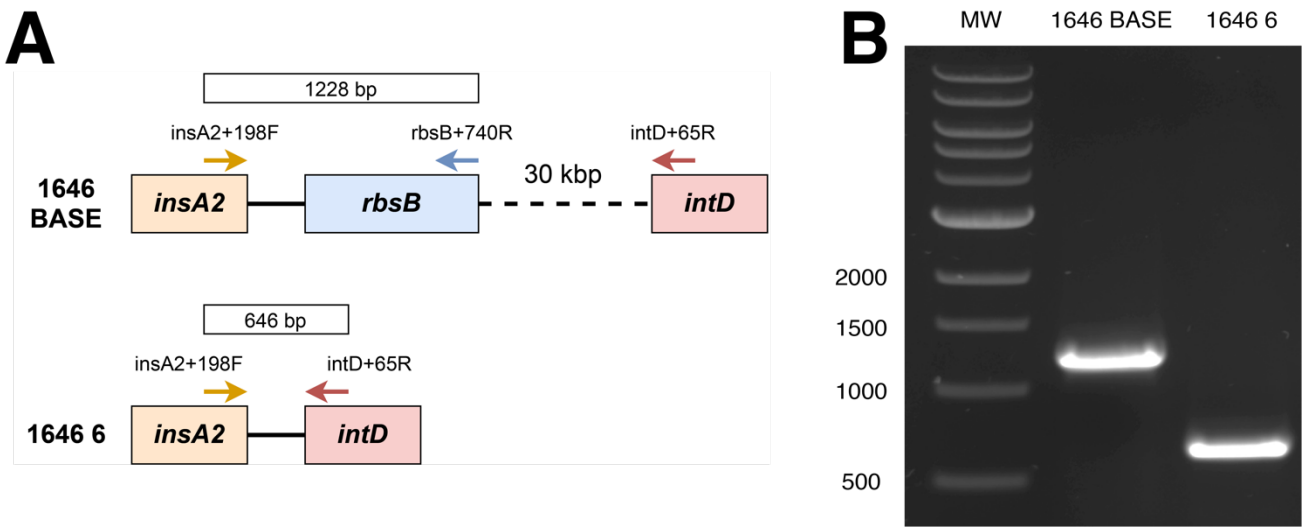



**Figure 44 PCR validation of 30kbp deletion in the *nfsB* region of Isolate 1646 6. (A)** Primers were designed to target the *insA2*, *rbsB,* and *intD* genes, the *nfsB* gene is located 8132bp downstream of the *rbsB* stop codon (not shown). All three genes were present in isolate 1646 BASE but PCR amplification will only yield a 1228 bp PCR product from the primer pair insA2+198F & rbsB+740R targeting *insA2* and *rbsB* respectively. In 1646 BASE, *intD* was located too far downstream for a PCR product to be generated from insA2+198F and intD+65R. However, isolate 1646 6 has a 30 kbp deletion which brings *intD* closer to *insA2,* allowing for a 646 bp PCR product to be generated from the primer pair insA2+198F and intD+65R. **(B)** Agarose gel showing the PCR products generated using the aforementioned primers with 1646 BASE (Column 2) and 1646 6 (Column 3) as the DNA template. Column 1 was loaded with 1 kb DNA ladder manufactured by New England Biolabs.

Isolate 2015 BASE had a complete deletion of *nfsA,* but an intact, wild-type *nfsB* (Figure 43B). The deletion of *nfsA* in the 2015 BASE could explain its higher MIC compared to 1646 BASE. Deletion of *nfsA* alone is known to bestow intermediate levels of nitrofurantoin resistance in *E. coli*. Compared to 2015 BASE, 2015 12 had an additional partial deletion in *nfsB*, which further increased its MIC (Figure 43B). According to cgMLST-based phylogenetic data, 1646 BASE, 1646 3, and 1646 6 were the same strain (ST131); this was

also true for 2015 BASE and 2015 12 (ST58) (Figure 42). This meant that the development of nitrofurantoin resistance within the *E. coli* isolates of patients 1646 and 2015 were probably due to the persisting bacteria acquiring resistance to nitrofurantoin (strain evolution) rather than being replaced by another $Nit^R$ strain (strain invasion). Patient 1646 was prescribed short-course nitrofurantoin therapy during the first and fourth month of AnTIC (Figure 41). Patient 2015 was also prescribed short-course nitrofurantoin, but there was no chronological record associated with this prescription (Figure 41).

Previous studies on nitrofurantoin resistance in *E. coli* have been limited to using $Nit^S$ laboratory [273] or clinical isolates [271] that were induced to become $Nit^R$ within laboratory settings. Having access to the evolutionary "snapshots" of $Nit^S$ (1646 BASE, 2015 BASE) and $Nit^R$ (1646 6, 2015 12) strains that naturally evolved nitrofurantoin resistance in a clinical setting has potential to aid to understanding the mechanism(s) underpinning the low incidence of nitrofurantoin resistance within *E. coli* isolates, a topic be explored in the next chapter.

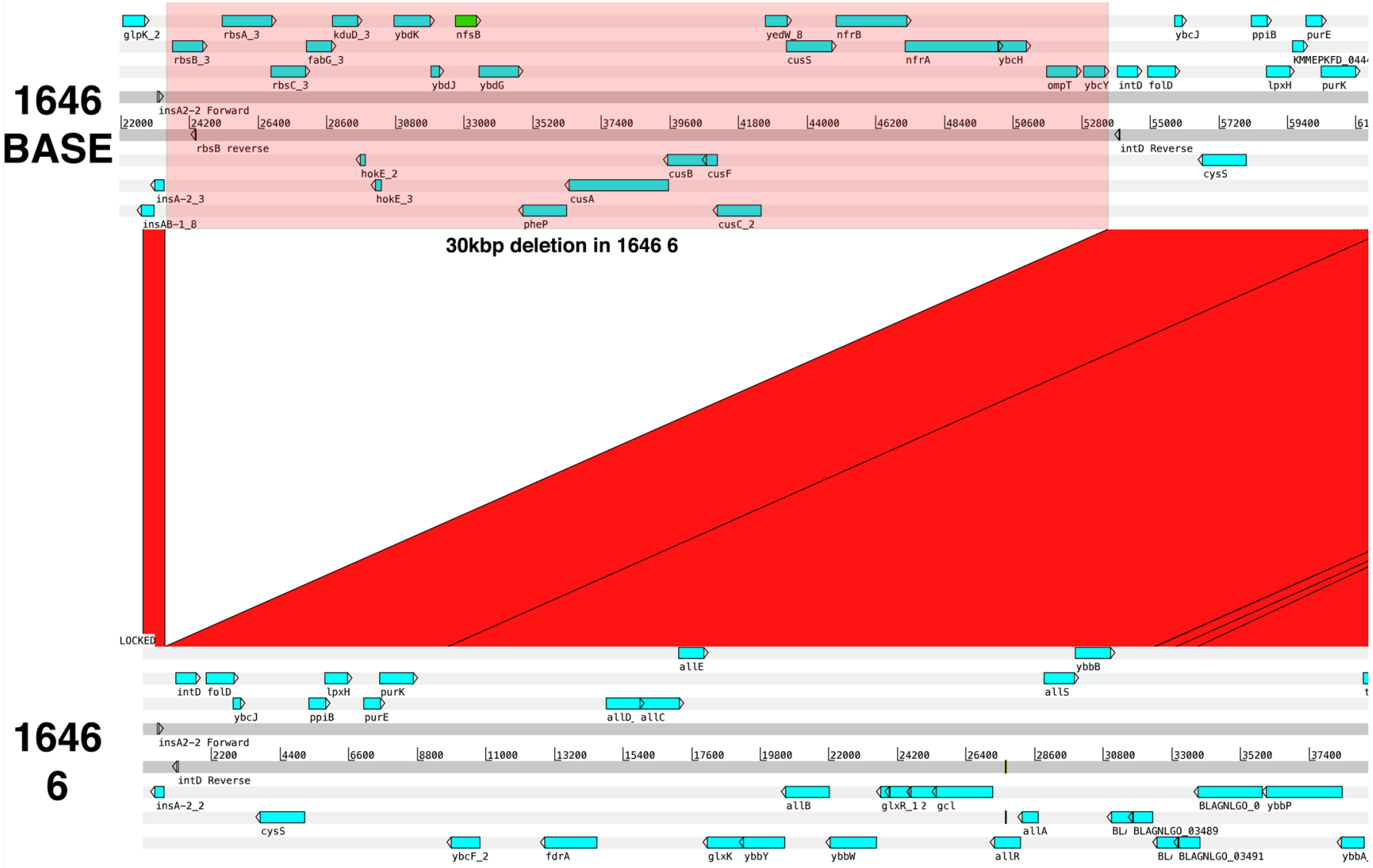



**Figure 45 Comparison of the *nfsB* region in Isolates 1646 BASE and 1646 6.** The red columns between 1646 BASE and 6 represent genomic areas with high sequence similarity between the two genomes. The region highlighted in pink represents the 30kbp deletion in 1646 6, there is no red column connecting the two genomes at this 30kbp area because of the deletion in 1646 6. The *nfsB* gene has been highlighted in green (1646 BASE). The figure was originally generated in Artemis and processed in Adobe Illustrator to highlight the genomic region associated with the 30kbp deletion.

### *3.4.2 Ciprofloxacin resistance in AnTIC E. coli isolates*

Ciprofloxacin and other fluoroquinolones accounted for 4/34 (11.8%) of all acute antibiotic prescriptions within the 17 AnTIC patients studied. Across the entire AnTIC dataset, ciprofloxacin and other fluoroquinolones accounted for 98/914 (10.7%) of all acute antibiotic prescriptions. For comparison, fluoroquinolones accounted for 4.1% of all UTI-related antibiotic prescriptions within the UK primary care setting from 2013-2015 [333]. As indicated in Figure 41 (purple boxes), and Table 29, ciprofloxacin was prescribed to patients for the management of acute UTIs (two non-prophylaxis and two prophylaxis patients). Three isolates were specified as being ciprofloxacin resistant ($Cip^R$) according to the AMR profile dataset (Table 29) but none of the isolates appeared to have acquired ciprofloxacin resistance during the study.

#### 3.4.2.1 Genetic mechanism of ciprofloxacin resistance in *E. coli*

Ciprofloxacin is a fluoroquinolone that exhibits its antimicrobial properties by binding to GyrA and ParC, preventing re-ligation of broken DNA strands during DNA replication. This leads to an accumulation of double-stranded breaks, which disrupt cell growth and eventually causing apoptosis [286]. Clinical resistance to ciprofloxacin (MIC > 0.5 μg/ml) in *E. coli* is associated with a two-step mutation in GyrA (S83L-D87N point mutations) that prevents the binding of ciprofloxacin to GyrA [287]. It should be noted that the first-step mutation (GyrA S83L) does support sub-clinical resistance to ciprofloxacin. Higher levels of resistance are associated with an additional point mutation in ParC (S80I or E84K) that is dependent on the presence of an existing GyrA mutation [288,289].

#### 3.4.2.2 *In silico* analysis of ciprofloxacin resistance in AnTIC isolates

The GyrA and ParC genes of the 47 *E. coli* isolates were analysed to determine whether the $Cip^R$ phenotypes were associated with the presence of specific point mutations in GyrA and ParC. All three $Cip^R$ isolates (1260 3, 1260 12, 2434 18) possessed the following mutations: GyrA S83L-D87N, ParC S80I (Table 29, highlighted in red). Of 44 ciprofloxacin sensitive ($Cip^S$) isolates, 15 possessed the GyrA S83L mutations whereas the remaining 29 $Cip^S$ isolates did not possess any relevant mutations in either GyrA or ParC (Table 29). The genotype data of all 47 isolates corresponded to the known mutational patterns of ciprofloxacin resistance, which negated the need to perform MIC assays to confirm their ciprofloxacin resistance phenotype.

There was no strong evidence to support the acquisition of ciprofloxacin resistance among the 47 sequenced *E. coli* isolates during the study. Unlike nitrofurantoin resistance, the emergence of $Cip^R$ isolates did not always correspond with ciprofloxacin usage. For example, a $Cip^R$ isolate (2434 18) emerged in patient 2434 despite the patient having no history of ciprofloxacin usage during the study (Table 29, Figure 41). Furthermore, 13/15 of the *E. coli* isolates with the GyrA S83L mutation was linked to six patients with no history of ciprofloxacin usage during the study (Table 29, Figure 41). The remaining two *E. coli* isolates (2433 9, 2877 9) originated in two patients (Patient 2433 and 2877) that were prescribed ciprofloxacin during the study (Table 29, Figure 41). However, these isolates emerged before the start of ciprofloxacin therapy. This is clearly demonstrated in Figure 41 where the isolates 2433 9 and 2877 9, represented by their strain numbers ST69 and ST767 respectively, were situated earlier in the timeline when compared to the positions of the purple boxes representing acute ciprofloxacin prescriptions,

The fact that all 15 $Cip^R$ isolates were isolated from patients not in receipt of ciprofloxacin suggests that low-level ciprofloxacin resistance is prevalent among *E. coli* isolates. Strain ST131 is notably overrepresented within this group with 9/15 isolates (Figure 46). Even without considering the clonal spread of successful MDR strains, ciprofloxacin resistance among urinary *E. coli* isolates is sufficiently commonplace, global incidence: 27%, UK: 15.3%. That is was, perhaps not a surprise to encounter *E. coli* isolates with GyrA mutations in patients with no ciprofloxacin usage [335].

Despite the higher proportional usage of ciprofloxacin in this study, the incidence of ciprofloxacin resistance among urinary *E. coli* isolates within the entire AnTIC dataset was lower compared to the average national incidence rate of CipR urinary *E. coli* isolates within the UK (12.5% vs 15.3%) [336].

**Table 29 Ciprofloxacin resistance phenotype and GyrA/ParC genotype of sequenced AnTIC isolates.**

| Patient | Donation month | Study Arm | Acute Cip. | CipR | GyrA | ParC |
|---|---|---|---|---|---|---|
| 13 | 3 | No prop. | No | S | | |
| | 6 | | | S | | |
| | 9 | | | S | | |
| 426 | BASE | No prop. | No | S | S83L | |
| | 3 | | | S | S83L | |
| | 9 | | | S | S83L | |
| | 18 | | | S | S83L | |
| 432 | BASE | No prop. | No | S | | |
| | 3 | | | S | | |
| | 6 | | | S | | |
| 824 | 9 | No prop. | No | S | | |
| | 18 | | | S | | |
| 1253 | 3 | No prop. | No | S | | |
| | 9 | | | S | | |
| | 12 | | | S | | |
| 1646 | BASE | No prop. | No | S | S83L | |
| | 3 | | | S | S83L | |
| | 6 | | | S | S83L | |
| 2433 | BASE | No prop. | Yes | S | | |
| | 3 | | | S | | |
| | 6 | | | S | | |
| | 9 | | | S | S83L | |
| | UTI 12 | | | S | | |
| | UTI 18 | | | S | | |
| 2877 | 3 | No prop. | Yes | S | | |
| | 9 | | | S | S83L | |
| 429 | 6 | Cefalexin prop. | No | S | S83L | |
| | 9 | | | S | S83L | |
| 1218 | 12 | Trimethoprim prop. | No | S | | |
| | 18 | | | S | | |
| 1260 | 3 | Trimethoprim prop. | Yes | R | S83L & D87N | S80I |
| | 12 | | | R | S83L & D87N | S80I |
| 2015 | BASE | Trimethoprim prop. | No | S | S83L | |
| | 12 | | | S | S83L | |
| 2417 | BASE | Nitro prop. | No | S | | |
| | 6 | | | S | | |
| 2434 | 12 | Trimethoprim prop. | No | S | | |
| | 18 | | | R | S83L & D87N | S80I |
| 2439 | BASE | Trimethoprim prop. | No | S | | |
| | 6 | | | S | S83L | |
| | 12 | | | S | S83L | |
| | 18 | | | S | | |
| 2470 | BASE | Trimethoprim prop. | Yes | S | | |
| | 3 | | | S | | |
| | 6 | | | S | | |
| 2889 | 6 | Trimethoprim prop. | No | S | | |
| | 9 | | | S | | |

The isolate's name is the combination of the patient's name and donation month separated by a space. For example, isolate 13 3 is Patient 13's *E. coli* isolate that was collected during month 3 of the study. The third column (study arm) shows which cohort the patient was assigned to during the AnTIC trial, no prop. (non-prophylaxis), cefalexin prop. (cefalexin prophylaxis), trimethoprim prop. (trimethoprim prophylaxis), Nitro prop. (nitrofurantoin prophylaxis). [Caption continues on the following page]

[Caption following on from the previous page]. The fourth column ($Cip^R$) shows whether an isolate was ciprofloxacin sensitive (S) or resistant (R). The fifth (GyrA) and sixth (ParC) columns shows the mutations that were present in the GyrA or ParC protein respectively, amino acid changes are shown for point mutations. Isolates highlighted in green denote nitrofurantoin sensitive ($Nit^S$) isolates. Isolates highlighted in pink denote nitrofurantoin resistant ($Nit^R$) isolates. Isolate highlighted in yellow denote $Nit^R$ isolates that were originally misclassified as being $Nit^S$. Isolates highlighted in red denote $Cip^R$ isolates.

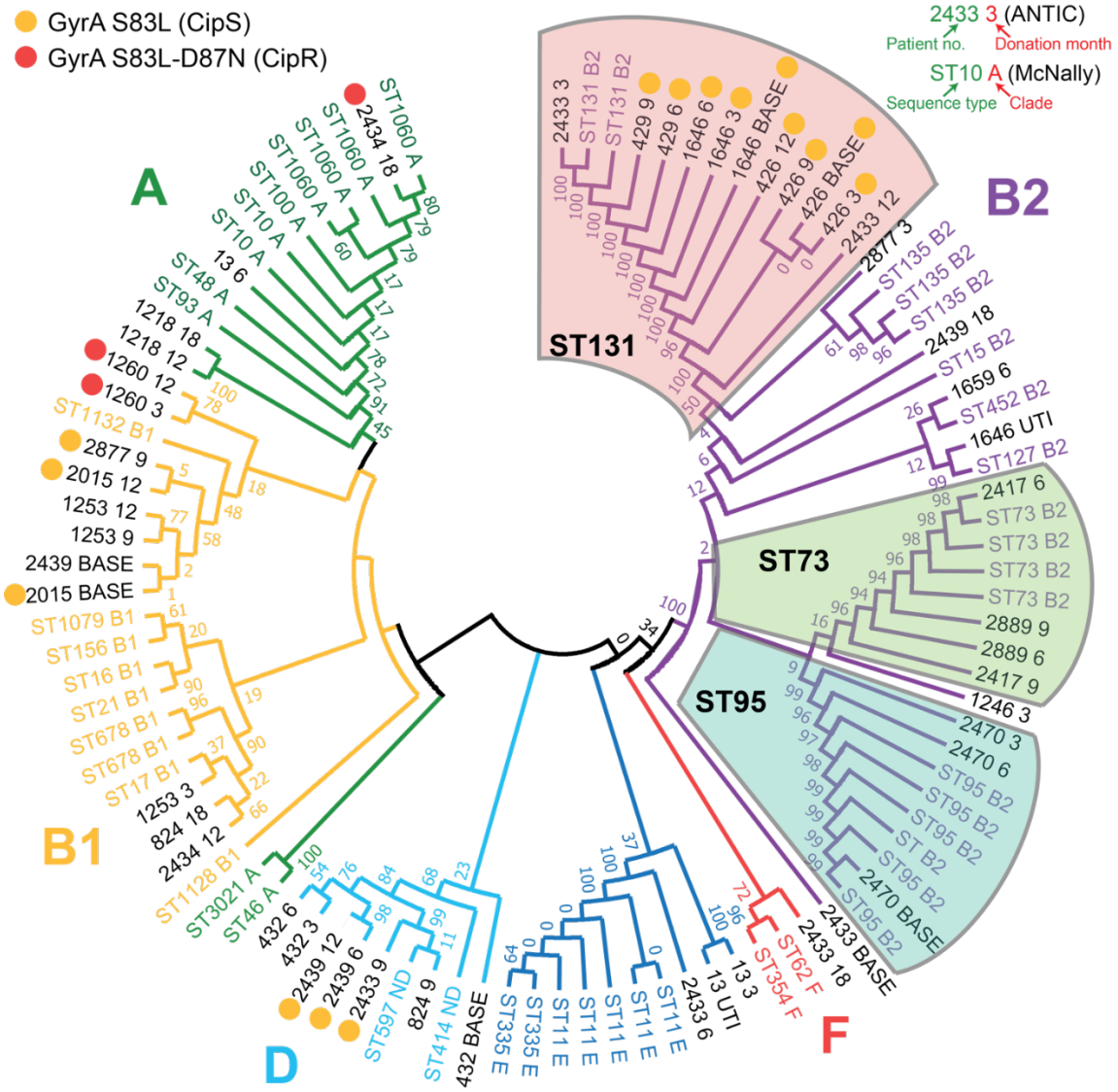



**Figure 46 7-gene MLST phylogenetic tree overlaid with relevant GyrA genotype data.** Isolates which possessed the first step S83L mutation were marked with a yellow circle whereas isolates which possessed both the S83L and D87N mutations were marked with a red circle.

## 3.5 Summary

All of the patients who were recruited into the AnTIC study were CISC patients who utilised urinary catheters on a regular, but intermittent basis and were clinically characterised as patients suffering complicated UTIs. However, the infrequent usage of urinary catheters among the general population means that the statistics cited in the literature or in public databases are mainly concerned with uncomplicated UTIs [337]. This does not mean that comparisons cannot be made between these two

groups, but in discussing any apparent statistical differences the effects of catheter usage on the urinary tract and its resident microflora should be considered. A key question is how can catheter usage drive this difference?

The urethral meatus is a microflora-rich region which includes the tissue surrounding the urethral opening and the distal urethra (Figure 2) [338]. This region cannot be avoided during catheter insertion, which means that individuals who practice intermittent catheterisation will transfer some of the meatus microflora directly into the bladder every time they need to void their bladder (20-150 CFUs per catheterisation, 4-6 times daily) [338–340].

Furthermore, catheter usage bypasses the mechanical and immunological barriers which protect the bladder against bacterial colonisation [338]. This increases the accessibility of the bladder environment to a wider variety of microorganisms and removes the need for such microbes to possess specific virulence factors to overcome these barriers [341,342]. With easy access, the bladder of CISC patients might therefore be a more competitive environment for microbes compared to uncomplicated UTI patients. In other words, catheter usage breaks down the physical barrier that normally exists between the bladder environment and the external urinary environment, which could effectively merge these two environments into a single compartment. The external urinary environment in both male and female individuals are proximal to other microflora rich regions besides the urinary meatus. In female patients, the external urinary environment is proximal to the vaginal environment whereas in male patients, it is proximal to the preputial environment [75,343].

Antibiotic prophylaxis treatment was shown through the AnTIC trial to reduce symptomatic UTIs in CISC patients by one case per year [14]. Moreover, data reported in this chapter supported this observation and showed that antibiotic prophylaxis reduced the incidence of bacterial colonisation. However, there was a subset of prophylaxis patients that still experienced persistent bacterial colonisation during AnTIC. Among these patients, *E. coli* colonisation was reduced, but not completely eliminated, and was accompanied by an increased incidence of minor uropathogens. AMR analyses showed that antibiotic resistance was more common among isolates that were classified as minor uropathogens compared to *E. coli* isolates, thereby providing a possible explanation for this observation.

However, a substantial number of AnTIC patients were colonised by *E. coli* despite being managed with antibiotic prophylaxis, and yet these remain asymptomatic. For a subset of these patients who acquired MDR bacteria during the AnTIC trial, antibiotic prophylaxis was associated with relatively stable urinary *E. coli* colonisation that linked to a reduce incidence of infections. The origin of these *E. coli* isolates was not explored, but it can be assumed they linked to the microflora of the urethral meatus, vaginal, and preputial environments. There is also the possibility of bacterial transfer between the gut and the urinary tract via the perineum, especially among females [8]. However, secondary findings from the AnTIC study downplays this possibility as there was no increase in resistance to any antibiotics among perianal *E. coli* isolates [14].

Interestingly, the gut microflora may constitute an important infection reservoir among non-catheterised patients with a murine study reporting selective depletion of UPEC intestinal reservoirs to be associated with decreased UTI incidence [180]. Therefore, the primary infection reservoir of UTI may differ between non-catheterised and catheterised patients. For example, among non-catheterised individuals, the gut might be the primary infection reservoir whereas among CISC individuals, the urinary meatus, vaginal and preputial environments may be the primary infection reservoir [338,344]. The frequent introduction of a wider variety of bacterial isolates into the bladder environment by catheterisation may greatly increase competitive pressure between urinary isolates, which may in turn affect the emergence and prevalence of antibiotic resistance [345,346]. Acquisition of antibiotic resistance is generally associated with a fitness cost, this fitness cost may differ depending on the mechanism of resistance and the genetic background of the bacterial isolate [347].

According to NICE guidelines, nitrofurantoin is recommended for UTI treatments in the UK [30]. Nitrofurantoin resistance among *E. coli* isolates recovered from AnTIC patients was higher when compared to that of the general population. This was unexpected given the lower proportional usage of nitrofurantoin among AnTIC patients. Even after excluding patients who were on nitrofurantoin prophylaxis during the study (they were not eligible for acute nitrofurantoin prescriptions), the overall incidence of $Nit^R$ *E. coli* isolates among the remaining patients was high compared to the general population (6.1% vs 2.1%) [334]. Conversely, the incidence of ciprofloxacin resistance among *E. coli* isolates from AnTIC patients was lower compared to the general population despite the higher proportional usage of ciprofloxacin among

AnTIC patients [333,348]. In *E. coli*, ciprofloxacin resistance comes with a higher fitness cost compared to nitrofurantoin resistance [349]. This means that any urinary $Cip^R$ isolates are possibly displaced quickly by fitter $Cip^S$ isolates following an acute therapy. However, $Nit^R$ isolates are more likely to persist in the bladder as they are still able to compete with $Nit^S$ isolates in an antibiotic-free environment.

The ability of antibiotic resistant strains to persist in such a competitive environment depends on their fitness in addition to their resistance. Among non-catheterised patients, the fitness of resistant isolates plays a less important role as the bladder is more of a “niche” environment and less accessible to potential competitors. Among CISC patients, there is a higher incidence of resistance against antibiotics where resistance does not substantially undermine fitness. Therefore, it might be necessary to reassess the types of treatments prescribed to CISC patients. This should not only improve treatment outcome but also minimise the emergence of antibiotic resistance.

Future studies involving CISC patients should take into consideration the microflora of the urinary meatus, vagina, and preputial space as the close proximity of these regions to the site of catheter insertion makes them potential reservoirs of infection and resistance. In line with this, it might also be worth investigating whether the application of topical antibiotics to these regions might decrease the incidence of CAUTI among CISC patients.

Previous studies which investigated this treatment option among non-catheterised subjects found that it was ineffective at reducing the incidence of UTI [350]. This is not surprising as it failed to target the primary infection reservoir of non-catheterised individuals, which is likely perianal or rectal [180]. The ease and acceptability of accessing the periurethral and perivaginal regions which constitute the primary infection reservoirs of CISC individuals makes the usage of topical antibiotics a viable and acceptable alternative to systemic antibiotics should it prove to be effective. This is an exciting clinical prospect as the effects of topical antibiotics are more localised and thus would be associated with a lower incidence of adverse outcomes and antibiotic resistance [351].

# Chapter 4 Investigating the fitness dynamics underlying the acquisition of nitrofurantoin resistance by *Escherichia coli*

## 4.1 Introduction

### 4.1.1 Introduction to nitrofurantoin, historical and recent clinical usage

Nitrofurantoin is a broad-spectrum antibiotic, which been in clinical usage since the mid-1950s to manage uncomplicated UTIs [257]. Its popularity waned in the 1970s with the introduction of trimethoprim and modern β-lactam antibiotics, which were more suitable for individuals with decreased renal function [257,258]. However, there has been renewed interest in nitrofurantoin as a therapeutic, as resistance amongst uropathogens against trimethoprim and β-lactams has become increasingly common over the last few decades [257]. The pharmacokinetics and pharmacodynamics of nitrofurantoin as well as the resistance mechanisms against it have been discussed (Section 1.7.2).

### 4.1.2 Investigating the low incidence of nitrofurantoin resistance

Nitrofurantoin is, however, unique among antibiotics in that it enjoys a low incidence of resistance despite having been in active clinical usage for over 60 years [257]. Although counter-intuitive this may explain why few efforts have been made to better understand the evolutionary factors driving nitrofurantoin resistance, despite the growing problem of antibiotic resistance worldwide. To date, only one study published in the 2008 Journal of Antimicrobial Chemotherapy has attempted to investigate the factors contributing to the low incidence of nitrofurantoin resistance among *E. coli* isolates [271].

This study found that the growth rate of clinical $Nit^R$ isolates was 6% slower compared to clinical $Nit^S$ isolates [271]. Resistant mutants generated by *in-vitro* selection also showed a 1-3% slower growth rate compared to their parental strain [271]. All resistant isolates were unable to grow at 200 µg/ml nitrofurantoin, which was presumed to be the minimum nitrofurantoin concentration in the urinary tract during treatment [271]. Molecular analyses, specifically PCR amplification and sequencing revealed that all resistant isolates possessed inactivating mutations in both *nfsA* and *nfsB* [271]. The study concluded that the fitness cost of inactivating *nfsA*/*nfsB* and more importantly, the inability of resistant isolates to grow at 200 µg/ml nitrofurantoin, function to prevent the establishment of resistant isolates in the urinary tract, thereby

contributing to the low incidence of nitrofurantoin resistance among urinary *E. coli* isolates [271].

The 2008 study is not without its weaknesses. Clinical isolates were compared without considering genetic background, thereby distorting any fitness difference between $Nit^S$ and $Nit^R$ isolates. Indeed, the innate fitness of an isolate and the fitness cost of antibiotic resistance can greatly differ depending on genetic background [352]. The only evolutionarily credible comparison made in the study was between the *in vitro* resistant mutants and their parental isolate, with a notably smaller fitness difference (3% vs 6%) between the two [271].

Additionally, a recent pharmacokinetic study reported the **maximum** urinary nitrofurantoin concentration during treatment to be in the region of 94 µg/ml [267], which is less than half the presumed minimum urinary nitrofurantoin concentration of 200 µg/ml used by Sandegren et al. (2008) [271]. Furthermore, urinary nitrofurantoin concentration does not remain constant but fluctuates in a downward pattern during therapy [267] (Figure 19). It could be argued from these data that even with optimal patient compliance, there will be extended periods when urinary nitrofurantoin concentration is low enough [267] to enable even single *nfsA* mutants (MIC: 12-16 µg/ml) to proliferate at a rate that is sufficient to maintain their presence in the urinary tract [271] (Figure 19). Such fluctuations may provide a window of opportunity for these mutants to establish themselves in the urinary tract by outcompeting their more sensitive parental isolates, providing a "stepping stone" for the eventual emergence of double mutants (with inactivated *nfsA* and *nfsB*) and high-level nitrofurantoin resistance.

Therefore, the aim of the studies reported in this chapter was to understand the evolutionary dynamics driving the low incidence of nitrofurantoin resistance. This was attempted by investigating the fitness impact of the *nfsA*/*nfsB* mutations that confer nitrofurantoin resistance in a systematic and evolutionarily meaningful manner.

## *4.2 Targeted mutagenesis of clinical and laboratory isolates*

As reported in the previous chapter, the *E. coli* isolates of patient 1646 developed nitrofurantoin resistance *in vivo* during the AnTIC study. The earlier isolates (1646 BASE and 1646 3) had intact *nfsA* and *nfsB* genes, with MIC assay findings confirming their $Nit^S$/$Nit^R$ phenotypes (Section 3.4.1.1, Figure 43). The 6-month

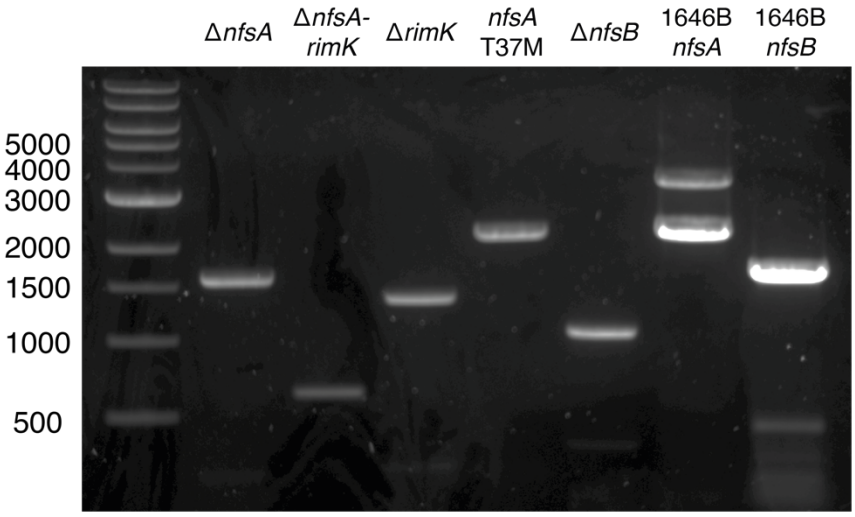
isolate of this patient, however, was characterised by an inactivating point mutation in *nfsA* and a large deletion in the *nfsB* region, resulting in high-level nitrofurantoin resistance as confirmed via MIC assay. Therefore, the *E. coli* isolates of Patient 1646 represented chronological “snapshots” of the same strain before and after the development of nitrofurantoin resistance. These clinical strains also represented a unique opportunity to investigate the fitness dynamics, if any, underlying the acquisition of nitrofurantoin resistance.

Mutants in *nfsA* and/or *nfsB*, were generated from the starting $Nit^S$ isolate of Patient 1646 (1646 BASE) and a laboratory *E. coli* strain (W3110) that is commonly used as the “wild-type” model organism in a wide range of microbiological research. By using two isolates of contrasting genetic backgrounds (clinical vs laboratory), it was possible to investigate the extent to which the phenotypic effects, impacting resistance and fitness of the *nfsA/nfsB* inactivating mutations were dependent on genetic background.

For these analyses, *nfsA* and *nfsB* mutants were generated by targeted mutagenesis rather than *in-vitro* selection. This approach was taken to reduce the risk of spontaneous mutations distorting the actual fitness cost of the *nfsA*/*nfsB* inactivating mutations. Inactivation of *nfsA* and *nfsB* was achieved by replacing these wild-type regions with DNA constructs generated by Gibson assembly or PCR amplification (Table 30; Figure 47).

**Table 30 Donor DNA templates generated from either Gibson assembly or PCR amplification**

| | |
|---|---|
| ***nfsA* donor DNA templates (Gibson)** | Δ*nfsA*, Δ*nfsA-rimK*, Δ*rimK* |
| ***nfsA* donor DNA template (1646 6)** | *nfsA* T37M (C110T) |
| ***nfsB* donor DNA templates (Gibson)** | Δ*nfsB* |
| ***nfsB* donor DNA template (1646 6)** | Δ*nfsB30kbp* |




| Donor DNA | Size |
|---|---|
| Δ*nfsA* | 1614 bp |
| Δ*nfsA-rimK* | 689 bp |
| Δ*rimK* | 1416 bp |
| *nfsA* T37M | 2298 bp |
| Δ*nfsB* | 1061 bp |
| 1646B *nfsA* | 2298 bp |
| 1646B *nfsB* | 1697 bp |

**Figure 47 Screening of donor DNA fragments via gel electrophoresis.** The upper band in 1646B *nfsA* (Lane 6) was likely caused by non-specific binding and amplification of the primers that were used to amplify the *nfsA* region. The Δ*nfsB30kbp* donor DNA is shown in Figure 44B, column 3 (1646 6).

Partial deletions, which inactivate *nfsA* and *nfsB* are well established as the causative factor of nitrofurantoin resistance in *E. coli* [273]. Hence, the *nfsA* and *nfsB* wild-type genes were inactivated via complete deletion (Δ*nfsA* and Δ*nfsB*) in the generated mutants; this ensured that they were completely non-functional. Additionally, Δ*rimK* and Δ*nfsA-rimK* mutations were also created as there were several Nit$^{R}$ AnTIC isolates identified that carried partial deletions in *rimK* in addition to complete/partial deletions in *nfsA*. The *rimK* gene encodes an enzyme involved in the post-translational modification of ribosomal protein S6 via the addition of glutamate residues [353]. Ribosomal protein S6 is a known component of the 30S ribosomal subunit, but other than that, little is known about its exact function within the ribosome and the effects of *rimK*-mediated glutamation on ribosomal function [353].

The proximity of *rimK* to *nfsA* and its involvement in what appears to be an essential cellular process could provide an alternative explanation to the low incidence of nitrofurantoin resistance. The acquisition of nitrofurantoin resistance is highly dependent on *nfsA* being inactivated, usually via partial/complete deletion. However, the random nature of deletions and the proximity of *rimK* to *nfsA* means that *rimK* may frequently be inactivated in addition to *nfsA*. Any modifications to *rimK* might

carry a substantial fitness cost due to its involvement in an essential cellular process. Therefore, the proximity of *nfsA* to *rimK* might constitute an evolutionary bottleneck in the acquisition of nitrofurantoin resistance. Investigating the fitness impact of the Δ*rimK* and Δ*nfsA-rimK* mutations allowed this hypothesis to be further explored.

The usual method for performing targeted mutagenesis involves using CRIPSR Knock-ins. However, it was decided that a novel two-step mutagenesis approach that employed lambda-red and CRISPR was better suited for the task [354]. The approach was a modification of the classic landing pad (LP) technique and substituted the I-SceI endonuclease and its cleavage sites with functional equivalents from the CRISPR system (Table 31) [355].

**Table 31 Comparison of the classic landing pad (LP) vs CRISPR LP technique**

| Classic landing pad (LP) technique | CRISPR landing pad (LP) technique |
|---|---|
| I-SceI (endonuclease + target sequence binding) | Cas9 (endonuclease)<br>gRNA (target sequence binding) |
| I-SceI cleavage site (target sequence)<br>Antibiotic resistance gene (selection gene) | *cat* (target sequence and selection gene) |

To generate the *nfsA* and *nfsB* mutants, Lambda red was first used to generate the intermediate strain from 1646 BASE and W3110 by replacing the *nfsA*/*nfsB* region with a chloramphenicol resistance (*cat*) gene. The function of the *cat* gene was twofold: (i) positive selection of successful recombinants and (ii) highly unique gRNA target sequence for CRISPR mutagenesis. The intermediate strain was then subjected to CRISPR mutagenesis that replaced the *cat* gene with an appropriate donor DNA template synthesised via Gibson assembly. The donor DNA templates were modified *nfsA*/*nfsB* regions that had specific gene deletions, which were associated with the acquisition of nitrofurantoin resistance (Table 30).

The mutants generated from 1646 BASE and W3110 are shown in Figure 48. It was only possible to generate viable mutants with the 30kbp deletion in the *nfsB* region from 1646 BASE. W3110 lacked the necessary homologous regions, i.e. *insA2* (Figure 44), within its *nfsB* region to allow for successful homologous recombination repair with the Δ*nfsB30kbp* donor DNA during CRISPR mutagenesis. The *nfsA* and *nfsB* regions of every generated mutant were amplified and sequenced to confirm successful mutagenesis.

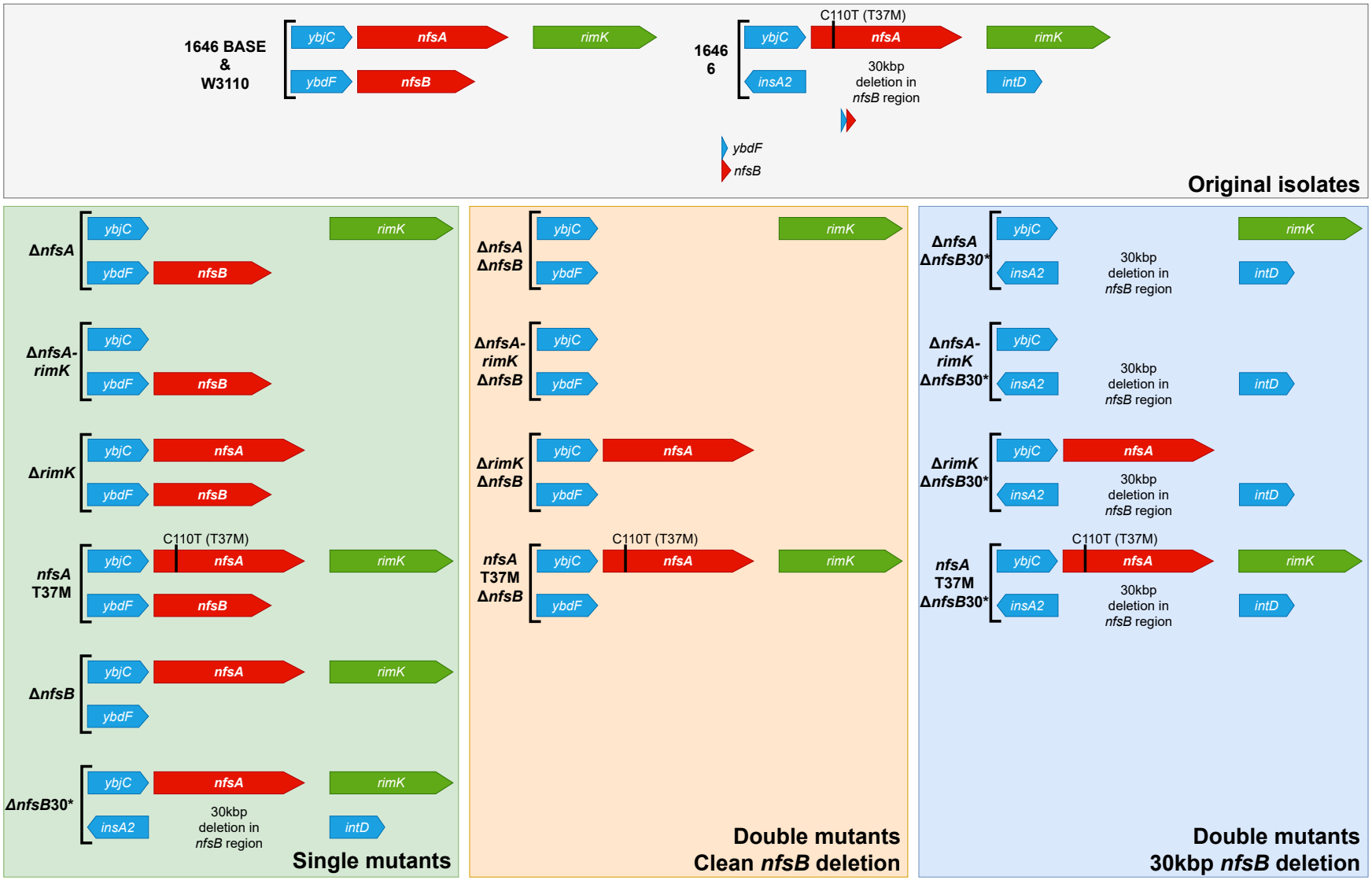



**Figure 48 *nfsA* and *nfsB* single and double mutants generated by lambda-red/CRISPR mutagenesis.** Mutants marked with * were only generated from 1646 BASE

## 4.3 MIC assay findings of nfsA/B mutants

The generated *nfsA* and *nfsB* mutants were subjected to MIC assays (Section 2.8) to quantitatively assess the level of nitrofurantoin resistance conferred by the different mutations. Isolates 1646 BASE and 1646 6 were included as $Nit^S$ and $Nit^R$ controls for the 1646 BASE mutants, and W3110 was the $Nit^S$ control for the W3110 mutants. Modified Clinical and Laboratory Standards Institute (CLSI) breakpoints were used to determine whether an isolate was $Nit^S$ or $Nit^R$ (Table 32) [356]. In clinical practice, isolates with intermediate resistance against nitrofurantoin ($Nit^I$) are treated as resistant isolates [356,357].

**Table 32 CLSI and modified CLSI MIC breakpoints.**

| | **CLSI** | **Modified CLSI** |
|---|---|---|
| Sensitive ($Nit^S$) | ≤ 32 µg/ml | ≤ 32 µg/ml |
| Intermediate ($Nit^I$) | 64 µg/ml | |
| Resistant ($Nit^R$) | ≥ 128 µg/ml | ≥64 µg/ml |

**Table 33 MIC assay findings of 1646B and W3110 *nfsA*/*nfsB* mutants.**

| **Isolate (Clinical *E. coli*)** | **MIC (µg/ml)** |
|---|---|
| 1646 BASE (1646B) | 8 |
| 1646 6 | 64 |
| 1646B Δ*nfsA* | 16 |
| 1646B Δ*nfsA-rimK* | 16 |
| 1646B Δ*rimK* | 8 |
| 1646B *nfsA* T37M | 16 |
| 1646B Δ*nfsB* | 8 |
| 1646B Δ*nfsA* Δ*nfsB* | 128 |
| 1646B Δ*nfsA-rimK* Δ*nfsB* | 128 |
| 1646B Δ*rimK* Δ*nfsB* | 64 |
| 1646B *nfsA* T37M Δ*nfsB* | 64 |

| **Isolate (Laboratory *E. coli*)** | **MIC (µg/ml)** |
|---|---|
| W3110 | 4 |
| | |
| W3110 Δ*nfsA* | 8 |
| W3110 Δ*nfsA-rimK* | 8 |
| W3110 Δ*rimK* | 4 |
| W3110 *nfsA* T37M | 8 |
| W3110 Δ*nfsB* | 4 |
| W3110 Δ*nfsA* Δ*nfsB* | 64 |
| W3110 Δ*nfsA-rimK* Δ*nfsB* | 64 |
| W3110 Δ*rimK* Δ*nfsB* | 32 |
| W3110 *nfsA* T37M Δ*nfsB* | 64 |

The 1646 BASE isolate had a MIC value of 8 µg/ml whereas 1646 6 had a MIC value that was eight times greater, i.e 64 µg/ml (Table 33). Single 1646 BASE (1646B) mutants with inactivated *nfsA* (Δ*nfsA*, Δ*nfsA-rimK*, *nfsA* T37M) supported MIC values of 16 µg/ml (Table 33). The other single 1646B mutants (Δ*rimK* and Δ*nfsB*) also had a MIC of 8 µg/ml, the same MIC as 1646 BASE. Collectively, these data indicated that only single mutants with inactivated *nfsA* had a higher MIC when compared to their parental isolate.

All 1646B double mutants were Nit$^R$ with MICs ≥64 µg/ml (Table 33). Double mutants in which *nfsA* was inactivated via complete deletion (Δ*nfsA* Δ*nfsB,* Δ*nfsA-rimK* Δ*nfsB*) had higher MICs (128 µg/ml) compared to 1646 6 and its corresponding mutant (*nfsA* T37M Δ*nfsB*) in which *nfsA* was inactivated via a point mutation. These data indicated that the inactivation of *nfsA* via deletion was more likely to completely eliminate NfsA nitroreductase expression/activity in comparison with inactivation via point mutation, thereby granting increased resistance. These data probably explained why there is a higher incidence of *nfsA* mutants with partial deletions compared to point mutations among clinical isolates. It was surprising to note that the MIC of the Δ*rimK* Δ*nfsB* mutant was comparable to 1646 6 despite having a functional *nfsA* gene (Table 33).

A similar pattern was observed for the W3110 single and double mutants. However, the W3110 parental isolate was notably less resistant against nitrofurantoin compared to the 1646 BASE parental isolate as evidenced by its lower MIC of (4 µg/ml vs 8 µg/ml) (Table 33). In fact the MICs of all W3110 mutants was half that of their 1646 BASE counterparts (Table 33).

Therefore, these data suggested that the level of nitrofurantoin resistance conferred by the *nfsA* and *nfsB* mutations was dependent on the innate resistance of the parental isolate, evidenced by the difference in MICs between comparable 1646 BASE (8 µg/ml) and W3110 mutants (4 µg/ml). As expected, double mutants in which both *nfsA* and *nfsB* were inactivated were Nit$^R$ and had a substantially higher MIC compared to their parental isolate or single mutants. Single mutants with inactivated *nfsA* were more resistant compared to their parental isolate, providing a selective advantage at lower nitrofurantoin concentrations.

### *4.4 Investigating the fitness impact of the nfsA and nfsB mutations*

The maximum bacterial growth rate in terms of doubling time was used to gauge the relative fitness of the mutants, with shorter doubling times corresponding to increased fitness. Mutants were compared to their corresponding parental isolates (1646 BASE or W3110) to determine whether the *nfsA/nfsB* mutations significantly impacted doubling times in the absence of nitrofurantoin and at increasing nitrofurantoin concentrations of 2, 4, 8, 16, 32, 64, and 128 µg/ml. Following this, the fitness of single mutants were also compared to their corresponding double mutant i.e. Δ*nfsA* vs Δ*nfsA* Δ*nfsB*, Δ*nfsA-rimK* vs Δ*nfsA-rimK* Δ*nfsB, nfsA* T37M vs *nfsA* T37M Δ*nfsB*.

Fitness data are shown in Figure 49 & Figure 50 and significance data are shown in Appendix A & Appendix B respectively. In the absence of nitrofurantoin (0 μg/ml), there was no significant difference in doubling time between any of the mutants and their corresponding parental isolates. Single and double mutants with functional *nfsA* (Δ*rimK*, Δ*nfsB*, Δ*nfsB*30, Δ*rimK* Δ*nfsB*, Δ*rimK* Δ*nfsB*30) had equivalent doubling times to their parental isolates at all viable concentrations (2, 4, and 8 μg/ml of nitrofurantoin). It was only single and double mutants with inactivated *nfsA* (Δ*nfsA*, Δ*nfsA-rimK, nfsA* T37M, Δ*nfsA* Δ*nfsB*, Δ*nfsA-rimK* Δ*nfsB, nfsA* T37M Δ*nfsB*) that showed a significant decrease in doubling time when compared to their parental isolates at 8 μg/ml of nitrofurantoin, with the average doubling time of these mutants reduced by 35% ($p<0.001$). At lower nitrofurantoin concentrations (2, 4 μg/ml), the doubling time of these mutants was similar to their parental isolates.

There was no significant difference in doubling times between the single and their corresponding double mutants at 2, 4, and 8 μg/ml of nitrofurantoin respectively (Figure 49, Figure 50). At 16 μg/ml of nitrofurantoin, only 1646 6; 1646 BASE single and double mutants with inactivated *nfsA*,  and W3110 double mutants with inactivated *nfsA* remained viable. For the 1646 BASE mutants, the double mutants with inactivated *nfsA* grew significantly faster (46%) compared to their single mutant counterparts by 46% (Figure 49 & Figure 50, Appendix A & Appendix B). A similar comparison could not be made for the corresponding W3110 mutants as the W3110 single mutants with inactivated *nfsA* were non-viable at 16 μg/ml of nitrofurantoin (Figure 50, Appendix B).

At 32 and 64 μg/ml, only double mutants with inactivated *nfsA* (Δ*nfsA* Δ*nfsB*, Δ*nfsA-rimK* Δ*nfsB, nfsA* T37M Δ*nfsB*) were able to proliferate (Figure 49, Figure 50). Surprisingly, W3110 Δ*nfsA* Δ*nfsB* and W3110 Δ*nfsA-rimK* Δ*nfsB* were able to proliferate at their minimum inhibitory concentration of 64 μg/ml (Table 33, Figure 50B). For some mutants, e.g. Δ*rimK* Δ*nfsB*, the MIC and growth curve assays suggested different inhibitory concentrations (Section 4.3, 4.4). At 128 μg/ml, none of the mutants/isolates were able to proliferate.

When the relevant comparisons were made, data suggested that the method by which *nfsA* (Δ*nfsA*, *nfsA* T37M) or *nfsB* (Δ*nfsB*, Δ*nfsB*30) was inactivated did not significantly impact doubling times in either single or double mutants at all tested concentrations (Figure 49,

Figure 50, Appendix A, Appendix B, Appendix C). Therefore, it was concluded that the inactivation of these genes was more important than the manner in which they were inactivated. Furthermore, the inactivation of *rimK* in addition to the inactivation of *nfsA* did not significantly improve nor diminish doubling time. It is likely that the inactivation of *rimK* in certain $Nit^R$ AnTIC isolates was coincidental due to its proximity to *nfsA*.

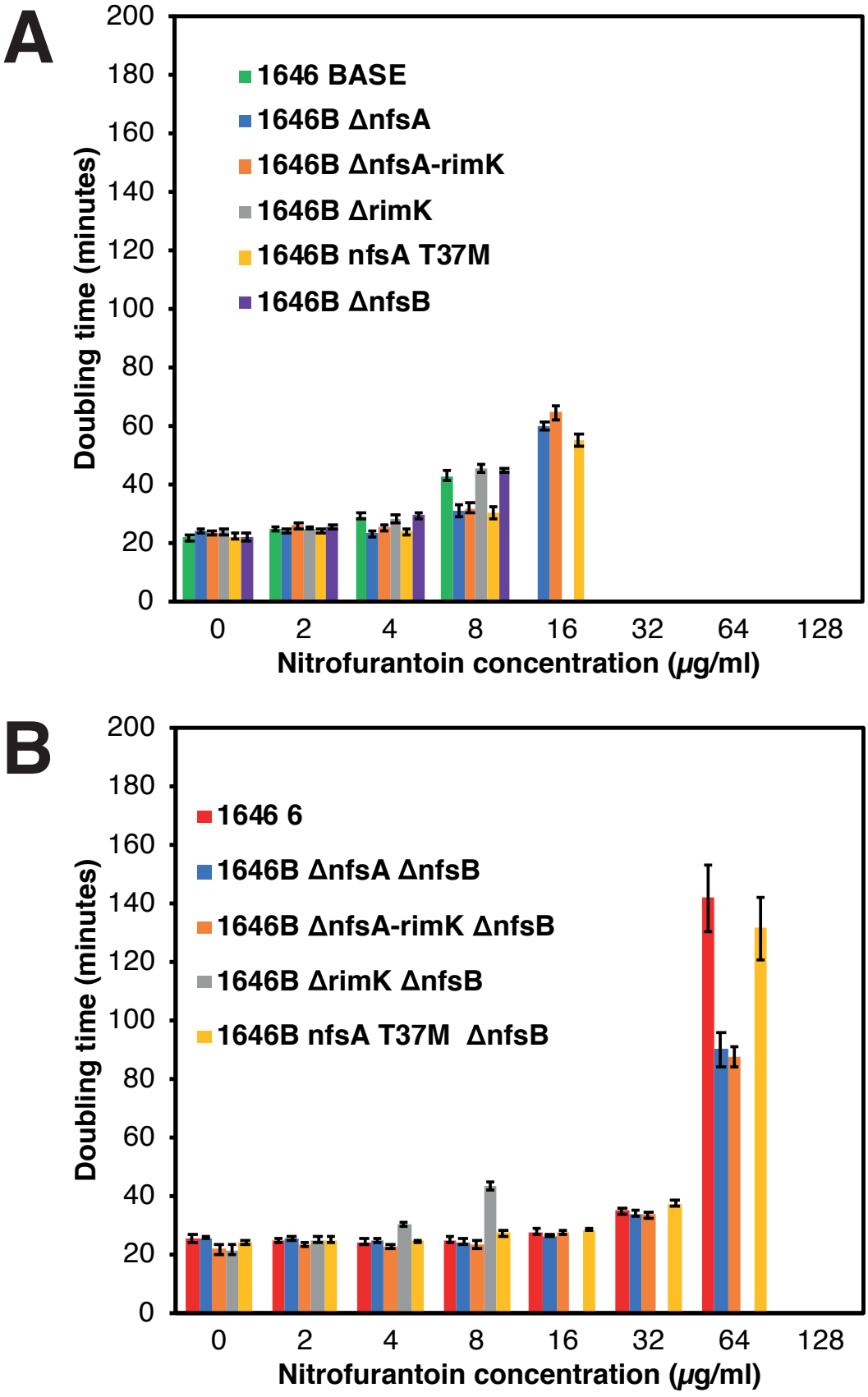



**Figure 49 Doubling time of *nfsA/nfsB* single (A) and double (B) mutants generated from 1646 BASE.** Within each antibiotic concentration, the growth rate (doubling time) of the isolates was statistically compare against each other. Refer to Appendix A for the doubling times and signficance data.

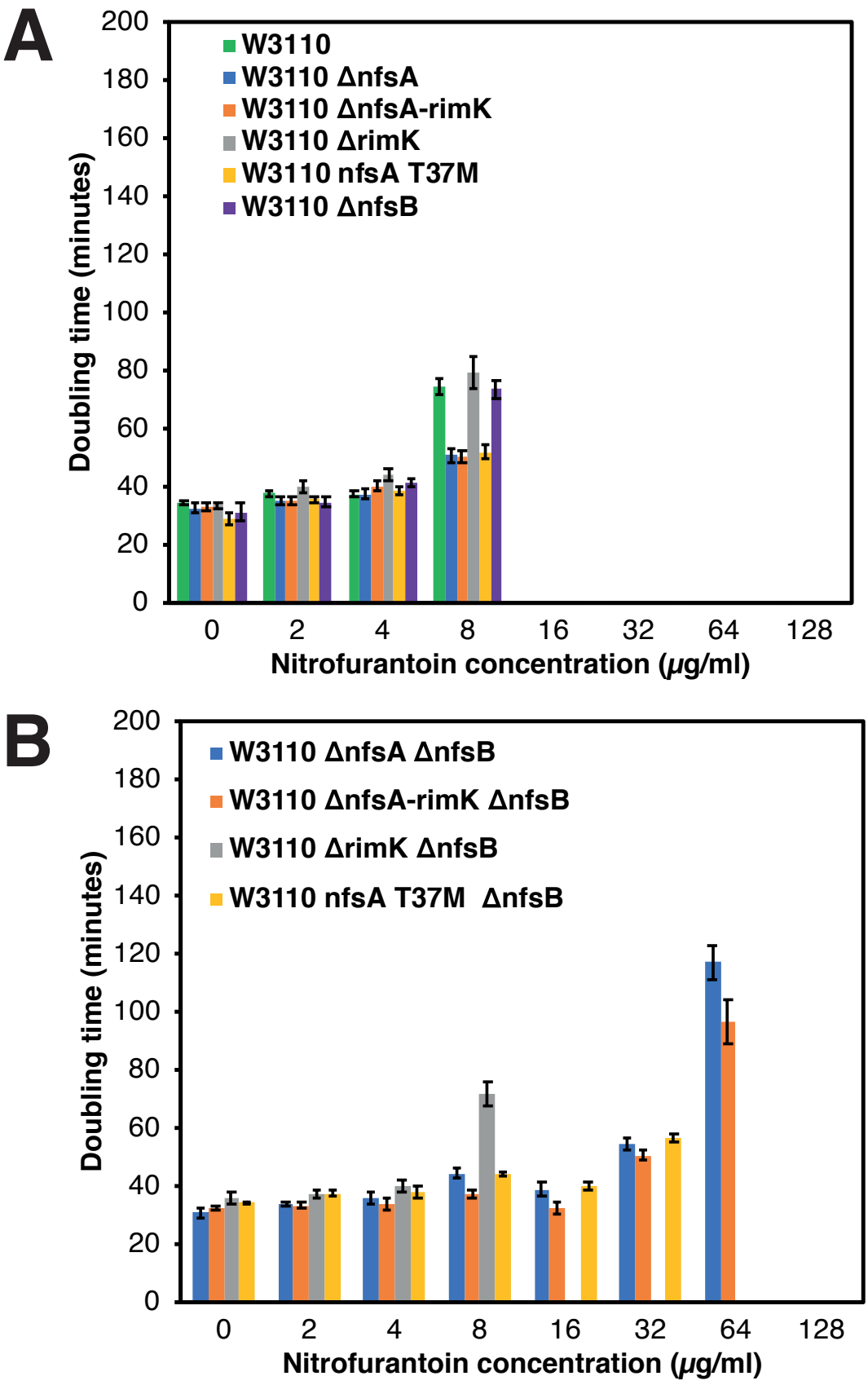



**Figure 50 Doubling time of *nfsA/nfsB* single (A) and double (B) mutants generated from W3110.** Within each antibiotic concentration, the growth rate (doubling time) of the isolates was statistically compare against each other. Refer to Appendix B for the doubling times and signficance data.

Contrary to the previous study by Sandegren et al. (2019) [271], mutations that inactivated *nfsA* and/or *nfsB* were not associated with a significant fitness cost as there was no growth difference between mutants and their parental isolates in the absence of nitrofurantoin. Moreover, when the isolates were grown in the presence of nitrofurantoin, certain mutations bestowed a fitness advantage. For example, at 8 μg/ml of nitrofurantoin, single mutants with inactivated *nfsA* (Δ*nfsA,* Δ*nfsA-rimK, nfsA* T37M) were fitter than isolates with intact *nfsA* (1646 BASE, W3110, Δ*rimK,* Δ*nfsB,* Δ*rimK* Δ*nfsB*) (Figure 49 & Figure 50). Similary, at 16 μg/ml of nitrofurantoin, 1646 mutants with both inactivated *nfsA* and *nfsB* (Δ*nfsA* Δ*nfsB*, Δ*nfsA-rimK* Δ*nfsB, nfsA* T37M Δ*nfsB*) were fitter than their corresponding single mutants (Δ*nfsA,* Δ*nfsA-rimK, nfsA* T37M respectively). Together with the MIC findings, these data support the two-step pathway in developing nitrofurantoin resistance, with the crucial first step being the inactivation of *nfsA* followed by the inactivation of *nfsB*.

Additionally, all isolates/mutants showed a substantial increase in doubling time at a nitrofurantoin concentration that was one step below (0.5x) their minimum inhibitory concentration. Using 1646 Δ*nfsA* Δ*nfsB* as an example, there was no substantial increase in its doubling time from 0-32 μg/ml of nitrofurantoin, but when the nitrofurantoin concentration was 64 μg/ml, its doubling time increased to 90.5 minutes, a 2.6-fold increase over its doubling time of 34.2 minutes at 32 μg/ml of nitrofurantoin (Figure 49B, Appendix A). Considering that the maximum urinary nitrofurantoin concentration has been reported to be in the region of 94 μg/ml [267], these *in vitro* data suggest that even $Nit^R$ isolates will struggle to grow clinically.

## *4.5 Investigating the competitive fitness of the nfsA and nfsB mutants*

Previous studies have demonstrated that maximum growth rate is only one of the many components of bacterial fitness with growth assays providing a guide of bacterial fitness [358]. To definitively gauge bacterial fitness, it is necessary to perform growth competition assays using conditions where isolates of interest are grown together in the same culture from a small seeding population [358]. Using a small seeding population maximises the duration of the exponential phase, allowing for better discrimination of fitness between the tested isolates [358].

Growth competition assays were performed with the 1646 BASE *nfsA*/*nfsB* mutants, 1646 BASE, and 1646 6. To differentiate between the competing isolates, Δ*uidA* mutants were

generated. The *uidA* gene encodes beta-glucuronidase, whose catalytic activity is used as an *E. coli* specific chromogenic marker on Biomerieux CPSE plates. CPSE plates are differential agar plates used routinely to identify UTI-associated pathogens. By deleting *uidA*, *E. coli* mutants form colourless colonies instead of the usual purple-coloured colonies associated with *E. coli*, allowing quick, easy visual differentiation of isolates (Figure 51).

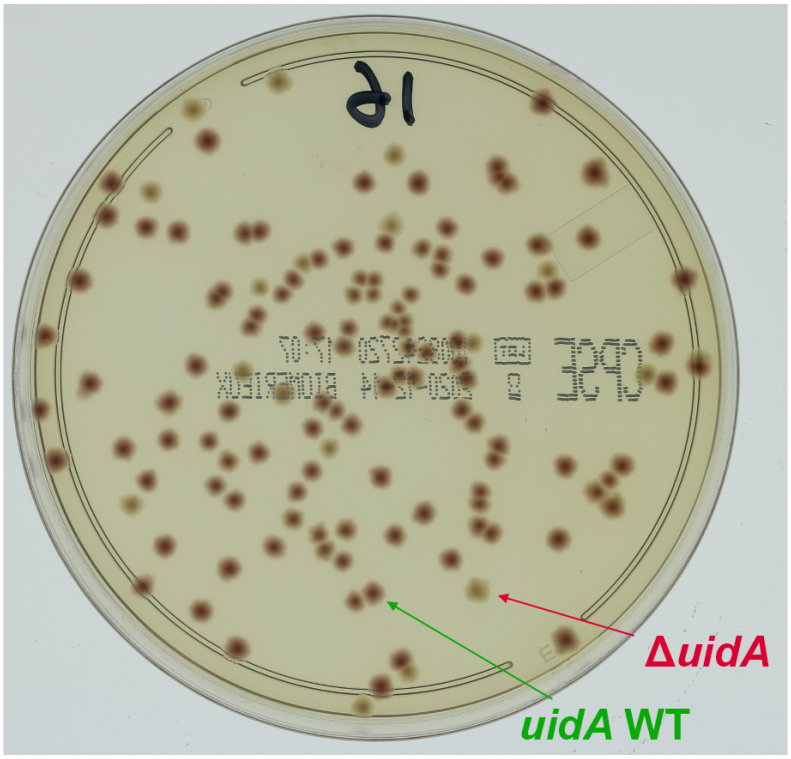



**Figure 51 *uidA* wild-type (WT) and Δ*uidA* colonies on a CPSE plate.** *uidA* WT colonies are easily distinguishable from Δ*uidA* colonies as they are noticeably darker to the naked eye especially with adequate backlighting.

The use of CPSE plates also integrated a robust quality-control measure into this assay as any bacterial/fungal contamination was easily identified. To ensure that the Δ*uidA* mutation did not have any fitness impact, growth competition assays were performed with the original isolates and their Δ*uidA* counterparts (Section 2.10). The competitive index was calculated for each competition pairing using the formula below. Positive values indicated that Isolate A had a fitness advantage whereas negative values indicated that Isolate B had a fitness advantage. For example, if a competition index score of -5 was assigned to a growth competition pair of A vs B, it meant that B had a fitness advantage of five log (100000X) compared to A.

$$\textbf{Competitive index} = \log \frac{\text{Isolate A CFU}}{\text{Isolate B CFU}}$$

A breakpoint value of +1/-1 was used as the threshold of fitness difference. Any competition pairings where a competitive index was below this threshold were regarded as having no competition between the tested isolates. This implied that there was no significant difference in fitness between the tested isolates in that competition pairing. In cases where only one of the paired isolates proliferated, the maximum competition index score of +7/-7 was immediately assigned. There was no fitness impact associated with the Δ*uidA* mutation as the calculated competition index for every competing pair was lower than the breakpoint value of +1/-1 (Figure 52).

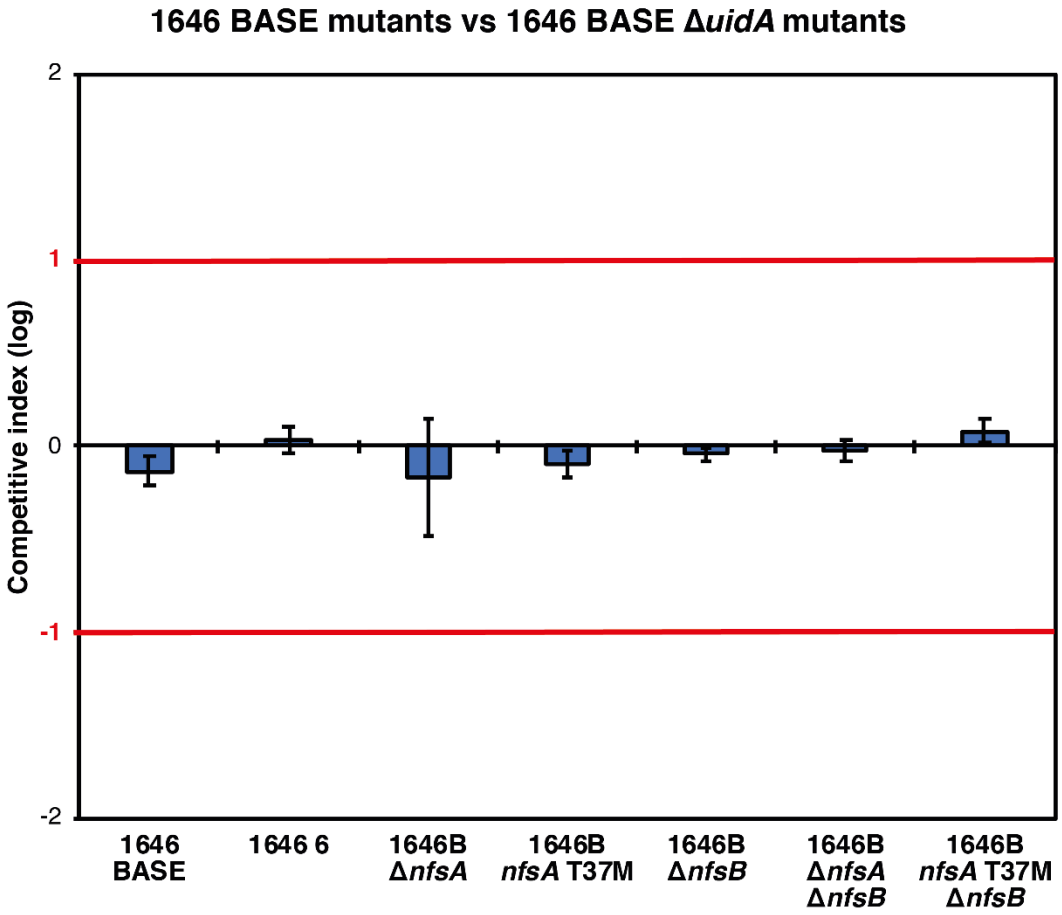



**Figure 52 Growth competition assay findings of *uidA* wild-type isolates vs Δ*uidA* mutants.** The competition index for each pairing was lower than +1/-1 (red line) indicating no significant competitive difference between the tested isolates.

### *4.5.1 Growth competition assay planning*

The Δ*uidA* mutants were generated from selected 1646 BASE *nfsA*/*nfsB* mutants. These mutants were chosen because they best represented potential evolutionary pathways in developing nitrofurantoin resistance (Figure 53). The Δ*nfsA*, *nfsA* T37M, and Δ*nfsB* mutants represented the three evolutionary pathways that a nitrofurantoin sensitive isolate could

undergo to become a first step (single) mutant. Likewise, the Δ*nfsA* Δ*nfsB* and *nfsA* T37M Δ*nfsB* mutants represented the evolutionary pathway that a viable first-step mutant (Δ*nfsA*, *nfsA* T37M) could undergo to become a second step (double) Nit$^{R}$ mutant (Figure 53).


**Figure 53 Evolutionary pathway of nitrofurantoin resistance based on literature findings.** Of the single mutants, only mutants with inactivated nfsA have ever been found among clinical and in-vitro spontaneous mutagenesis isolates. Therefore, the inactivation of nfsA is likely the only viable first-step mutation towards the acquisition of nitrofurantoin resistance. It is highly unlikely that nfsB mutants will ever be able to establish themselves within a population as they lack any substantial fitness/resistance advantage compared to their parental isolates, making nfsB mutations an evolutionary dead-end.

Growth competition assays were performed in two different settings: (i) antibiotic-free growth media to mimic an ideal environment where the tested isolates should achieve maximum growth and (ii) growth media with 8 μg/ml of nitrofurantoin. The latter was used to generate an antibiotic-induced stress that still allowed for the Nit$^{S}$ parental isolates to grow and compete with their **first step** and **second step** mutant counterparts, as suggested by the growth curve findings (Figure 49), but would highlight any fitness differences between these two groups.

The 1646 BASE isolate was paired against first and second step mutants to determine (i) whether the *nfsA*/*nfsB* mutations had any impact on fitness and (ii) whether the increased resistance of the mutants with inactivated *nfsA* translated to a significant fitness advantage at lower nitrofurantoin concentrations at which parental isolate and its mutants were able to grow. Mutants were also paired against each other to gauge their relative fitness under antibiotic-induced stress. The overall research question was to determine whether it is fitness or selection that drives the acquisition of both low, emergence of first step mutants, and high level, emergence of second step mutants, nitrofurantoin resistance.

### *4.5.2 Growth competition assay findings*

Data resulting from these experiments are shown in Figure 54A-E. At 0 μg/ml, there was no fitness difference between all tested isolates. At 8 μg/ml, only the single (Δ*nfsA*, *nfsA* T37M) and double mutants (Δ*nfsA* Δ*nfsB*, *nfsA* T37M Δ*nfsB*) with inactivated *nfsA* were

able to grow within their monoculture controls ( Figure 54F). Since 1646 BASE failed to grow at 8 μg/ml, there was no competition between it and its single/double mutants with inactivated *nfsA* ( Figure 54A,B,F). This pattern was also observed when the Δ*nfsB* single mutant was compared to the single mutants with inactivated *nfsA*, with the former failing to proliferate at 8 μg/ml ( Figure 54D,F). Positive selection of the single mutants with inactivated *nfsA* was therefore attributed to selective rather than fitness advantage at 8 μg/ml.

There was no fitness difference between the two single mutants with inactivated *nfsA* (Δ*nfsA* vs *nfsA* T37M) ( Figure 54D). There was also no fitness difference between the two double mutants with inactivated *nfsA* (Δ*nfsA* Δ*nfsB* vs *nfsA* T37M Δ*nfsB*) ( Figure 54E). These data suggested that *nfsA* inactivation was key and the manner of *nfsA* inactivation was irrelevant in terms of microbial fitness.

Single mutants with inactivated *nfsA* were also paired against their corresponding double mutants (i.e. Δ*nfsA* vs Δ*nfsA* Δ*nfsB*, *nfsA* T37M vs *nfsA* T37M Δ*nfsB*). However, it was impossible to isolate a single colony of the single mutant (Δ*nfsA* or *nfsA* T37M) from the growth competition cultures even at the lowest possible dilution (1:1000) where single colonies can be reliably discerned on CPSE plates. To address this limitation, data from the monoculture controls were examined. These data showed that double mutants had a 6-fold log fitness advantage over their corresponding single mutants at 8 μg/ml of nitrofurantoin ( Figure 54C,F).

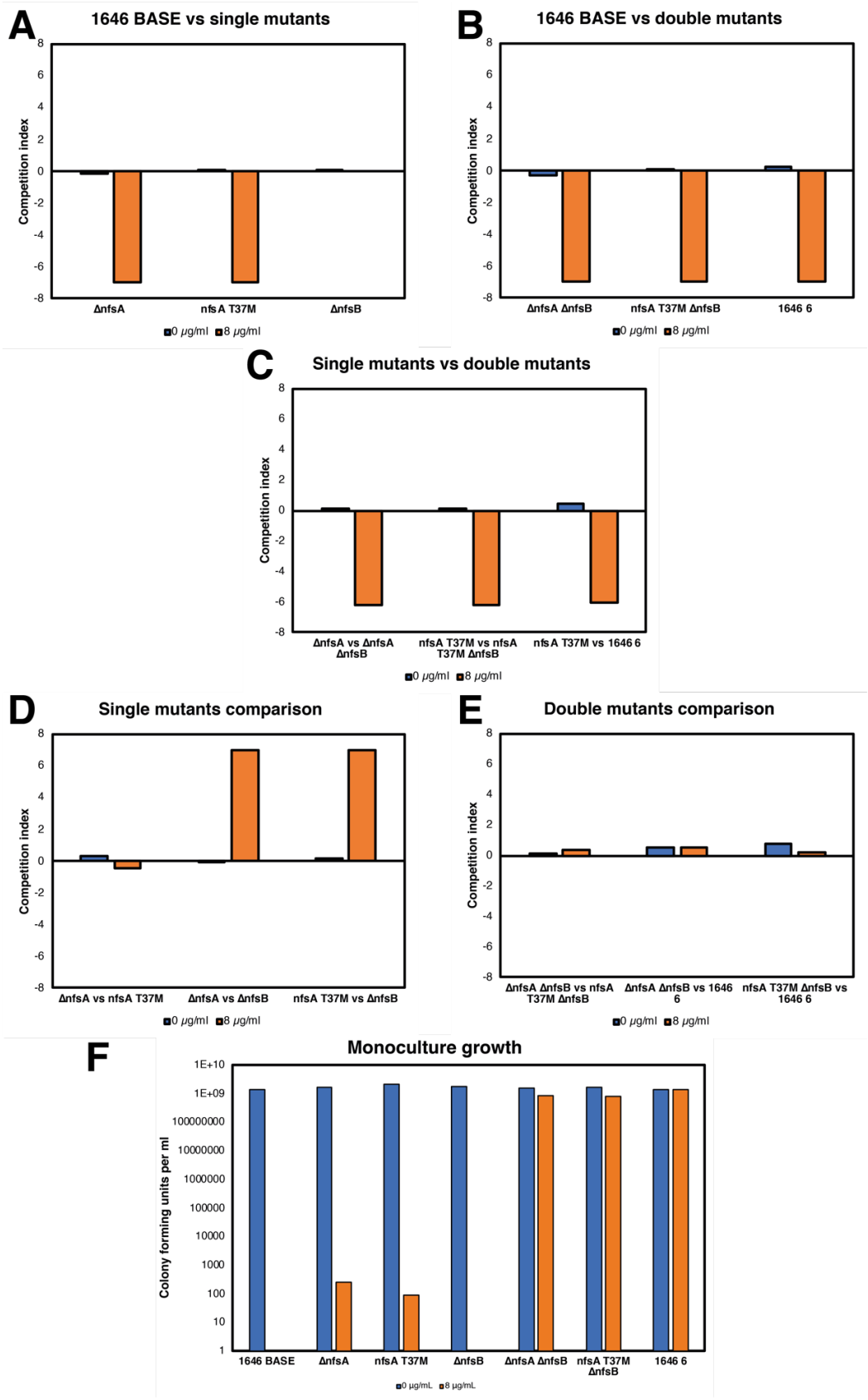



**Figure 54 Growth competition findings (A-E) and monoculture control findings (F).** For A and B, the x-axis shows the mutants that were competed against 1646 BASE, positive y-values indicate a growth advantage for 1646 BASE. For C,D,E, the x-axis shows the competition pairings, positive y-values indicate a growth advantage for the first of the pair. Monoculture control findings (Figure **55**) was used to determine whether an isolate was able to proliferate at 8 μg/ml and thus able to compete with its paired isolate.

Most mutations which confer antibiotic resistance affect important biological processes and are often associated with a substantial fitness cost in an antibiotic-free environment [359]. However, the growth competition data reported in this chapter provided evidence that mutations which inactivate *nfsA/nfsB* do not carry a fitness cost. Hence the low incidence of resistance among *E. coli* isolates cannot be attributed to any fitness difference between $Nit^R$ and $Nit^S$ isolates.

Data did however flag discrepancies between the outcomes of the MICs, growth curves, and growth competition assays in relation to the perceived minimum inhibitory concentrations for certain isolates (Table 34). For example, 1646 BASE was able to grow at 8 μg/ml of nitrofurantoin in the growth curve assay, but failed to do so in the MIC and the growth competition assays (Table 33; Figure 49A & Figure 52F). These discrepancies can be explained by a phenomenon known as the inoculum effect in which larger inoculum sizes increase the minimum inhibitory concentration of the bacterial culture[360]. Table 34 shows the inoculum size that was used for each of the assays mentioned in this chapter and the minimum inhibitory concentration of 1646 BASE inferred from the findings of each assay.

**Table 34 Assays, inoculum size, and MIC.**

| Assay | Inoculum size | MIC (1646 BASE) |
|---|---|---|
| MIC assay | ~$10^5$ cells in 200μl growth media, ~$10^5$-$10^6$ CFU/ml | 8 μg/ml |
| Growth curve assay | ~$10^6$ cells in 200μl growth media, ~$10^6$-$10^7$ CFU/ml | 16 μg/ml |
| Growth competition assay | ~25-50 cells in 5ml growth media, 5-10 CFU/ml | 8 μg/ml |

The inoculum size used for the growth curve assay was substantially larger compared to that used for the MIC and growth competition assays thereby explaining why the MIC of most tested isolates measured higher in the growth curve assay.

A concentration of 8 μg/ml of nitrofurantoin was chosen as the antibiotic concentration to induce bacterial stress within the growth competition assays because the growth curve data indicated that all of the 1646B isolates/mutants (including $Nit^S$ isolates) should be able to grow at this concentration (Figure 49). Furthermore, this was the concentration where there was the largest difference in growth rate between the $Nit^S$ isolates and single/double mutants with inactivated *nfsA* (Figure 49, Appendix A). However, the failure of 1646 BASE to grow at 8 μg/ml of nitrofurantoin in the monoculture controls ( Figure 54F) indicated that

there was no competition taking place between the $Nit^S$ isolates and the mutants with inactivated *nfsA* since competition can only take place between two viable isolates. Therefore, the data suggested 8 μg/ml of nitrofurantoin was too high an antibiotic concentration to compare the fitness of the mutants against their parental isolate. At this concentration, the mutants were able to grow only because they possessed a selective (increased resistance to nitrofurantoin) rather than a fitness advantage.

Although CPSE plates provided a straightforward way to differentiate the competing isolates, they were ineffective in competition pairings where one isolate vastly outcompeted the other. In such cases, data from the monoculture controls were used to estimate the population of the weaker isolate. Comparison of these data ( Figure 54F) against the competition data ( Figure 54A-E) revealed that the monoculture control data mirrored the competition data at both 0 and 8 μg/ml. Therefore, there was no need to perform additional growth competition assays as the monoculture control data by itself could be used to confidently predict whether a pairing might be competitive or not.

Isolate 1646 BASE and its mutants were therefore grown in monocultures following the growth competition assay protocols at 0, 2, 4, 8, and 16 μg/ml of nitrofurantoin (Figure 55). All isolates were able to grow at both 0 and 2 μg/ml of nitrofurantoin with no substantial difference in endpoint growth. At 4 μg/ml of nitrofurantoin, 1646 BASE failed to grow whereas mutants with inactivated *nfsA* (Δ*nfsA*, *nfsA* T37M, *nfsA* Δ*nfsB*, *nfsA* T37M Δ*nfsB*) were able to maintain normal growth. Surprisingly, the Δ*nfsB* single mutant which shared similar growth characteristics to 1646 BASE in previous experiments (Sections 4.3, 4.4) managed to grow, albeit poorly, at this higher nitrofurantoin concentration (compare 1646 BASE and 1646B Δ*nfsB* in Table 33, Figure 49A, Figure 55). At 16 μg/ml of nitrofurantoin, only double mutants with inactivated *nfsA* were able to grow, with similar endpoint growth as when they were growing at 0 μg/ml (Figure 55). Due to the limited range of concentrations (2 μg/ml of nitrofurantoin) where there might be actual competition between 1646 BASE and the *nfsA* mutants, the acquisition of nitrofurantoin resistance appears to be primarily driven by differences in resistance (selection-driven) between the mutants and their parental isolates rather than fitness difference.

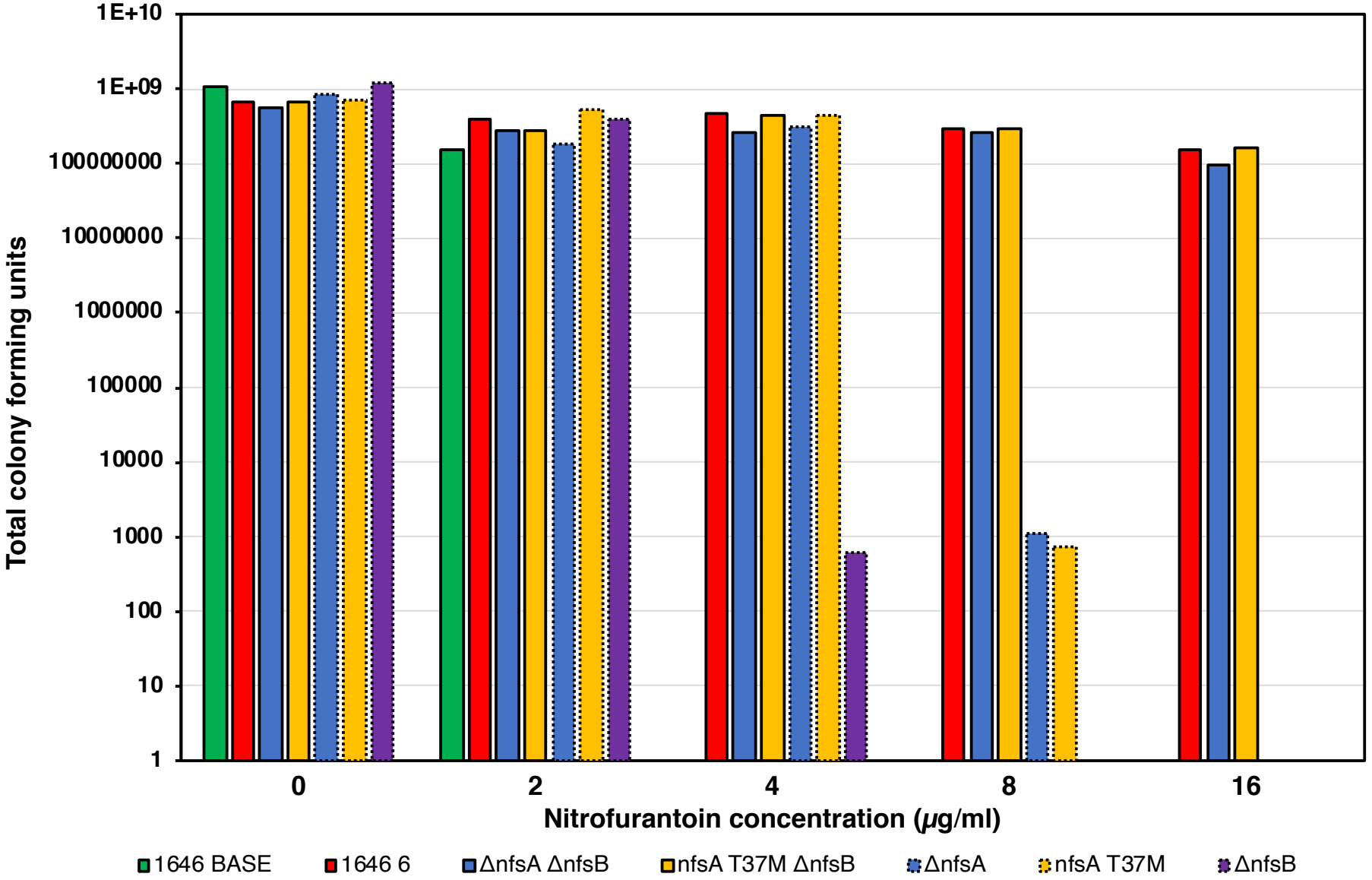



**Figure 55 Monoculture findings of 1646B, 1646 6, and 1646B *nfsA/nfsB* mutants at low nitrofurantoin concentration (0 -16 μg/ml).**

## 4.6 Discussion

Traditionally, bacterial resistance was assumed to be driven by high antibiotic concentrations (traditional selective window) that exceeded the MIC of sensitive isolates, allowing for resistant isolates to grow uncontested (Figure 56, highlighted red). Recently, multiple studies have successfully isolated resistant mutants from growth conditions where antibiotic concentrations were low enough for the proliferation of sensitive isolates [361–363]. For most antibiotics, bacterial isolates will still experience some degree of growth reduction even at concentrations just below their MIC even though their growth is not completely inhibited [364]. A resistant isolate will have a competitive advantage across an antibiotic concentration range (sub-MIC selective window) in which the growth reduction of the sensitive isolates is larger than that of resistant isolates (cost of resistance) (Figure 56, highlighted yellow).

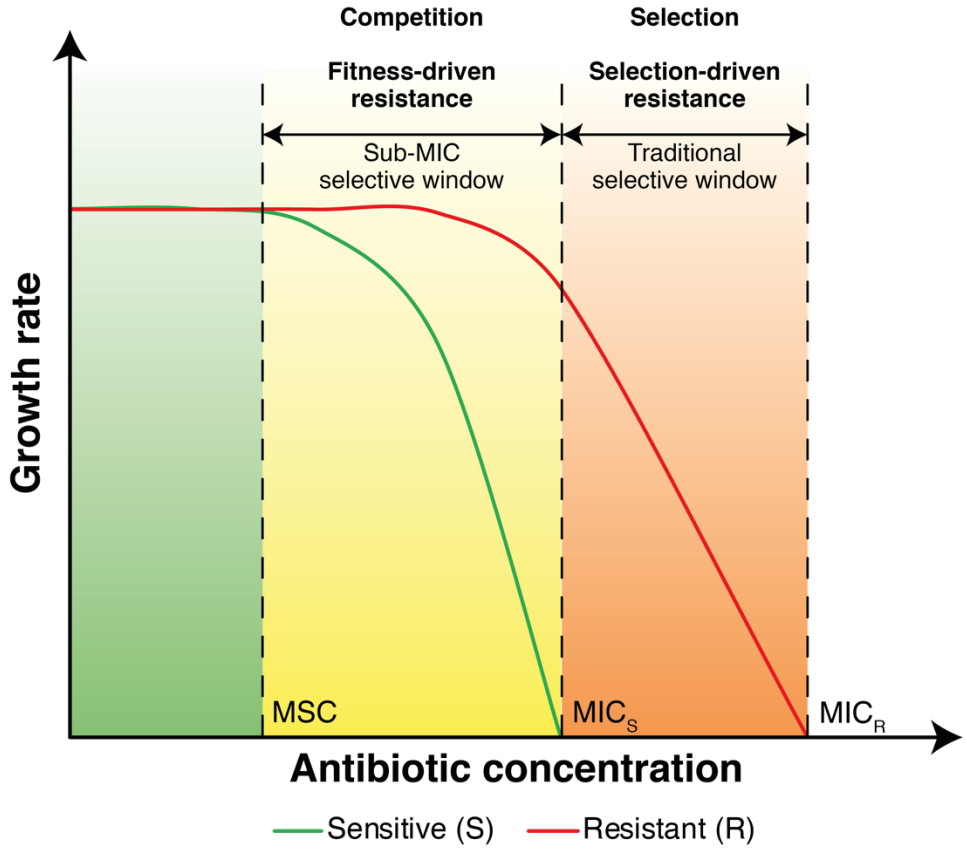



**Figure 56 How differences in growth rate between sensitive and resistant isolates can drive the acquisition of antibiotic resistance.** At the lowest antibiotic concentrations, there will be no significant difference in growth rate between the sensitive and resistant isolates. At higher antibiotic concentrations (below the MIC of the sensitive isolate), there will be a divergence in growth rate. The resistant isolates will have a fitness advantage over sensitive isolates, allowing the former to establish themselves in the population. At even higher antibiotic concentrations (above the MIC of the sensitive isolates but below the MIC of the resistant isolates), only resistant isolates will be able to grow as they possess a selective advantage over sensitive isolates at these high antibiotic concentrations. Adapted from Gullberg et al. (2011) [362]

The width of the traditional and sub-MIC selective windows may influence the incidence of antibiotic resistance through the ability of resistant mutants to successfully establish themselves within a population. One hypothesis is that the wider the selective windows, the greater the incidence of resistance. However, the width of these selective windows is not universal, but changes depending on the antibiotic-bacterial combination. Therefore, the pharmacokinetics & pharmacodynamics of an antibiotic, which influences its antimicrobial efficacy, and the genetic background of a bacteria, which determines its innate resistance, can influence the incidence of resistance within a bacterial population.

With regards to nitrofurantoin resistance, *in vitro* growth and competition assay findings reported in this thesis argue that the selective windows are wide, with selection-driven resistance playing a more dominant role compared to fitness-driven resistance (Figure 56). However, this argument is only applicable when comparing Nit$^R$ double mutants with

inactivated *nfsA* (Δ*nfsA* Δ*nfsB*, *nfsA* T37M Δ*nfsB*) against their Nit$^S$ parental isolates. In reality, this is an evolutionarily unfair comparison as these Nit$^R$ mutants cannot emerge directly from their Nit$^S$ parental isolates. High level nitrofurantoin resistance requires the inactivation of two genes, which are located so far apart from each other (287kbp distance) that the likelihood of both being simultaneously inactivated through a single natural genetic event is non-existent [273]. Furthermore, there are multiple essential genes between *nfsA* and *nfsB*, which prevents them from being inactivated in a single large deletion.

Therefore, the inactivation of these genes is much more likely to occur in a stepwise manner, starting with *nfsA* followed by *nfsB*. It can be speculated that the emergence of Nit$^R$ mutants is highly dependent on the ability of intermediate mutants with low level nitrofurantoin resistance to establish themselves within a population of sensitive isolates. In fact, when the growth of mutants (Δ*nfsA*, *nfsA* T37M) modelling these intermediate scenarios was compared to their parental isolate (1646 BASE), it was discovered that their selective windows were extremely narrow. Even at concentrations where they were strongly selected (>4 μg/ml), their low-level resistance did not provide enough protection to sufficiently offset the heavy antibiotic-induced growth reduction (Figure 49A). It can be argued that without sufficient growth, these mutants will still fail to establish themselves in the environment even though there is a strong selective pressure favouring their growth. More so in the urinary environment where frequent bladder voiding necessitates a minimum doubling time of 36 minutes to just maintain the presence of an isolate in that environment [271].

Furthermore, urinary nitrofurantoin concentration rarely falls to within this narrow concentration window, limiting the window of opportunity for the selection of these mutants (Figure 57A). Even when urinary nitrofurantoin concentration does fall to within this window, it still trends towards the higher end (~8 μg/ml) where there is strong growth reduction (Figure 57**)**. Therefore, the weak performance of the intermediate mutants (Δ*nfsA*, *nfsA* T37M) under nitrofurantoin-induced stress and the unique pharmacokinetics of nitrofurantoin, arguably, prevents these mutants from establishing themselves in a sufficiently successful manner, undermining the widespread emergence of double mutants. Intermediate mutants are the weak link in the chain which undermines the emergence of nitrofurantoin resistance in *E. coli*, thereby explaining the low incidence of resistance.

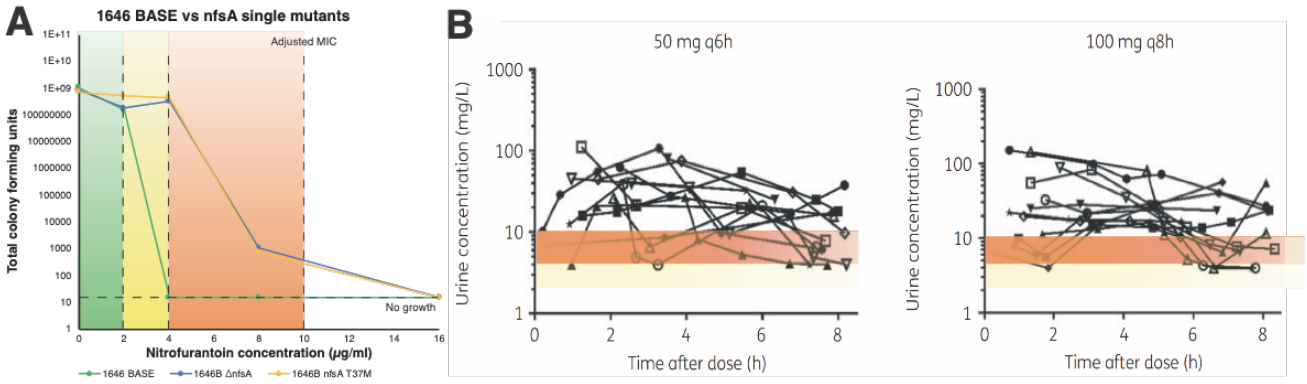



**Figure 57 Concentration windows where single mutants have a fitness (highlighted yellow) or selective advantage (highlighted) over their parental isolate , 1646 BASE (A) Correlation of concentration windows with nitrofurantoin pharmacokinetics data (B).** Due to the poor growth of the *nfsA* single mutants at 8 μg/ml, their MIC has been adjusted to 10 μg/ml. (B) was adapted from Huttner et al. (2019) [267]

# Chapter 5 The role of *Proteus mirabilis* in the host-microbial interactions of UTI

## 5.1 Introduction to the BUTI dataset

The Biomarkers in Urinary Tract Infection (BUTI) was a clinical study, which investigated microbiological and immunological components associated with uncomplicated recurrent UTIs in a cohort of 30 community dwelling older patients aged ≥ 65 years [142]. The study also included a control cohort consisting of 15 healthy older age matched individuals with no prior history of UTI in the previous three years or longer. Each patient in the rUTI cohort submitted a mid-stream urine sample every fortnight for a period of 6 months (Figure 58). Those in the control cohort submitted three urine samples throughout the study, spaced two weeks apart.

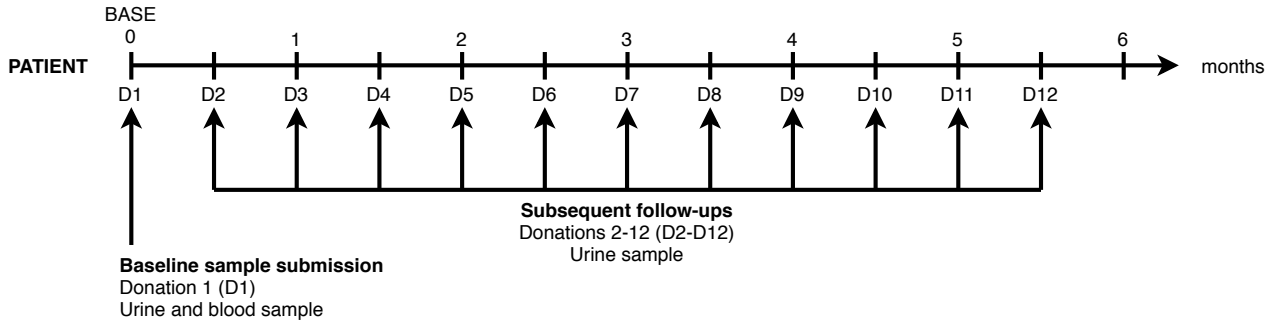



**Figure 58 BUTI study timeline.** Patients in the rUTI (target) group submitted a urine and blood sample at the beginning of the study and a urine sample every 2 weeks thereafter over the course of around 6 months.

The BUTI dataset consisted of three main components (Figure 59), relating to (i) urine samples, (ii) clinical data, and (iii) blood samples. Unfiltered urine samples were collected during this study and analysed using Siemens Multistix 10SG Urinalysis strips (dipstick analysis) before being plated on chromogenic agar plate (bioMérieux CPS3 or CPSE plates) to screen for uropathogens (Figure 59A). Of the uropathogens, only *E. coli, P. mirabilis and E. faecalis* were quantified where present (Figure 59A). Full plate washes from these chromogenic plates were stored at -80°C to allow for future microbiological analyses. *E. coli* colonies were isolated and subjected to additional analyses: (i) 7-gene Achtman MLST, (ii) motility assays, and (iii) antimicrobial resistance (AMR) assays, with the aim of studying their longitudinal colonisation pattern (Figure 59A). Filtered urine samples were also analysed for urinary creatinine and cytokines concentrations (Figure 59A).

Clinical data from the 30 rUTI patients provided information regarding their existing medical conditions, antibiotic prescriptions, and self-reported/GP-confirmed UTI symptoms during the study (Figure 59B). Blood samples were subjected to serum and TLR genotype analysis (Figure 59C). The TLR genotype analysis screened for four common TLR SNPs which were known to affect the immune response of the urinary tract: TLR1 T1805G, TLR2 G2258A, TLR4 A896G, TLR5 C1174T respectively.

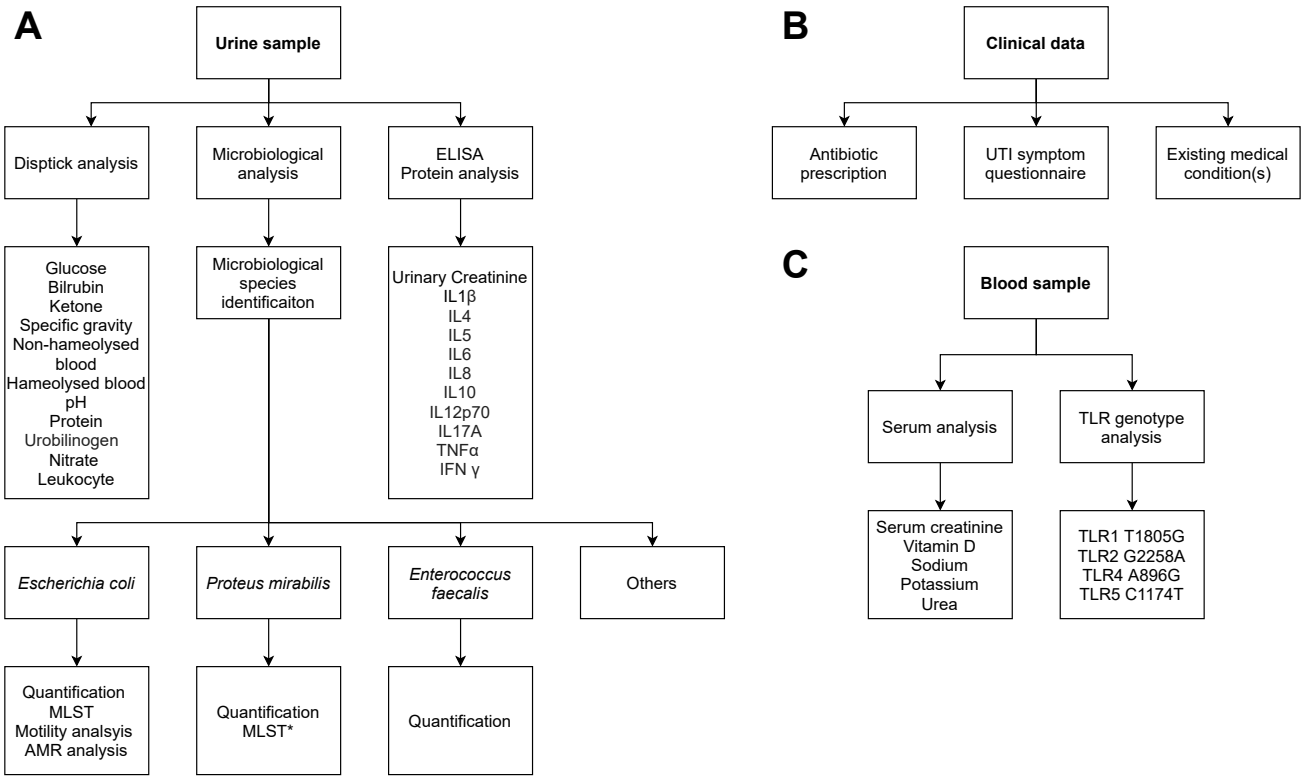



**Figure 59 Overview of the BUTI dataset organised by data source.** (A) Urine sample data (B) Clinical data (C) Blood sample data.

## 5.2 Revisiting the BUTI dataset

*E. coli* colonisation was investigated in the original BUTI study due to the status of *E. coli* as the dominant causative organism in UTIs. However, no research was undertaken on *P. mirabilis*. *P. mirabilis* is not as prevalent as *E. coli* in uncomplicated UTIs (Figure 9) [209], in fact it is quite rare. Hence its isolation from the urines of these older patients was interesting and justified further investigation.

### 5.2.1 Microbiological overview of BUTI

Identification and quantification of the three major bacterial groups (*E. coli, P. mirabilis, E. faecalis*) were performed on chromogenic agar plates. The main limitation of these plates is that they lack the required specificity to identify *E. faecalis* isolates. Therefore, no data

relevant to *E. faecalis* was collected. Microbiological analysis of the urine samples (n=360) revealed that *E. coli* was found in 51.1% of all urine samples compared to *P. mirabilis* in 18.6% of samples (Figure 60). *E. coli* was also more likely to exhibit clinical levels of colonisation (≥ $10^5$ CFU/ml) compared to *P. mirabilis* (56.5% vs 25.4%) (Figure 60).

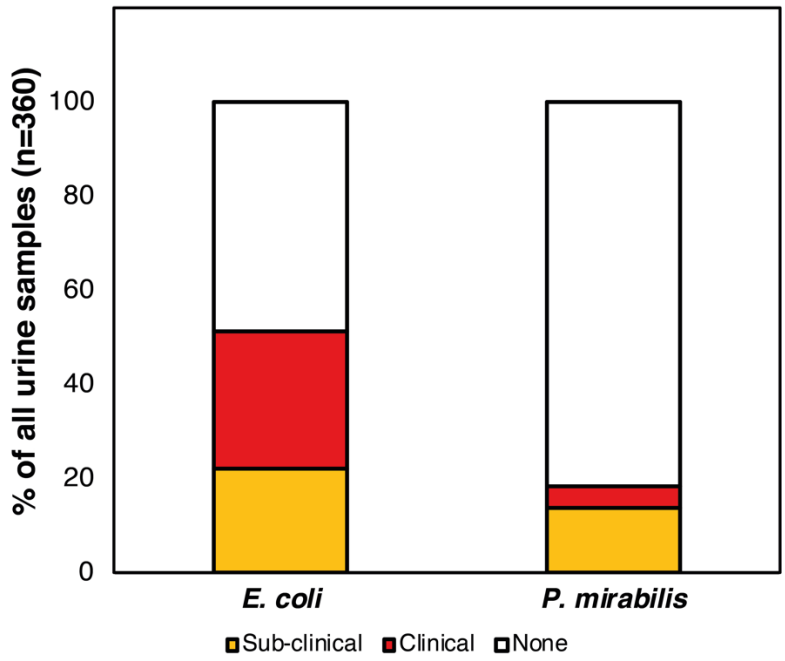



**Figure 60 Incidence of *E. coli* and *P. mirabilis* across all 360 BUTI urine samples.** None: no colonisation, Sub-clinical: < $10^5$ CFU/ml, Clinical: ≥ $10^5$ CFU/ml

### *5.2.2 Immunological background of BUTI participants*

As discussed in the introduction (Section 1.2.7), TLR SNPs can influence the innate immune response of the urinary tract by increasing/reducing an individual's susceptibility to UTI and/or bacterial colonisation [151]. To date it has been reported through genetic studies that TLR1 1805T is associated with protection from pyelonephritis [365]; TLR2 2258A is associated with increased susceptibility to asymptomatic bacteriuria [151]; TLR4 896G is associated with an increased risk of UTI [365]; and TLR5 1174T is associated with an increased risk of rUTI [365] (Section 1.2.7). Unsurprisingly, SNPs that do not impact TLR functioning are usually protective whereas SNPs that impair TLR functioning are detrimental (Table 6) [151,366,367].

**Table 35 Number of patients with specific TLR genotypes for the 4 screened TLR SNPs.**

| TLR SNP | AA Homozygous dominant | Aa Heterozygous complete dominant | aa Homozygous recessive | AB Heterozygous codominant | BB Homozygous dominant |
|---|---|---|---|---|---|
| TLR1 G1805T | **12** | N/A | N/A | **16** | **2** |
| TLR2 G2258A | **28** | **2** | **0** | N/A | N/A |
| TLR4 A896G | **1** | **4** | **25** | N/A | N/A |
| TLR5 C1174T | **0** | **4** | **26** | N/A | N/A |

For each TLR SNP (TLR1 G1805T, TLR2 G2258A, TLR4 A896G, TLR5 C1174T, the colour of the alleles (A,G,C,T) corresponds to the whether the allele is dominant (red font), recessive (green font), codominant or (blue font). Of the 4 TLR SNPs, only TLR1 G1805T has two codominant alleles whereas the rest of the TLR SNPs have a dominant and a recessive allele.

The 30 BUTI patients were analysed for TLR SNPs, data suggested that there were 12 patients with potentially non-functional TLR1 (TLR1 1805GG genotype) and 18 patients with potentially functional TLR1 (TLR1 1805GT/TT genotype) (Table 6, row highlighted green in Table 35). The high frequency of this SNP in the study population (18/30 patients possessed the T allele) and the fact that TLR1 is expressed in the urothelium [368] prompted an interest in exploring whether TLR1 played a role in the pathogenesis of rUTIs.

### *5.2.3 Does TLR genotype influence bacterial colonisation?*

The BUTI rUTI patient microbiological data was initially segregated based on the two patient groups, **TLR1 1805G** and **TLR1 1805T** to represent patients with potentially non-functional and functional receptors respectively, and to investigate whether the TLR1 genotype was associated with any significant changes in bacterial colonisation.

Contrary to a previous study [151], bacterial colonisation amongst TLR1 1805G patients was significantly lower compared to TLR1 1805T patients (84.7% vs 95.6%, $p<0.001$) (Figure 61). This difference could be attributed to the significantly lower incidence of *P. mirabilis* colonisation among TLR1 1805G patients compared to TLR1 1805T patients (0.7% vs 23.5%, $p<0.001$) (Figure 61). Interestingly, there was no significant difference in the incidence of *E. coli* colonisation between these two groups. *P. mirabilis* colonisation appeared to be almost exclusively limited to TLR1 1805T patients as there were only four

urine samples that were positive for *P. mirabilis* from TLR1 1805G patients compared to 63 positive samples from TLR1 1805T patients.

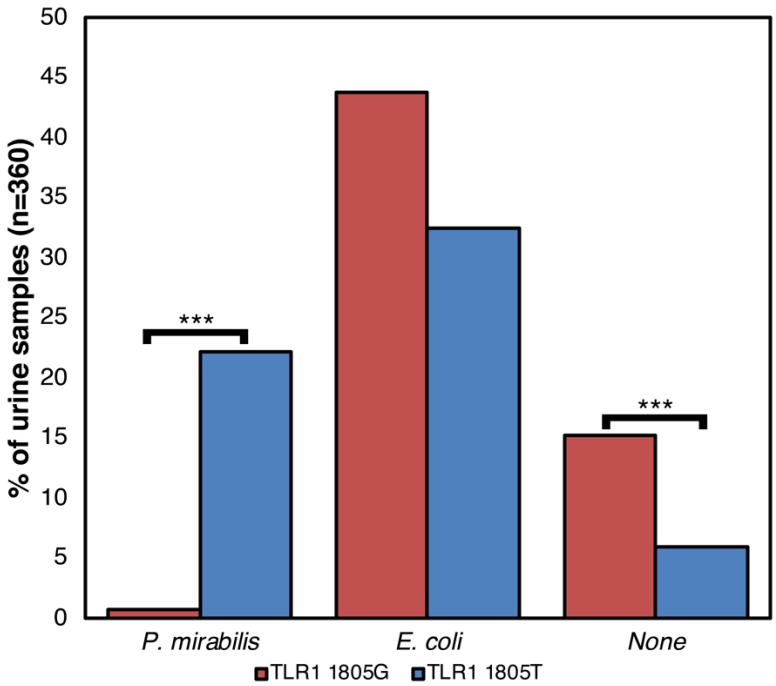



**Figure 61 Breakdown of bacterial colonisation of TLR1 1805G vs 1805T patients.** *** p<0.001.

Urine samples from the TLR1 1805G patients were compared to samples from TLR1 1805T patients. The concentration of the tested cytokines was normalised against urinary creatinine concentration to account for variations in renal function between the patients. TLR1 1805G patients were found to have significantly higher levels of IL4 (0.92 vs 0.52 pg/mg, p=0.014), TNF$\alpha$ (14.5 vs 6.3 pg/mg, p=0.001), and IFN$\gamma$ (29.6 vs 15.6 pg/mg, p=0.009) compared to TLR1 1805T patients.

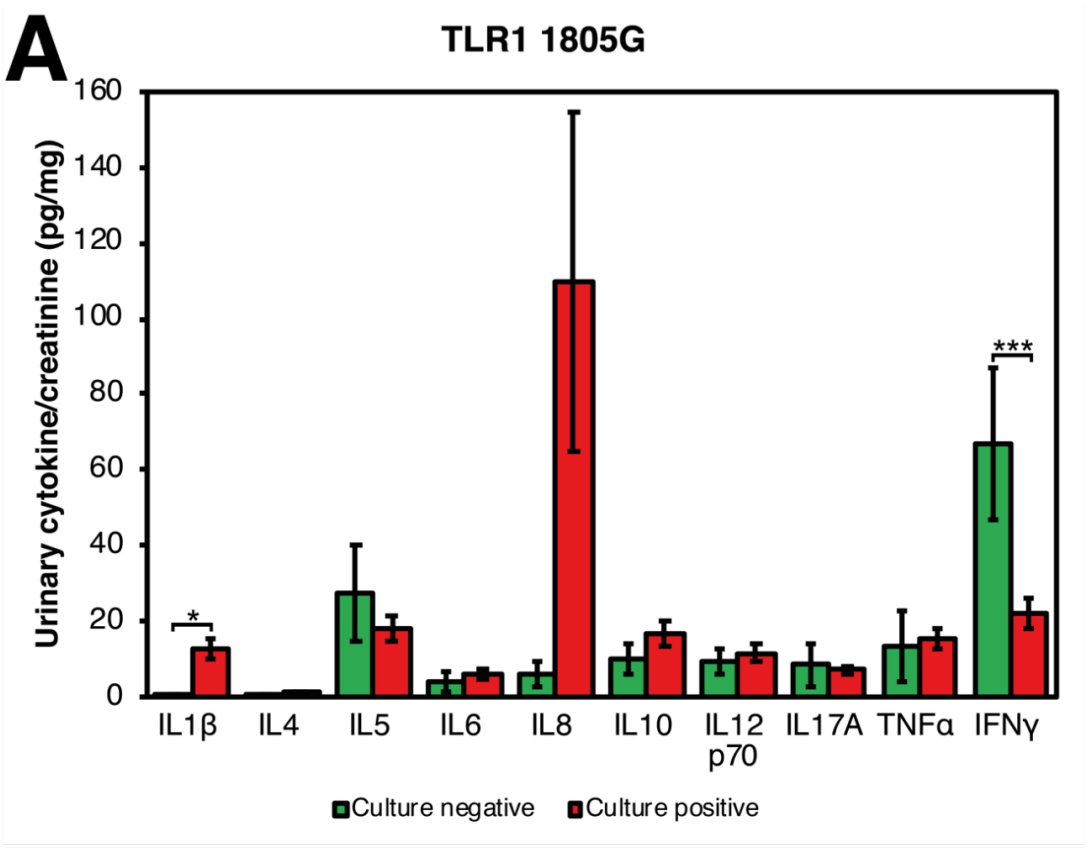
### *5.2.4 Does TLR genotype influence the immune response?*

To explore whether TLR1 function affected the cytokine response to bacterial colonisation, urine cytokine measurements associated with TLR1 1805G patients were compared to those of TLR1 1805T patients. Within each TLR patient group, culture positive urine samples were compared to culture negative urine samples. The concentrations of the tested cytokines were normalised against urinary creatinine concentration to account for interindividual variations in renal function.

For TLR1 1805G patients, culture positive samples had significantly elevated levels of IL1$\beta$ (10.8 vs 0.52 pg/mg, p=0.013) and reduced levels of IFN$\gamma$ (22.3 vs 67.0 pg/mg, p=0.0005)

(Figure 62A). Although there was a substantial difference in IL8 secretion between culture negative and positive samples from TLR1 1805G patients, this was not statistically significant (5.9 vs 109.5 pg/mg , p=0.29) (Figure 62A). However, the low number of culture negative samples from TLR1 1805G patients (n=22) and the large variation in IL8 levels amongst culture positive samples (standard deviation: 459.2) could have decreased statistical power to an extent that rendered the difference in IL8 levels between these two sets of samples to become insignificant. For TLR1 1805T patients, there were no significant differences in cytokine concentrations between culture positive and culture negative samples (Figure 62B).

.



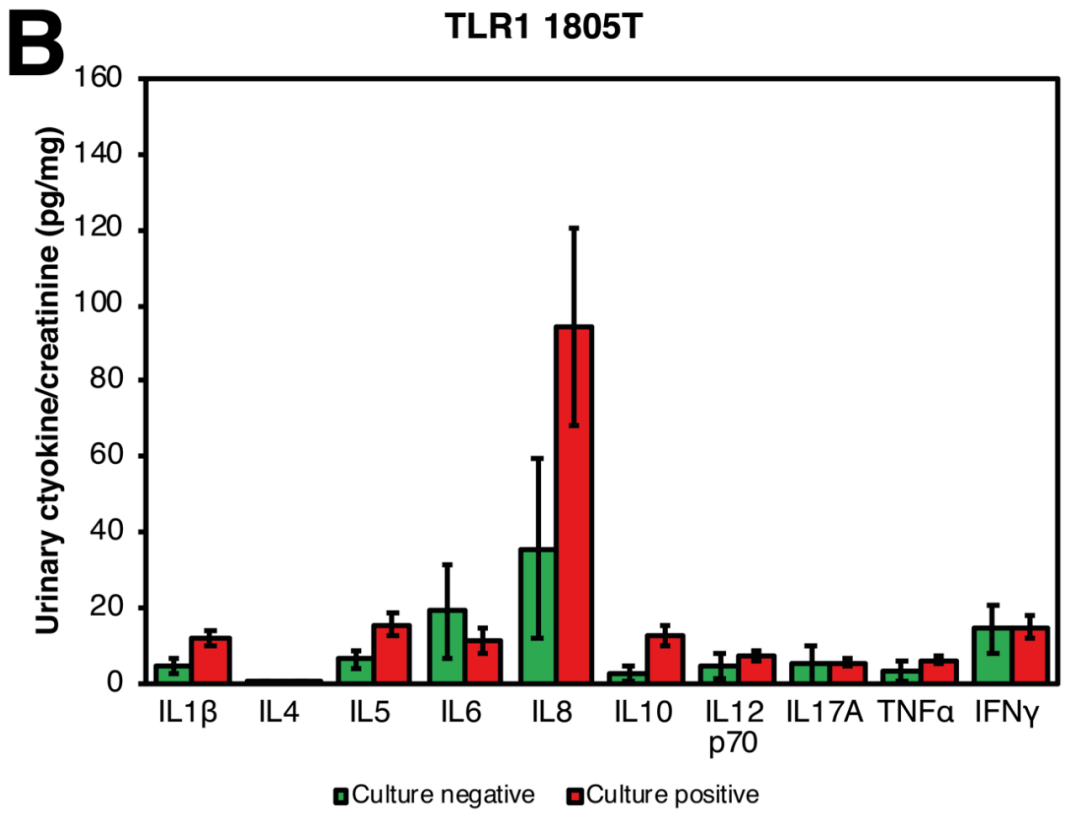




**Figure 62 Average urinary cytokine levels of culture negative (green bars) vs culture positive (red bars) samples from TLR1 1805G (A) and TLR1 1805T (B) patients.** There was a significant change in IL1β and IFNγ levels for TLR1 1805G patients when the urinary tract was colonised (culture positive). However, the cytokine levels for TLR1 1805T patients remained unchanged. TNFα, Tumour Necrosis Factor alpha; IFNγ, Interferon gamma. * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

### 5.2.5 Summary

Analyses of BUTI patient data suggested a strong association between *P. mirabilis* colonisation and TLR1 1805T rUTI patients, with 63/67 (94%) of *P. mirabilis* isolates originating from this patient group. Urinary cytokine data suggested that the innate immune response to bacterial colonisation was not as robust among TLR1 1805T patients compared to TLR1 1805G patients (Figure 62).

The decrease in pro-inflammatory IFNγ amongst culture positive samples from TLR1 1805T patients was however surprising as bacterial colonisation is usually associated with an increase in the secretion of pro-inflammatory cytokines. However, IFNγ promotes the secretion of indoleamine 2,3-dioxygenase (IDO) from urothelial cells and polymorphonuclear leukocytes which can interfere with the clearance of extracellular bacteria due to its immunosuppressive effects. Taken together, the increased incidence of *P. mirabilis* colonisation among TLR1 1805T patients may have resulted from the lack of an effective innate immune response to *P. mirabilis* via TLR1, thereby preventing bacterial clearance

## 5.3 Investigating the immunological mechanism underlying P. mirabilis colonisation in TLR1 1805T patients

To further investigate the immunological mechanisms, *in-vitro* experiments were performed using the RT4 immortalised cell line to model the urothelial response against *E. coli* and *P. mirabilis*. The RT4 cell line was profiled for TLR SNPs and expression of these TLRs was confirmed by RT-PCR (Table 36, Figure 63). Moreover, the TLR profile of this RT4 cell line matches that of a TLR1 1805G patient (Table 36).

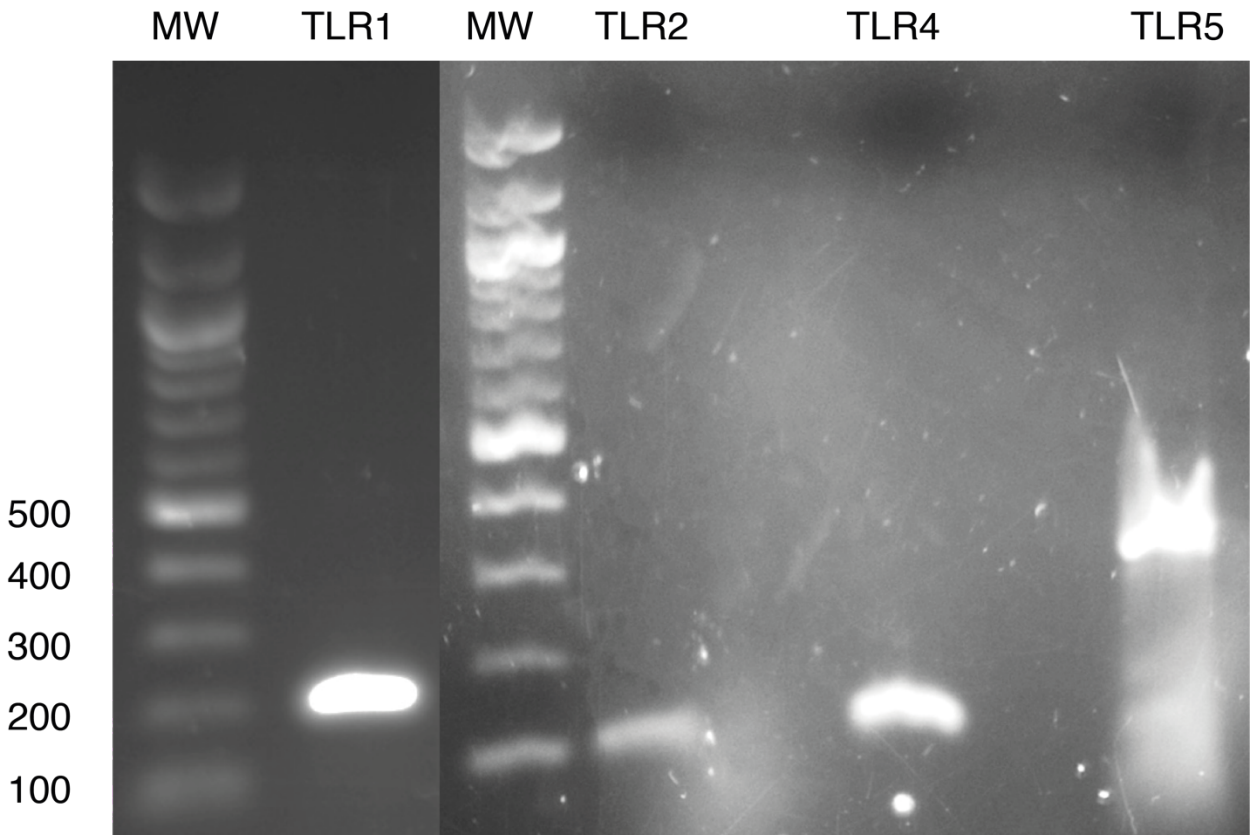



**Figure 63 Gel electrophoresis picture of TLR 1,2,4,5 messenger RNA (mRNA) reverse transcriptase PCR products to assess TLR expression in RT4 cells.** The presence of bands confirmed the expression of TLR 1,2,4,5. Molecular weight (MW): New England Biolabs (NEB) 100 bp DNA ladder. The original agarose gel pictures are shown in Appendix D and Appendix E.

**Table 36 TLR profile and mRNA expression of RT4 cells**

| TLR | RT4 genotype | RT4 mRNA expression |
|---|---|---|
| TLR1 | TLR1 1805GG | Yes |
| TLR2 | TLR2 2258GG | Yes |
| TLR4 | TLR4 896AA | Yes |
| TLR5 | TLR5 1174CC | Yes |

### *5.3.1 Cytokine profiling of RT4 cells against flagellin, a known agonist.*

To confirm the RT4 cells were responsive to PAMPs, they were challenged with flagellin, a known immunostimulatory agent associated with uropathogens (50 ng/ml flagellin) [158], and the cytokines IL1β, IL5, IL6, IL8, IL10 measured. These cytokines were selected to reflect the urothelial response based on existing literary evidence [116,369]. However, of those analysed, only IL8 displayed a significant change (increase) in concentration following exposure to flagellin (Figure 64). This finding modelled previous *in vitro* and clinical data in relation to UPEC [370]. Hence IL8 was chosen as a quantifiable indicator of the urothelial response.

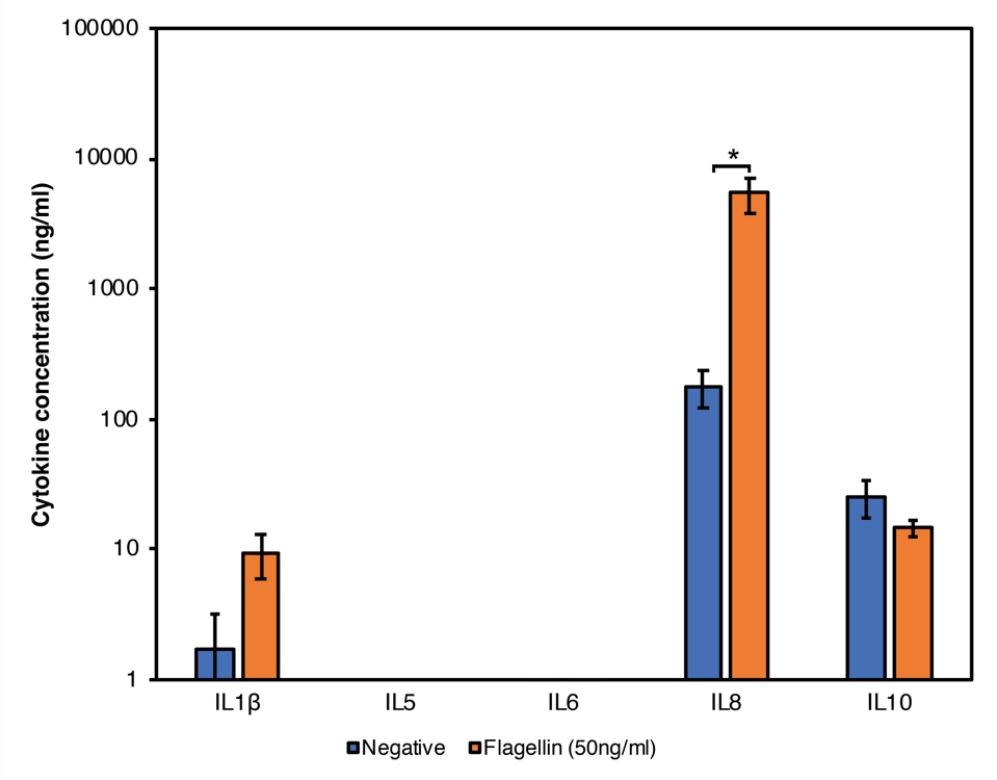



**Figure 64 Concentration of IL1β, IL5, IL6, IL8, IL10 following TLR agonist challenge with negative control (blue bars) and 50ng/ml flagellin (orange bars).** Negative control is sterile phosphate buffered saline (PBS). n=6 biological repeats

### *5.3.2 Uroepithelial response against heat-killed bacteria*

To explore the urothelial response to uropathogenic bacteria, RT4 cells were challenged with heat-killed *E. coli* (TPA 2743, NCTC 10418) and *P. mirabilis* (UTI100) using the following inocula: $10^1$, $10^2$, $10^3$, $10^4$, $10^5$, and $10^6$ CFU/ml. The *E. coli* isolate is a laboratory model whereas the *P. mirabilis* isolate (UTI 100) was a clinical isolate from a BUTI patient who experienced clinical levels of persistent *P. mirabilis* colonisation during the study [142]. Interleukin 8 (IL8) ELISA results are shown in Figure 65. These responses to *E. coli* and *P. mirabilis* showed a dose-dependent increase, peaking at an inoculum of $10^5$ CFU/ml for *E. coli* (10.1 µg/ml) and $10^6$ CFU/ml for *P. mirabilis* (7.4 µg/ml). At inoculum doses of $10^3$, $10^4$, $10^5$ CFU/ml respectively, the IL8 response to *E. coli* was significantly higher than that of *P. mirabilis* (Figure 65). However, at $10^6$ CFU/ml challenge, the IL8 response was comparable

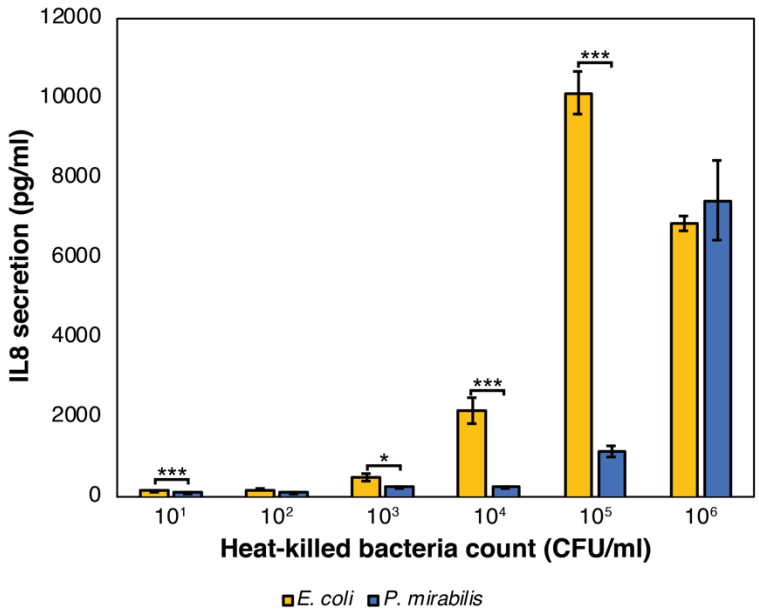



**Figure 65 IL8 secretion following challenge with heat-killed preparations of *E. coli* (yellow) and *P. mirabilis* (blue).** * p<0.05, **, p<0.01, *** p<0.001. 1 experimental repeat with 6 biological repeats at each concentraiton for each sample.

The urothelial response is mediated by TLRs which recognise different bacterial components [98]. In relation to motility, it is known that this response is predominated by TLR5 signalling which is driven by receptor recognition of flagellin [158,371]. *E. coli* and *P. mirabilis* are both motile bacteria characterised by flagella constructed from constituent flagellin proteins that share the same highly conserved amino acid regions, which are recognised by TLR5 [201,372]. It was therefore reasonable to hypothesise that the observed differences in IL8 secretion between heat-killed *E. coli* and *P. mirabilis* were potentially driven by differences in the TLR5 response [207,371,373].

### *5.3.3 Uroepithelial response against purified flagellin preparations*

To explore this further, RT4 cells were challenged with purified flagellin (Section 2.6.1) prepared from a motile *E. coli* isolate (TPA 3408) and *P. mirabilis* isolate UTI 100. The purified flagellin preparations are shown in Figure 66. The following concentrations of purified flagellin were used in the challenges: 0.025, 0.25, 2.5, 25 and 50 ng/ml respectively.

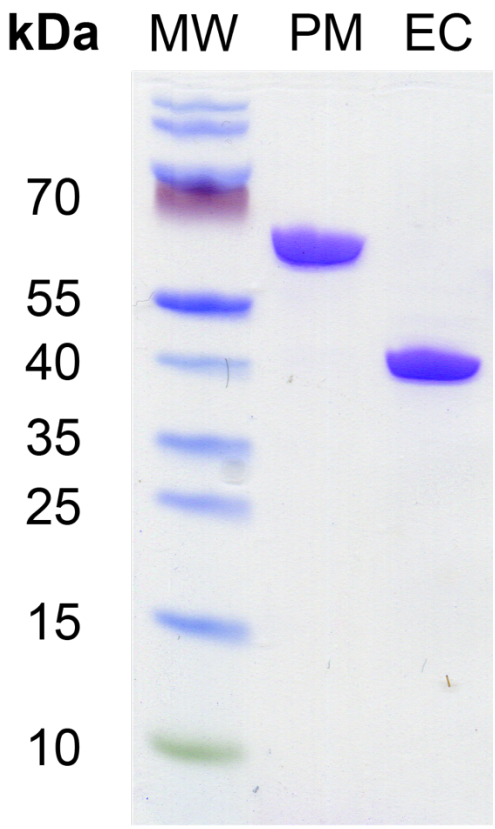



**Figure 66 SDS-PAGE gel showing purified flagellin preparations from *Proteus mirabilis* UTI 100 (PM) and *Escherichia coli* TPA 3408 (EC).** The two gel lanes with the *P. mirabilis* flagellin preparations were purified from the same isolate. Molecular weight (MW): Sigma-Aldrich mPAGE Color Protein Standard. An unaltered, annotated gel picture can be found in Appendix F.

As expected, both *E. coli* and *P. mirabilis* flagellin were able to induce a urothelial response (Figure 67). However, while the peak IL8 response to *P. mirabilis* flagellin was observed at 2.5 and 25 ng/ml respectively, the peak response to *E. coli* flagellin occurred at 25 ng/ml (Figure 67). Although the IL8 response against *P. mirabilis* flagellin was significantly higher compared to *E. coli* at 2.5 ng/ml (13.7 ng/ml vs 8.2 ng/ml, p=0.00003), this was reversed at the higher flagellin concentrations of 25 ng/ml (13.4 ng/ml vs 23.4 ng/ml, p=0.00002) and 50 ng/ml (9.5 ng/ml vs 17.5 ng/ml, p=0.00005) (Figure 67) respectively (Figure 67). These data suggested that the RT4 urothelial response to *P. mirabilis* flagellin plateaued at a concentration of 2.5 ng/ml and did not change despite a 10-fold increase in flagellin concentration.

However, the difference in the upper limit of the urothelial response towards *E. coli* and *P. mirabilis* flagellin was not as substantial when compared to the heat-killed preparations (compare 25 ng/ml findings from Figure 67 against $10^5$ findings from Figure 65). This suggests that there might be other bacterial components besides flagellin that could influence the urothelial IL8 response.

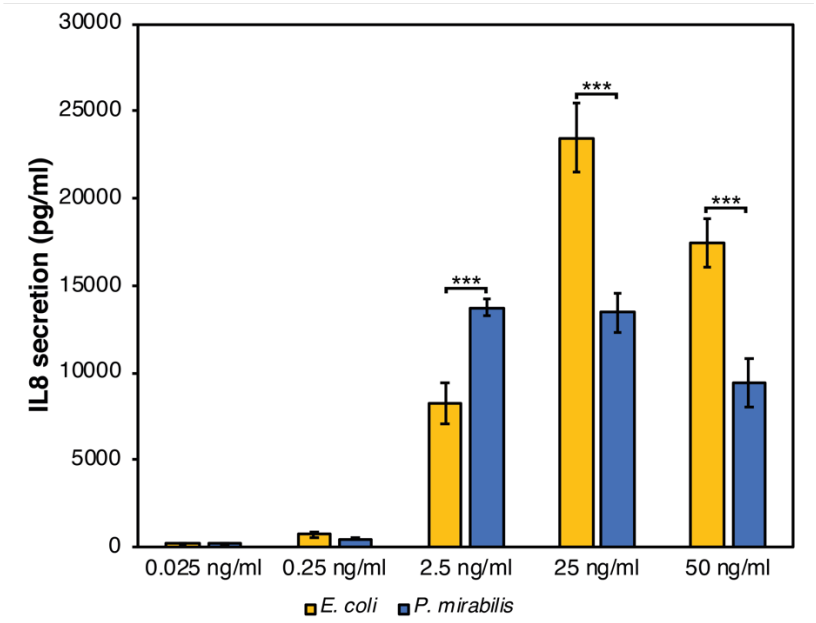



**Figure 67 IL8 secretion after exposure to *E. coli* (yellow) and *P. mirabilis* (blue) flagellin preparations.** * p<0.05, **, p<0.01, *** p<0.001. 2 experimental repeats with 6 biological repeats at each concentration for each sample.

### *5.3.4 Uroepithelial response against lipopeptides (TLR1/2/6 agonists)*

Data shown in Table 35, Section 5.2.2 indicated that 26/30 BUTI patients carried a functional TLR5 receptor while the remaining four patients carried the SNP which supported non-functional TLR5. Hence most of the BUTI patients would have been able to detect and presumably respond to bacterial flagellin. However, 18 of the BUTI patients each carried a functional TLR1 receptor (TLR1 1805T), which was associated with a higher incidence of *P. mirabilis* colonisation (23.5% vs 0.7%, p<0.001, Figure 60) compared to patients with non-functional TLR1 receptors (TLR1 1805G). All 30 BUTI patients each had functional TLR2 receptor, this is an important consideration as TLR1 forms a heterodimer with TLR2 during PAMP recognition (Section 1.2.3).

*In vitro* challenge experiments were therefore expanded to explore the contribution, if any, of these receptors in the global urothelial response to uropathogens. These challenges involved using Pam2CSK6 (Pam2) and Pam3CSK4 (Pam3). Pam2 and Pam3 are synthetic analogues of diacyl and triacyl lipopeptides which are expressed on the plasma membrane of Gram +ve and Gram -ve bacteria respectively [374]. Therefore, Pam2 should stimulate TLR2-6 signalling whereas Pam3 should stimulate TLR1-2 signalling (Table 4). However, challenging RT4 cells with these synthetic lipopeptides (100 ng/ml) did not generate as potent a urothelial IL8 response as flagellin (Figure 68), but the response to Pam2 was significantly greater than Pam3. This result was expected as RT4 cells are genotypically

TLR1 1805GG and TLR2 2258GG, i.e. are not predicted to have functional TLR1, but they do have functional TLR2 (Table 36).

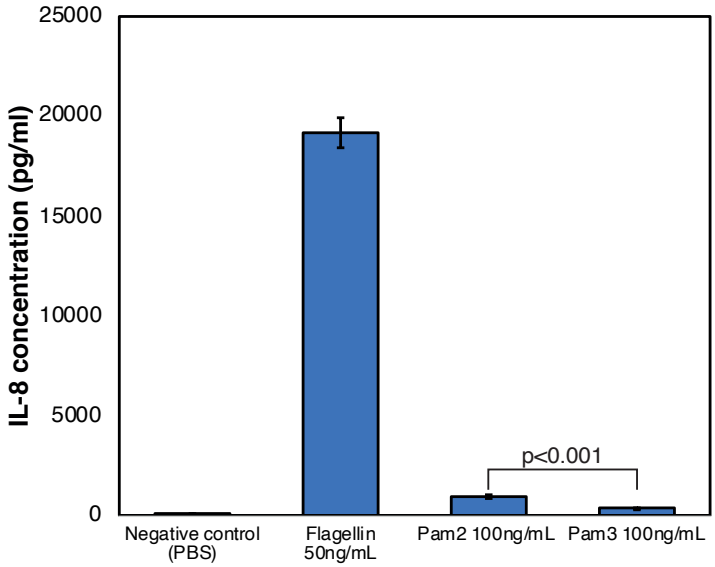



**Figure 68 IL8 response against flagellin, Pam2 and Pam3.** 3 experimental repeats with 6 biological repeats for each sample.

Since TLR signalling does not occur in isolation [375], co-challenges were performed in which RT4 cells were simultaneously challenged with both *E. coli* flagellin and either Pam2 or Pam3, and IL8 concentrations measured. A flagellin only control was included as the baseline reference to determine the presence of any synergistic or antagonistic effects (Figure 69, red bars).

Pam2 and flagellin co-challenges (Figure 69, green bars) elevated the urothelial IL8 response at 0.025 (1600 pg/ml vs 450 pg/ml, $p<0.001$) and 0.25 ng/ml (1400 pg/ml vs 850 pg/ml, $p<0.01$) of flagellin respectively (green vs red bars in Figure 69). At higher flagellin concentrations of 1, 2.5, 10, 25, 50 ng/ml respectively, the combination of Pam2 and flagellin did not significantly alter the urothelial IL8 response (Figure 69). The Pam3 and flagellin combination was however, associated with a significant decrease in urothelial IL8 response at 25 ng/ml of flagellin (19000 pg/ml vs 23400 pg/ml, $p=0.01$) (blue vs red bars in Figure 69).

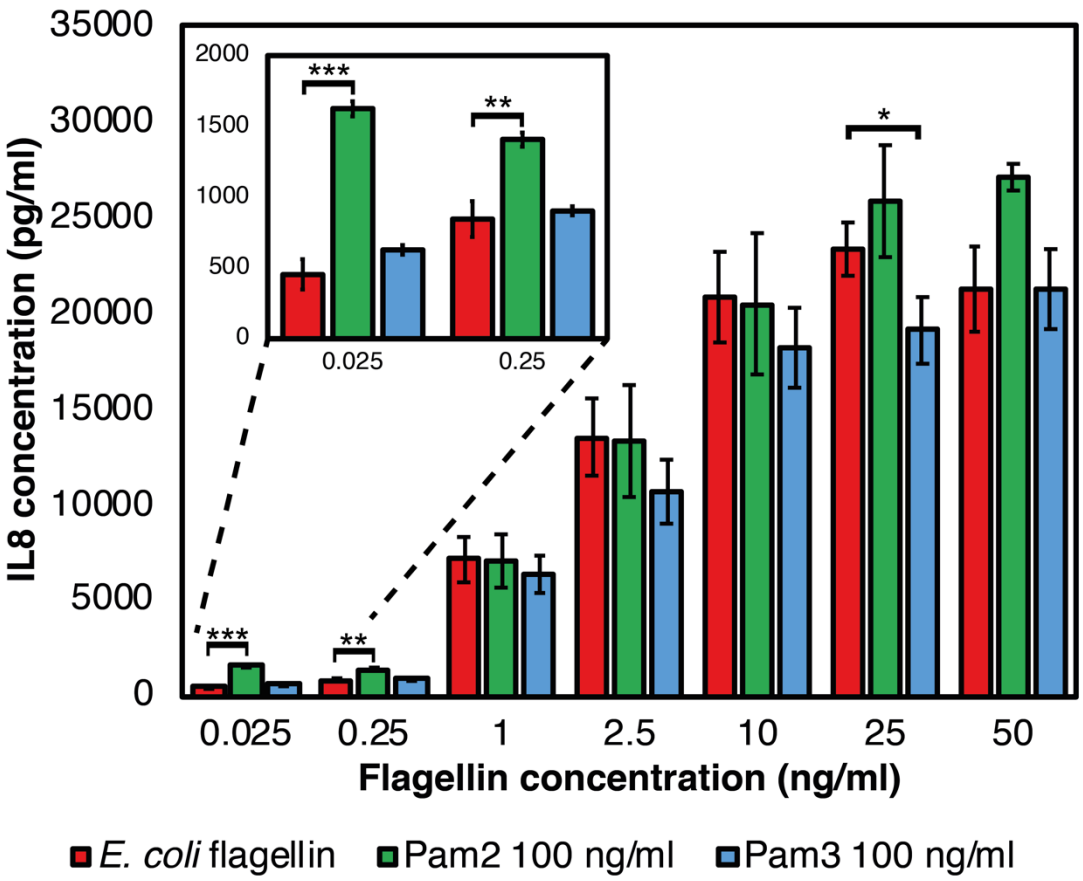



**Figure 69 TLR1/2/6 (Pam2 and Pam3) and TLR5 (flagellin) co-challenge assay findings.** The flagellin only control (red bars) was the baseline and was compared to cochallenges of flagellin + Pam2 100 (green bars) and flagellin + Pam3 100 (red bars). The IL8 findings at 0.025 and 0.25 ng/ml flagellin are also shown in smaller chart (top-left) on a different y-axis scale for ease of viewing. 2 experimental repeats with 6 biological repeats were performed for each concentration and agonist combination.

### *5.3.5 Summary*

RT4 cells expressed potentially non-functional TLR1 but functional TLR 2, 4, 5, and were thus comparable to TLR1 1805G patients. Challenges with heat-killed bacteria, to assess overall urothelial response, and purified flagellin preparations, to assess TLR5-specific responses demonstrated that *E. coli* induced a stronger IL8 urothelial response compared to *P. mirabilis*. Challenges with synthetic lipopeptide analogues, Pam2 and Pam3 to assess TLR1-specific and TLR2-specific responses respectively, induced a weak IL8 urothelial response compared to flag ellin. However, co-challenges of flagellin and Pam2/Pam3 lipopeptides, to assess TLR1/TLR2 and TLR5 crosstalk, demonstrated that Pam2 elevated the IL8 urothelial response at low flagellin concentrations whereas Pam3 lowered this response at a high flagellin concentrations.

It was surprising to observe that TLR1 signalling antagonised TLR5 signalling in RT4 cells despite lacking functional TLR1. It would be interesting to explore whether this antagonism

is more/less prominent in cells with functional TLR1 (TLR1 1805T genotype), and hence exploited by *P. mirabilis* to evade the immune response, allowing it to persist in the urinary tract of TLR1 1805T patients.

### *5.4 Investigating the nature of P. mirabilis colonisation in TLR1 1805T patients*

To determine whether TLR1 1805T BUTI patients were predisposed to persistent *P. mirabilis* colonisation compared to TLR1 1805G BUTI patients, the microbiological data generated from donated urine samples was analysed in a longitudinal manner. *P. mirabilis* colonisation was identified in 2/12 TLR1 1805G patients (non-functional TLR1), with both patients displaying sporadic *P. mirabilis* colonisation (<2 consecutive urine samples with *P. mirabilis*) (Figure 70A).

Amongst TLR 1805T patients, *P. mirabilis* colonisation was identified in 8/18 patients, with all eight patients displaying persistent *P. mirabilis* colonisation (≥2 consecutive urine samples with *P. mirabilis*) (Figure 70B). Four of these 8 patients (UTI 100, 139, 376, 755) were particularly interesting as *P. mirabilis* was identified in almost all of their urine samples, suggesting long-term colonisation. Similar to characterising *E. coli* colonisation in AnTIC (Section 3.3), it was necessary to genotype the *P. mirabilis* isolates in order to establish whether colonisation in these patients was stable or unstable.

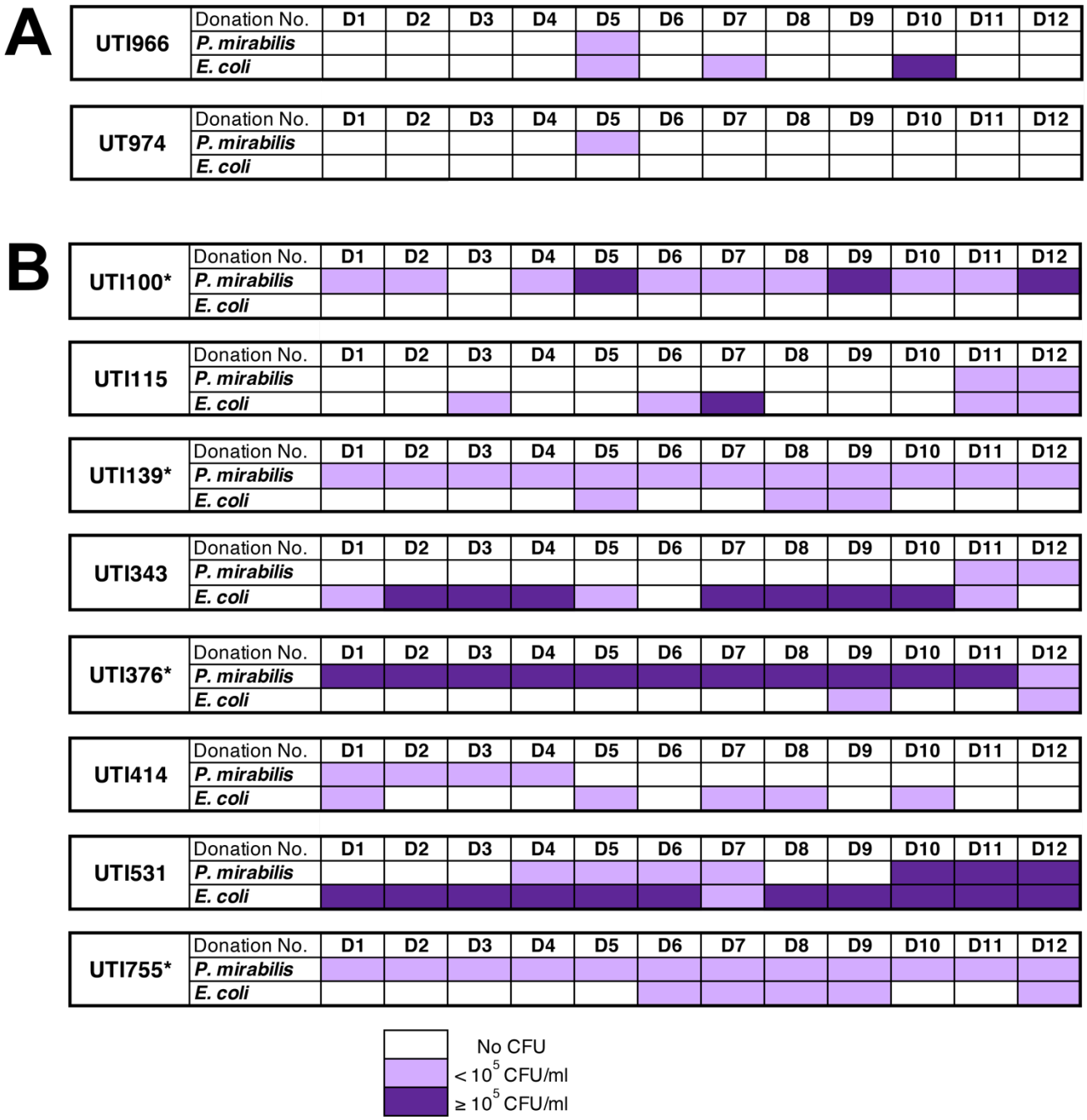



**Figure 70 Microbiological colonisation timeline of (A) TLR1 1805G and TLR1 1805T (B) patients with *P. mirabilis* colonisation.** The *P. mirabilis* and *E. coli* rows show the urine samples (D1-D12, Donation no. row) where they were absent (white cells) or present (light purple or dark purple cells). Light purple cells denote sub-clinical levels of colonisation in which $<10^5$ colony forming units (CFU/ml) of bacterial isolates were present whereas dark purple cells denote clinical levels of colonisation in which $\geq 10^5$ CFU/ml of bacterial isolates were present. (A) The two TLR1 1805G who carried *P. mirabilis* during the BUTI study displayed sporadic colonisation (B) The eight TLR1 1805T patients who carried *P. mirabilis* during the BUTI study displayed persistent *P. mirabilis* colonisation, in which there were at least two consecutive urine samples with *P. mirabilis.*

### 5.4.1 *Developing a novel MLST scheme for P. mirabilis*

MLST was used to genotype the BUTI *P. mirabilis* isolates, but due to the lack of a publicly available MLST scheme for this species, it was necessary to develop a novel MLST scheme. Most MLST schemes use seven genes for genotyping, which is both costly and time consuming (Section 1.8.3.3). It was therefore decided that the novel MLST scheme for *P. mirabilis* should employ just four genes. However, the viability and robustness of a 4-gene MLST scheme to generate phylogenetic data of sufficient resolution for research purposes first needed to be tested through exploiting an existing MLST scheme, i.e. the 7-gene Achtman *E. coli* MLST scheme (Figure 71).

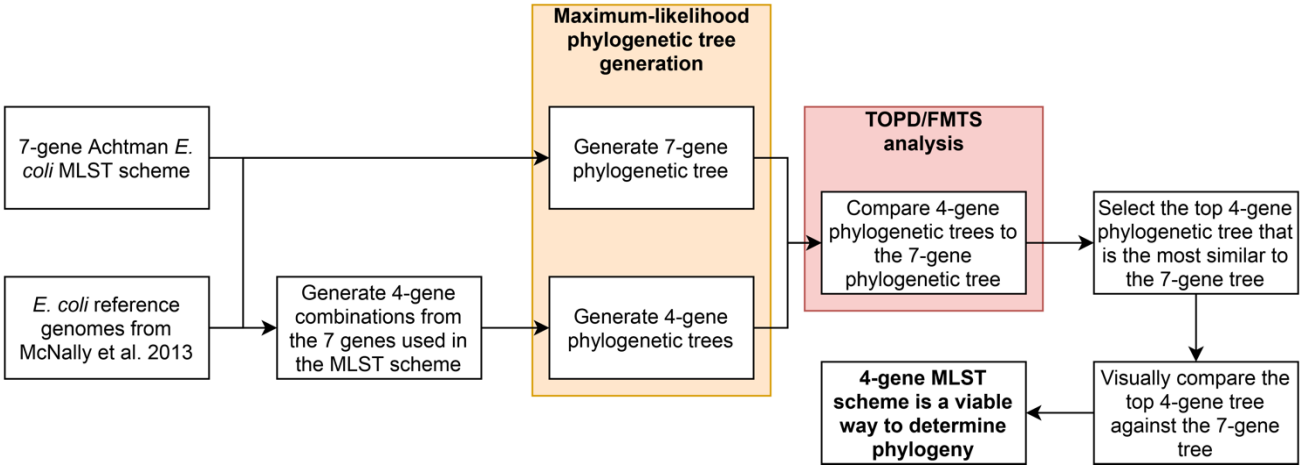



**Figure 71 Workflow to determine the viability of a 4-gene MLST scheme**

*E. coli* genomes (n=49) from an *E. coli* evolutionary study [317] and representative of the phylogenetic diversity within *E. coli* were used as reference genomes. Four-gene combinations were generated from the list of seven genes used in the Achtman scheme. Maximum-likelihood phylogenetic trees were constructed for each of the 36 4-gene combinations and compared against the "reference" 7-gene phylogenetic tree using TOPD/FMTS [376]. The TOPD/FMTS analysis performs pairwise comparison of a reference phylogenetic tree (7-gene) against query phylogenetic trees (4-gene) and generates two scores for each comparison. The first is the disagreement score, which is the number of taxa that were removed (trimmed) from both the reference and query trees so that they are as identical as possible to each other [376]. The second is the split distance score, i.e. the sum of the difference in phylogenetic distance for the remaining taxa between the "trimmed" reference and query trees, that were generated when calculating the disagreement score [376]. The lower the disagreement and split distance scores are, the more similar the query tree is to the reference tree [376].

Figure 72 shows the split distance score of the 4-gene trees when compared against the 7-gene tree. The leftmost bar with the lowest split distance score (0.224) is the 4-gene combination that generated the most similar 4-gene tree (*fumC, gyrB, icd, recA*). This tree was visually compared against the reference 7-gene tree (Figure 73A & B). Data indicated that the clonal complexes and clade structure of the 49 *E. coli* genomes were preserved in the 4-gene tree. These data justified the development of a 4-gene MLST scheme for *P. mirabilis*.

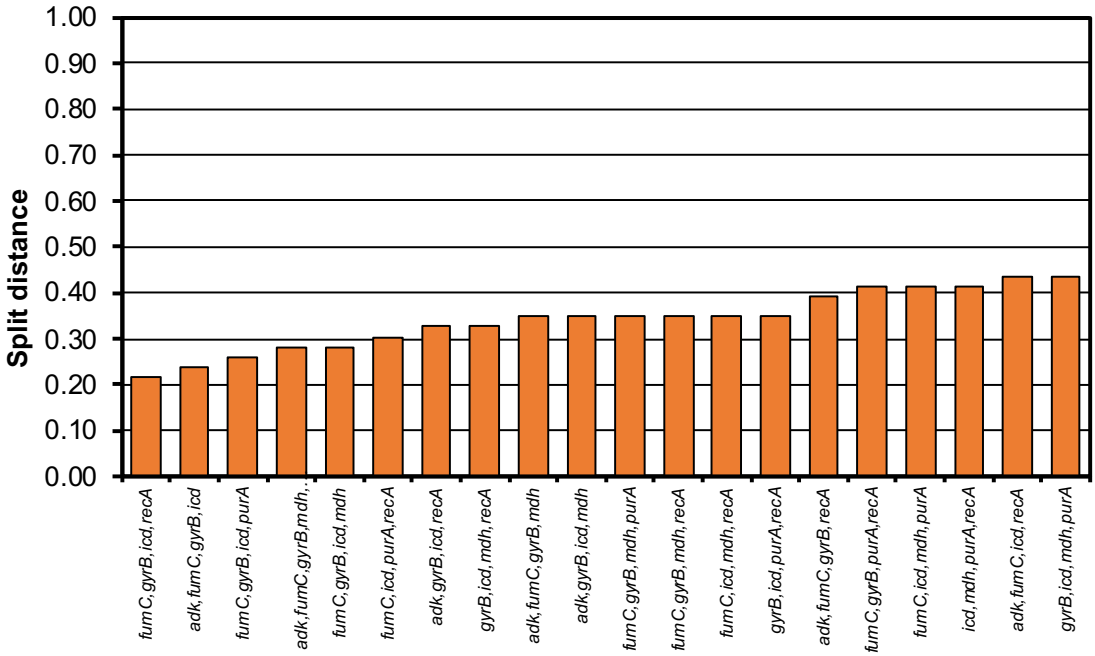



**Figure 72 Split distance scores of the top 20 4-gene trees when compared against the reference 7-gene tree.** Lower split distance score = more similar.

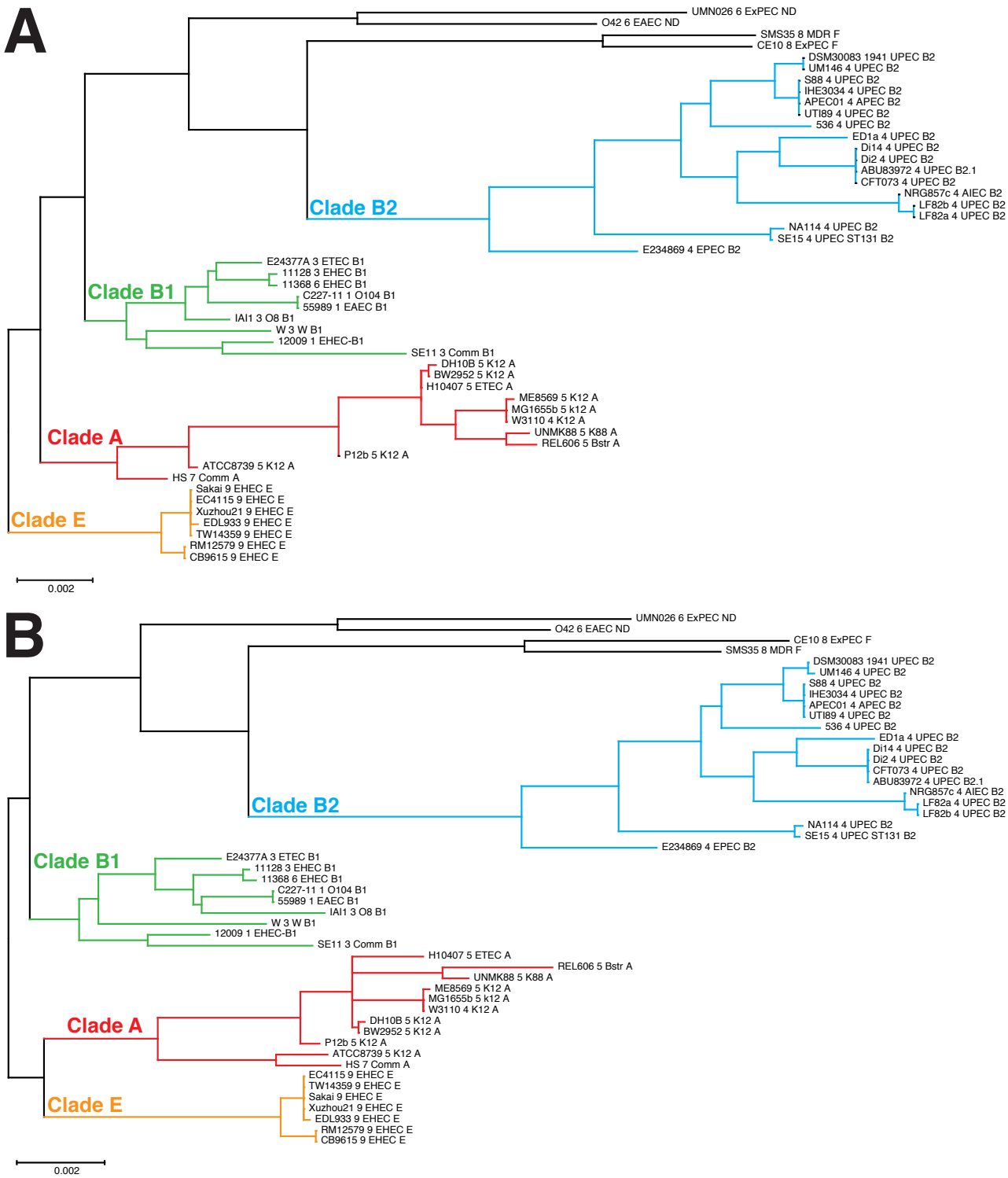



**Figure 73 Comparison of the most similar 4-gene tree (A) to the reference 7-gene tree (B).** The major *E. coli* clades (B2, B1, A, E) have been demarcated with coloured boxes. The pathotype (ExPEC, EAEC, UPEC, APEC, EHEC) and clade (B2, B1, A, E) of each *E. coli* isolate has also been included in the taxa entry. When comparing the 4-gene tree against the 7-gene tree, take note of the similarity in clade grouping between these two trees.

The next step was to choose four housekeeping genes that were evolutionary informative so that the novel MLST scheme generated robust phylogenetic data (Figure 74). Data mining was performed on an online MLST database (pubMLST) to generate a list of

housekeeping genes able to genotype Gram-negative bacteria (Figure 74). An original list of 93 genes was filtered for genes that were used in at least two MLST schemes, reducing the number to 24 (Figure 74, Appendix G). Of these 24 genes, 20 were found in *P. mirabilis* and there were two genes (*trp* and *mut*) that had two homologues (Appendix G). In total, there were 22 candidate genes that passed the initial filter (Appendix G; Figure 74).

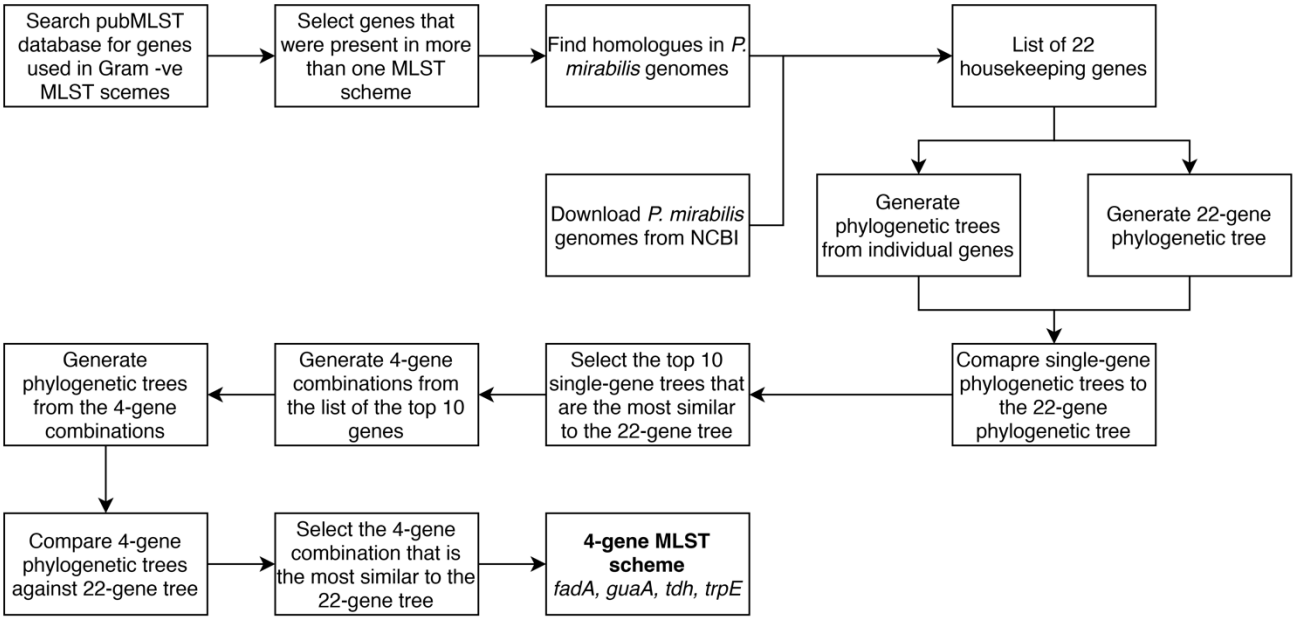



**Figure 74 Workflow to develop a novel 4-gene MLST scheme for *P. mirabilis***

Collectively, these 22 genes were used to generate a 22-gene phylogenetic tree from 64 *P. mirabilis* NCBI reference genomes (Figure 74). This 22-gene tree was used as a reference to determine which 4-gene combination was able to generate a phylogenetic tree that most closely resembled the 22-gene tree. Similarity was gauged using the split-distance scores generated by TOPD/FMTS analyses, with lower scores corresponding to greater similarity. However, generating 4-gene combinations from a list of 22 genes was not computationally viable as this would have generated 7315 unique combinations.

To reduce this number to a more manageable figure, it was decided that an additional filter should be implemented. Single-gene phylogenetic trees were generated from each of the 22 genes and compared against the 22-gene tree (Figure 74 & Figure 75). The top 10 genes that produced the most similar tree to the 22-gene tree (denoted by a red box in Figure 75) were used to create 4-gene combinations (210 unique combinations). Phylogenetic trees were then generated for each of the 210 4-gene combinations and compared against the 22-gene tree using TOPD (Figure 74, Figure 76).

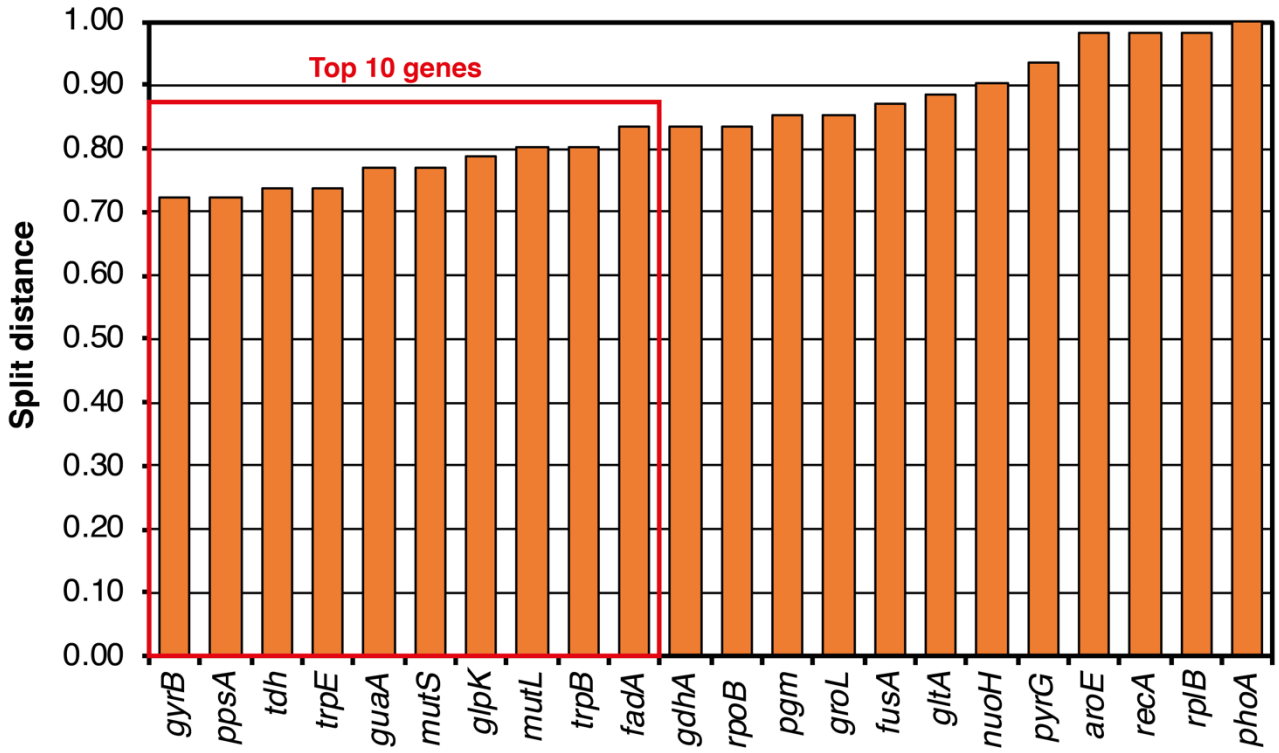



**Figure 75 Split distance scores of single-gene trees when compared against the reference 22-gene tree.** Lower split distance score = more similar.

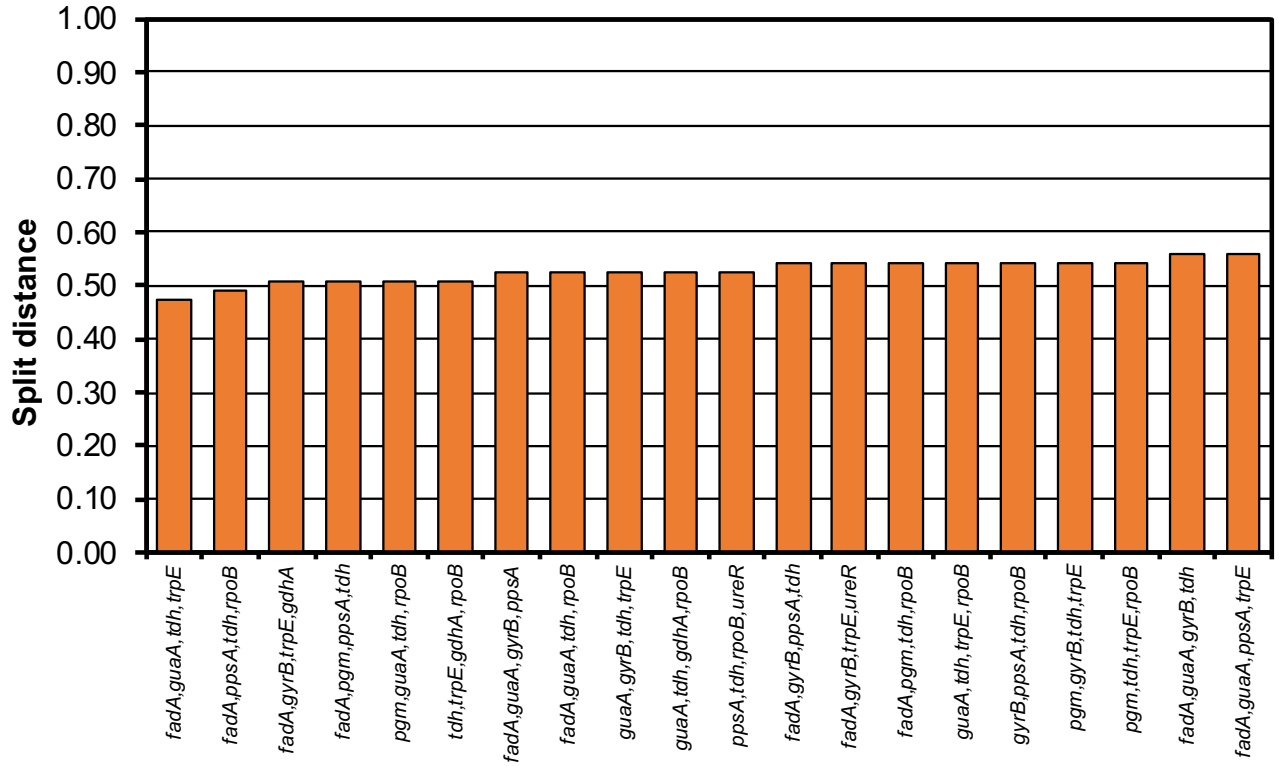



**Figure 76 Split distnace scores of the top 20 4-gene trees when compared against the reference 22-gene tree.** Lower split distance score = more similar. The range of split distance scores for all 210 trees is 0.48-0.75.

The 4-gene combination that generated the most similar phylogenetic tree to the reference 22-gene tree was *fadA, guaA, tdh, trpE*. Table 37 summarises the characteristics of each of these four genes. With the MLST scheme finalised, *in silico* MLST was performed on the 64 reference *P. mirabilis* genomes to generate a blastn MLST dataset (Appendix H).

**Table 37 Summary of the genes used in the novel 4-gene MLST scheme for *P. mirabilis*.**

| Gene | Fragment length | Nucleotide polymorphism | Function |
|---|---|---|---|
| *fadA* | 589 bp | 6.45% | 3-ketoacyl-CoA thisolase: Involved in fatty acid metabolism and lipid catabolism |
| *guaA* | 625 bp | 5.92% | GMP synthase: catalyses the synthesis of GMP from XMP |
| *tdh* | 469 bp | 7.68% | L-threonine 3-dehyodrogenase: catalyses the $NAD^+$-dependent oxidation of L-threonine to 2-amino-3-ketobutyrate. |
| *trpE* | 635 bp | 6.93% | Anthranilate synthase component 1: catalyses the biosynthesis of anthranilate from chorismite and L-glutamine. Anthranilate is an intermediate in the biosynthesis of L-tryptophan |

### *5.4.2 Genotyping P. mirabilis and comparing colonisation stability against E. coli*

A total of 70 unique *P. mirabilis* isolates (64 NCBI reference, 3 BUTI, 3 bovine) were genotyped using the novel customised MLST scheme. The 64 NCBI reference genomes were identical to those used earlier, the three BUTI isolates originated from patients UTI 100, 139, and 755 (Figure 78), while the three bovine isolates originated from a veterinary study that investigated AMR carriage in the bovine rumen. From this dataset of 70 isolates, 30 *fadA* alleles*, 38 guaA* alleles, 39 *tdh* alleles, and 32 *trpE* alleles were identified; with 49 unique allele combinations (PM numbers, Appendix H).

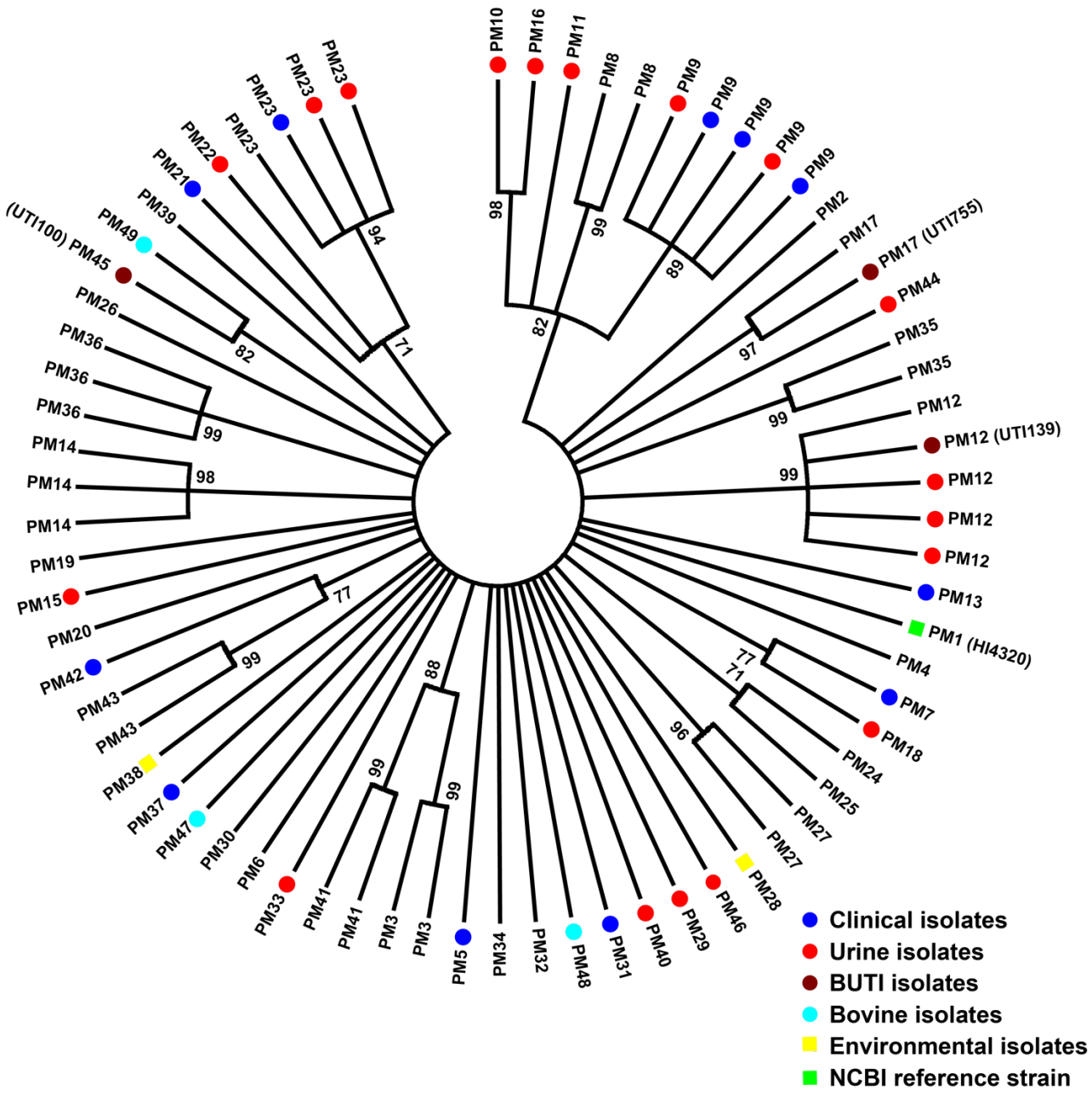



**Figure 77 Maximum-likelihood phylogenetic tree of *P. mirabilis* isolates , comprising of 64 NCBI reference isolates, 3 BUTI isolates, and 3 bovine isolates generated using the 4-gene MLST scheme.**

A maximum-likelihood phylogenetic tree was generated from the concatenated MLST nucleotide sequences of the 70 *P. mirabilis* isolates (Figure 77). Unlike with *E. coli* (Section 1.4), there appeared to be no phylogenetic clustering of *P. mirabilis* isolates based on isolation source or disease association (Figure 77). This lack of clustering suggests that urinary tract colonisation is a more universal trait in *P. mirabilis* compared to *E. coli*. This may be driven by robust flagellar and urease expression within *P. mirabilis* [209].Additionally, based on the MLST data, all three BUTI patients (UTI 100, 139, and 755) were found to be

colonised by their respective starting *P. mirabilis* strain throughout the study period, with Patient UTI 100 colonised by PM45, UTI 139 by PM12, and UTI 755 by PM17 (Figure 78).

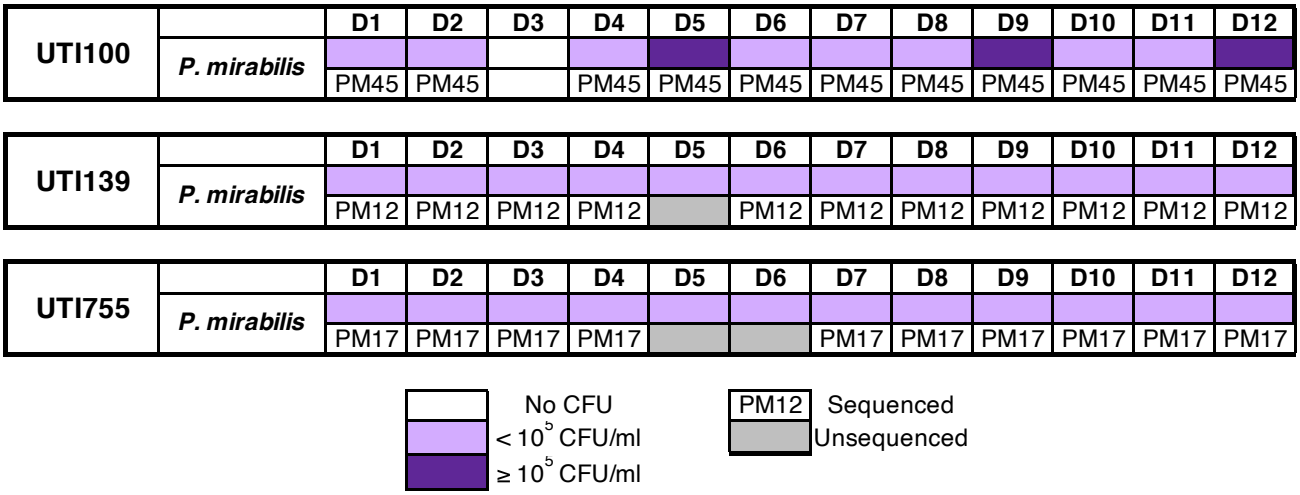



**Figure 78 Colonisation timeline of the three TLR1 1805T patients whose *P. mirabilis* isolates have been genotyped using the MLST scheme.** *P. mirabilis* colonisation was remarkably stable in all three patients.

### *5.4.3 Summary*

MLST genotyping of *P. mirabilis* isolates from three BUTI TLR 1805T patients (UTI 100, 139, 755) with persistent *P. mirabilis* colonisation (Figure 70) demonstrated stable colonisation of each of these patients, with no strain changes throughout the study period (Figure 78). Such stable colonisation suggests some degree of immune tolerance as the presence of bacteria in the urinary tract, especially highly flagellated bacteria (i.e. *P. mirabilis*) should induce a robust pro-inflammatory immune response as TLR5 is highly expressed in the urothelium (Section 1.2.3).

## *5.5 Discussion*

BUTI TLR1 1805T patients were predisposed to *P. mirabilis* colonisation (8/16 patients, 50% incidence) compared to TLR1 1805G patients (2/12 patients, 16.7% incidence). Although *P. mirabilis* colonisation in all eight TLR1 1805T patients was persistent (≥2 consecutive urine samples with *P. mirabilis*), there was a subset of patients (UTI 100, 139, 376, 755) where colonisation was unusually persistent, with *P. mirabilis* being present in all or nearly all of their urine samples. MLST genotyping of *P. mirabilis* isolates from three of these four patients (UTI 100, 139, 755) revealed that *P. mirabilis* colonisation in each patient was remarkably stable.

Efforts were made to investigate the immunological mechanism underlying this phenomenon. Analyses of BUTI cytokine data revealed that urinary IL1β among TLR1 1805G patients was significantly elevated during bacterial colonisation whereas TLR1 1850T patients did not experience any significant changes in cytokine secretion during bacterial colonisation (Figure 62). This was surprising as TLR1 1805T patients express functional TLR1 receptors, which TLR1 1805G patients lack. However, TLR signalling does more than just inducing a pro-inflammatory immune response as it can polarise the immune response if multiple TLRs are activated, which occurs frequently *in-vivo* as most pathogens express more than one PAMP, i.e. *P. mirabilis* possesses triacyl lipopeptides (TLR1-2 agonist), LPS (TLR4 agonist), and flagellin (TLR5 agonist).

During an infection, the immune response can be polarised into one of three major types: type 1, 2 and 3. All three responses are mediated by innate lymphoid cells (ILCs) and T lymphocytes, which differentiate into one of three effector subtypes (Table 38). Type 1 response is effective against intracellular bacteria, protozoa and viruses, and is induced by the secretion of IL12 and IFNγ. Type 2 response is effective against helminths and venoms and is induced by the secretion of IL4. Type 3 response is effective against extracellular bacteria and fungi, and is induced by the secretion of IL1β and IL23 (Figure 11.5 Janeway) [103]. Effector functions of the Type 3 response include the production of antimicrobial peptides by epithelial cells, increased epithelial cell turnover, and recruitment of neutrophils via IL8 secretion from stromal and epithelial cells [103,377]. Therefore, the Type 3 response is the most ideal in clearing *P. mirabilis* and other extracellular bacteria from the urinary tract. However, the effector subtypes which mediate the Type 3 response (ILC3, $T_H17$) exhibit greater plasticity compared to those which mediate the Type 1 and 2 responses [103,378]. Depending on prevailing cytokines, the Type 3 effector cells can transition into Type 1 effector cells. Furthermore, cytokines that induce the type 1 and 2 responses can antagonise type 3 effector functions [379].

**Table 38 Immune response polarisation types and the relevant ILC/T-lymphocyte effector subtype.**

| Immune response | ILC effector subtype | T lymphocyte effector subtype |
|---|---|---|
| Type 1 | ILC1 | $T_H1$ |
| Type 2 | ILC2 | $T_H2$ |
| Type 3 | ILC3 | $T_H17$ |

Each TLR member promotes or inhibits one of three immune response type. TLR 3, 7, 8, & 9 are associated with promoting the Type 1 response while strongly inhibiting Type 2 and 3 responses; TLR 1, 2, & 6 are associated with the Type 2 response while strongly inhibiting Type 1 and 3 responses; and TLR 4 & 5 are associated with the type 3 response while weakly inhibiting Type 1 and Type 2 responses [380]. It can be hypothesised that TLR1 signalling amongst TLR1 1805T patients could inhibit IL1β secretion as it induces a Type 2 response that antagonises the Type 3 response that is induced by TLR5, thus explaining why IL1β remained unchanged when bacterial colonisation was present (Figure 62B). However, TLR1 1805G patients lack functional TLR1 and thus the Type 3 response is not inhibited, as evidenced by the significant increase in IL1β (one of the principal cytokines of the Type 3 response) when bacterial colonisation was present (Figure 62A). This increase in IL1β is accompanied by a substantial drop in IFNγ as the Type 3 response inhibits the induction of the Type 1 response, of which IFNγ is the principal cytokine.

The TLR1 receptor detects triacyl lipopeptides that is expressed by all Gram-negative bacteria (Section 1.2.3) [374]. It is therefore curious as to why the persistent and stable colonisation exhibited by *P. mirabilis* was not shared by *E. coli* amongst TLR1 1805T patients. It can be speculated that *P. mirabilis* triacyl lipopeptide induces a more potent TLR1 response compared to *E. coli,* thereby inducing a strong type 2 response that prevents *P. mirabilis* from being cleared from the urinary tract. This hypothesis could be tested in a future experiment *in-vitro* by generating $TLR1^+$ mutants from RT4 cells ($TLR1^-$), which could be challenged with purified lipoproteins from *E. coli* and *P. mirabilis.*

To conclude, the significantly higher incidence of *P. mirabilis* colonisation amongst TLR1 1805T patients may relate to a strong TLR1-induced Type 2 immune response, preventing an effective Type 3 immune response from being mounted against *P. mirabilis* colonisation. The lack of TLR1 signalling among TLR1 1805G patients allows TLR5 signalling to induce an effective Type 3 immune response against *P. mirabilis* in an unimpeded manner, thereby greatly reducing the incidence of *P. mirabilis* colonisation in this group.

# Chapter 6 General Discussion

## 6.1 Key findings

1. Antibiotic prophylaxis reduced the incidence of UTI by stabilising *E. coli* colonisation in a subset of patients in addition to reducing *E. coli* colonisation in others.
2. The TLR1 1805T SNP was strongly associated with stable yet clinically asymptomatic *P. mirabilis* colonisation. A type 2 immune response promoted by TLR1 signalling could be responsible for this phenomenon
3. Contrary to existing findings, the low incidence of nitrofurantoin resistance is not due to the fitness cost of inactivating *nfsA* and *nfsB.* Instead, viable first-step mutants have a limited window of opportunity to outcompete their parental isolates, preventing the establishment of a first-step mutant population, which has a knock-on effect on the emergence of second-step, nitrofurantoin resistant mutants.

## 6.2 Discussion

The aims of this research were to investigate the host-microbial interactions that underlie UTIs by analysing data and clinical isolates collected from two longitudinal clinical studies, AnTIC and BUTI. The main difference between these two studies was that AnTIC focused on catheter-associated UTI as its study population consisted of CISC patients, whereas BUTI focused on uncomplicated UTI and asymptomatic bacteriuria as its study population consisted of elderly community-dwelling patients. Analysing the two study populations provided a more holistic view of UTI with comparisons between these populations allow the identification of common host-microbial interactions.

Firstly, the impact of antibiotic prophylaxis on bacterial colonisation of the urinary tract in clean intermittent self-catheterised (CISC) patients was investigated using data from the AnTIC study. A subset of these patients were managed with antibiotic prophylaxis, while the remainder were managed with short course antibiotics (non-prophylaxis). Antibiotic prophylaxis significantly reduced bacterial colonisation, especially that of *E. coli*, but interestingly also increased the relative incidence of minor uropathogens. A comparable observation has been reported in a recent urinary microbiome study by Vaughan MH et al. (2020) [381].

Despite the reduction in *E. coli* colonisation, there were prophylaxis patients who still carried *E. coli* in a persistent manner. Forty-seven (47) *E. coli* isolates from a subset of nine prophylaxis and eight non-prophylaxis patients who developed MDR UTI (n=17 patients) during AnTIC were sequenced and analysed to investigate the impact of antibiotic prophylaxis on *E. coli* colonisation. MLST and cgMLST phylogenetic analyses demonstrated that antibiotic prophylaxis stabilised *E. coli* colonisation, i.e. the same *E. coli* strain persisted in the urinary tract for an extended period. Within this dataset of 17 patients, antibiotic prophylaxis was also associated with a lower frequency of symptomatic UTIs compared to non-prophylaxis (0.75 vs 2.63 average cases, $p<0.017$). These findings suggest that antibiotic prophylaxis reduces the frequency of symptomatic UTIs by not only clearing bacteria from the urinary tract, but also through stabilising *E. coli* colonisation. Therefore, stabilising the colonisation of non-pathogenic bacteria within the lower urinary tract with antibiotic prophylaxis while controversial could be a viable approach to manage rUTIs. This approach to rUTI management is supported by a comparable study, which demonstrated a significant reduction in the frequency of symptomatic UTIs amongst rUTI patients who were deliberately inoculated with a non-pathogenic *E. coli* strain that was able to stably colonise the lower urinary tract [19].

The association between stable bacterial colonisation and a reduction in UTI frequency was also observed in a different study (BUTI) but with a different bacterial species. Instead of antibiotic prophylaxis promoting stable *E. coli* colonisation, it was potentially a host-specific TLR1 1805T SNP that was associated with stable and clinically asymptomatic *P. mirabilis* colonisation. MLST genotyping of *P. mirabilis* isolates demonstrated that three patients with the TLR1 1805T SNP were stably colonised by *P. mirabilis*, with the starting *P. mirabilis* strain being present within the urinary tract of each patient throughout the study. Another surprising observation was that none of the three patients were clinically diagnosed with an UTI throughout the entire study. Cytokine data suggested that TLR1 1805T patients were unable to mount a strong Type 3 immune response against extracellular bacteria due to the possible antagonistic effects of the Type 2 immune response induced by their functional TLR1 receptors. This diminished immune response may not only have prevented the complete clearance of *P. mirabilis* from the urinary tract, but could have also promoted a tolerogenic response towards *P. mirabilis*, allowing it to persist in the urinary tract without inducing a pro-inflammatory immune response. This would explain why TLR1 1805T patients with *P. mirabilis* colonisation remained clinically asymptomatic. Therefore, for TLR1

1805T patients, *P. mirabilis* may behave as a commensal organism as its presence is well-tolerated and such colonisation may also prevent pathogenic bacterial isolates from infecting the urinary tract via competitive means, e.g. nutrient competition due to its high motility and environmental modification via urease activity.

However, stable long-term colonisation by a single bacterial strain increases the likelihood of that strain developing antibiotic resistance, especially if antibiotic treatment was used to achieve this state. In AnTIC, bacterial isolates from prophylaxis patients revealed a significant incidence of antibiotic resistance [14]. Nitrofurantoin and trimethoprim are commonly used to manage UTIs in the UK and are both suitable for short-course and prophylaxis usage. Yet, the increasing incidence of trimethoprim resistance among uropathogens has led to nitrofurantoin being advised (NICE guidelines) [30] and used as the first-line drug for UTI management. Contrary to most antibiotics, nitrofurantoin resistance has remained low amongst *E. coli* (2-3%) despite its clinical usage over decades. Whether this trend will persist is uncertain, but the increasing usage of nitrofurantoin in UTI management means nitrofurantoin resistance needs to be further investigated and understood. Indeed, identifying the factors responsible for the low incidence of nitrofurantoin resistance should facilitate the development of appropriate prescribing measures to maximise the benefits of nitrofurantoin therapy for UTI patients (especially in prophylaxis usage) but also minimise the development of resistance before it becomes too widespread, thereby extending the lifespan of the drug.

Therefore, a key part of this thesis was to explore nitrofurantoin resistance by investigating the fitness dynamics underlying the acquisition of nitrofurantoin resistance. The acquisition of nitrofurantoin resistance involves a two-step evolutionary pathway, which involves the inactivation of *nfsA* followed by *nfsB.* Mutants were generated *in vitro* to reflect the potential evolutionary paths towards nitrofurantoin resistance using a nitrofurantoin sensitive AnTIC clinical isolate (1646 BASE), developed nitrofurantoin resistance *in vivo* six months into the AnTIC study (Figure 41). The fitness of these mutants was assessed in the absence and presence of nitrofurantoin to determine the cost-benefit trade-off of acquiring nitrofurantoin resistance. Contrary to an existing study, inactivation of *nfsA* and *nfsB* did not incur a fitness cost as the growth rates of the mutants and their parental isolates *in* vitro were identical in antibiotic free media. This behaviour continued until a concentration of 8 μg/ml of nitrofurantoin where mutants carrying inactivated *nfsA* started to outgrow their parental

isolates. At 16 μg/ml of nitrofurantoin, the parental isolates were no longer viable, but mutants carrying inactivated *nfsA* remained viable. However, the average doubling time of single mutants carrying inactivated *nfsA*, which represented the evolutionary ideal first-step mutants, at 16 μg/ml of nitrofurantoin was 60 minutes (Section 4.4), which exceeded the 36 minute minimum doubling time required for sustained bladder colonisation [271]. Interestingly, double mutants carrying both inactivated *nfsA* and *nfsB*, which represented the evolutionary ideal second-step mutants, were able to stay below this 36-minute threshold at 16 and 32 μg/ml of nitrofurantoin, with doubling times of 27.7 and 35.2 minutes respectively. It was only at 64 μg/ml of nitrofurantoin where their average doubling times increased to 103.4 minutes.

Therefore, evolutionarily viable single mutants with inactivated *nfsA* were only able to outcompete their parental isolates across a narrow range of nitrofurantoin concentration (~8 μg/ml). Findings generated from growth competition assays were comparable, but suggested an even lower nitrofurantoin concentration, ranging from 2 to 4 μg/ml. Such low levels of nitrofurantoin will not exist clinically in urine during short-course therapy, thereby, arguably, limiting the emergence of nitrofurantoin resistant mutants within patients treated for acute UTIs. Thus, the low incidence of nitrofurantoin resistance amongst *E. coli* isolates relates potentially to the inability of first-step mutants, single mutants with inactivated *nfsA*, to outcompete their parental isolates and establish themselves within the urinary tract of patients being managed with short-course nitrofurantoin therapy.

However, low-dose antibiotic prophylaxis in which a single 100mg dose of nitrofurantoin is taken daily could provide a sufficient window of opportunity for first-step mutants to establish themselves within a bacterial population. The single dosing practice inherent to low-dose antibiotic prophylaxis means that urinary nitrofurantoin levels will not be maintained by additional nitrofurantoin doses during the day, and potentially, will eventually reach the low concentration range where first-step mutants can outcompete their parental isolates. This model therefore suggests an increased likelihood of nitrofurantoin resistant mutants emerging amongst patients managed with low-dose nitrofurantoin prophylaxis. Furthermore, nitrofurantoin only displays bactericidal effects (4-log reduction) against *E. coli* when its urinary concentration is 8-fold greater than the MIC of an *E. coli* isolate, and is maintained for at least 6 hours [269]. However, the maximum urinary concentration of nitrofurantoin only reaches ~ 94.1μg/ml following a single 100mg dose [267], which is

substantially below the bactericidal 8-fold MIC threshold for first-step mutants (MIC: 32 μg/ml, 8-fold MIC threshold: 256 μg/ml) and $Nit^S$ isolates (MIC: 16 μg/ml, 8-fold MIC threshold: 128 μg/ml). Therefore, the bactericidal effects of nitrofurantoin cannot be relied on to eliminate a first-step mutant once it emerges within a population as urinary nitrofurantoin levels fail to reach levels that are high enough to exert bactericidal effects on first-step mutants let alone $Nit^S$ isolates.

This presents a dilemma as (i) AnTIC has demonstrated that low-dose antibiotic prophylaxis is effective at reducing the frequency of symptomatic UTIs amongst CISC patients suffering from rUTIs and (ii) nitrofurantoin is still highly effective against *E. coli,* the main causative agent of UTI. However, these data suggest that the widespread use of low-dose nitrofurantoin prophylaxis will probably create reservoirs of resistant mutants that will disseminate into the wider *E. coli* population resulting in an increase in the incidence of nitrofurantoin resistant *E. coli* isolates. Since the acquisition of nitrofurantoin resistance does not impact bacterial fitness, reducing nitrofurantoin usage is unlikely to reverse this trend once it gains momentum, which fortunately has still yet to happen.

It could be argued that this scenario has already happened with trimethoprim as evidenced by the explosive increase in trimethoprim resistance during the 1990s and 2000s due to its excessive usage in the decades prior [382]. Fortunately, trimethoprim resistance does come with a fitness cost and as a result of trimethoprim being supplanted by nitrofurantoin in the management of UTIs, there has been a substantial reduction in the incidence of trimethoprim resistant urinary isolates [383]. However, any increase in the incidence of nitrofurantoin resistance will likely be permanent, thus raising the urgency for better stewardship this antibiotic before that happens.

Based on these data, the question needs to be asked whether nitrofurantoin should be used as a low-dose antibiotic prophylaxis treatment to manage recurrent UTIs. There are already other antibiotics that can be used for low-dose prophylaxis such as trimethoprim and cephalexin where it should be possible to manage the incidence of resistance if antibiotic cycling/rotation is practiced on a national level.

These *in vitro* and bioinformatic studies have shown that stable bacterial colonisation induced by antibiotic prophylaxis or a host-specific factor (TLR1 1805T SNP) is associated with a reduction in UTI symptoms. Although the usage of antibiotic prophylaxis to promote

stable bacterial colonisation appears promising, care is potentially needed when using nitrofurantoin in this manner. The pharmaceutical properties of nitrofurantoin and the fitness dynamics underlying the evolutionary acquisition of nitrofurantoin resistance in *E. coli* can interact in ways that could increase the likelihood of nitrofurantoin resistant mutants emerging amongst patients managed with low dose nitrofurantoin prophylaxis. Although there already exists a pharmacokinetic study which assessed urinary nitrofurantoin levels amongst young healthy females, there is a need to conduct further research on its pharmacokinetics across a wider range of patients, including elderly and catheterised individuals, and dosing methods so that current prescription strategies can be optimised to maximise efficacy while minimising the emergence of resistance. Optimising nitrofurantoin usage will also require a better understanding of the competitive fitness of first-step mutants against their parental isolates at low nitrofurantoin concentrations (2-8 μg/ml).

Lastly, the discovery of a potential host-specific factor that was associated with asymptomatic and stable bacterial colonisation in BUTI raises the possibility of an alternative to antibiotic usage to manage rUTIs. A clinical study has demonstrated the efficacy of urinary inoculation with an ASB *E. coli* strain (Strain 83972) to reduce the frequency of symptomatic UTIs amongst individuals with treatment resistant rUTIs. The efficacy of such inoculations may be enhanced by tailoring the bacterial isolate based on the host's immune profile (TLR SNP profiling) to maximise the possibility of long-term asymptomatic colonisation. However, it will be prudent to better understand the immunological mechanisms driving this phenomenon before further developing urinary inoculations for rUTI management. Future analyses should, for example, investigate the effects of the TLR1 1805T SNP on the immune response towards *P. mirabilis* and *E. coli* using a mixed cell culture of urothelial and antigen presenting cells.

In conclusion, promoting stable and asymptomatic bacterial colonisation, although controversial, could be a viable approach to managing rUTIs. Although this is achievable via low-dose nitrofurantoin prophylaxis, pharmacokinetics data [267] suggest that such usage will lead to extended periods of low urinary nitrofurantoin levels, which potentially favours the emergence of resistant mutants (Section 4.4). Urinary inoculation could be a promising alternative as demonstrated by the efficacy of *E. coli* strain 83972 at reducing rUTIs [19]. However, the efficacy of such therapies may be determined by host-specific factors, i.e. TLR SNPs. This is highlighted by microbiological and MLST data from this study, which

suggested that a subset of BUTI patients with the TLR1 1805T SNP experienced stable and asymptomatic *P. mirabilis* colonisation during the study. Therefore, investigating both the host and microbial components and how they interact are essential to understanding the pathogenesis of UTI, and thus improve its management.

# Chapter 7 References


1. Tan, C. W. & Chlebicki, M. P. Urinary tract infections in adults. *Singapore Med. J.* **57**, 485–490 (2016).

2. Foxman, B. Epidemiology of urinary tract infections: Incidence, morbidity, and economic costs. *Am. J. Med.* **113**, 5–13 (2002).

3. Bjorling, D. E., Wang, Z. Y. & Bushman, W. Models of inflammation of the lower urinary tract. *Neurourology and Urodynamics* vol. 30 673–682 (2011).

4. Beahm, N. P., Nicolle, L. E., Bursey, A., Smyth, D. J. & Tsuyuki, R. T. The assessment and management of urinary tract infections in adults: Guidelines for pharmacists. *Can. Pharm. J.* **150**, 298–305 (2017).

5. MIHAĽOVÁ, M., CELEC, P. & BREZA, J. Urosepsis. *Lek. Obz.* **6**, 364–368 (2020).

6. Gupta, K. & Trautner, B. W. Diagnosis and management of recurrent urinary tract infections in non-pregnant women. *BMJ* **346**, f3140–f3140 (2013).

7. Schmiemann, G., Kniehl, E., Gebhardt, K., Matejczyk, M. M. & Hummers-Pradier, E. The diagnosis of urinary tract infection: a systematic review. *Dtsch. Arztebl.* **107**, 361–367 (2010).

8. Hickling, D. R., Sun, T.-T. & Wu, X.-R. Anatomy and Physiology of the Urinary Tract: Relation to Host Defense and Microbial Infection. *Microbiol. Spectr.* **3**, (2015).

9. Storme, O., Tirán Saucedo, J., Garcia-Mora, A., Dehesa-Dávila, M. & Naber, K. G. Risk factors and predisposing conditions for urinary tract infection. *Ther. Adv. Urol.* **11**, 175628721881438 (2019).

10. NICE. *Urinary tract infection (recurrent): antimicrobial prescribing NICE guideline*. (2018).

11. Foxman, B. Recurring urinary tract infection: Incidence and risk factors. *Am. J. Public Health* **80**, 331–333 (1990).

12. Naziri, Z., Derakhshandeh, A., Borchaloee, A. S., Poormaleknia, M. & Azimzadeh, N.

Treatment failure in urinary tract infections: A warning witness for virulent multi-drug resistant ESBL-producing Escherichia coli. *Infect. Drug Resist.* **13**, 1839–1850 (2020).

13. Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR)*. www.facebook.com/PublicHealthEngland (2016).

14. Fisher, H. *et al.* Continuous low-dose antibiotic prophylaxis for adults with repeated urinary tract infections (AnTIC): a randomised, open-label trial. *Lancet Infect. Dis.* **18**, 957–968 (2018).

15. Albert, X. *et al.* Antibiotics for preventing recurrent urinary tract infection in non-pregnant women. *Cochrane Database Syst. Rev.* **2004**, (2004).

16. Cormican, M., Murphy, A. W. & Vellinga, A. Interpreting asymptomatic bacteriuria: Testing for and treating bacteriuria in children and non-pregnant adults without specific symptoms of urinary tract infection or sepsis is of uncertain benefit. *BMJ* **343**, (2011).

17. Givler, D. N. & Givler, A. *Asymptomatic Bacteriuria*. *StatPearls* (StatPearls Publishing, 2021).

18. Roos, V., Ulett, G. C., Schembri, M. A. & Klemm, P. The asymptomatic bacteriuria Escherichia coli strain 83972 outcompetes uropathogenic E. coli strains in human urine. *Infect. Immun.* **74**, 615–624 (2006).

19. Sundén, F., Håkansson, L., Ljunggren, E. & Wullt, B. Escherichia coli 83972 Bacteriuria Protects Against Recurrent Lower Urinary Tract Infections in Patients With Incomplete Bladder Emptying. *J. Urol.* **184**, 179–185 (2010).

20. Llor, C. *et al.* The adherence of GPs to guidelines for the diagnosis and treatment of lower urinary tract infections in women is poor. *Fam. Pract.* **28**, 294–299 (2011).

21. Car, J. *et al.* Urinary tract infections in women: diagnosis and management in primary care. *BMJ* **332**, 94–7 (2006).

22. Public Health England. *Diagnosis of urinary tract infections*. (2017).

23. Heytens, S. *et al.* Women with symptoms of a urinary tract infection but a negative urine culture: PCR-based quantification of Escherichia coli suggests infection in most cases. *Clin. Microbiol. Infect.* **23**, 647–652 (2017).

24. Demilie, T., Beyene, G., Melaku, S. & Tsegaye, W. Diagnostic accuracy of rapid urine dipstick test to predict urinary tract infection among pregnant women in Felege Hiwot Referral Hospital, Bahir Dar, North West Ethiopia. *BMC Res. Notes 2014 71* **7**, 1–5 (2014).

25. Mambatta, A. *et al.* Reliability of dipstick assay in predicting urinary tract infection. *J. Fam. Med. Prim. Care* **4**, 265 (2015).

26. Eigbefoh, J. O., Isabu, P., Okpere, E. & Abebe, J. The diagnostic accuracy of the rapid dipstick test to predict asymptomatic urinary tract infection of pregnancy. *J. Obstet. Gynaecol. (Lahore).* **28**, 490–495 (1991).

27. Patel, H. D., Livsey, S. A., Swann, R. A. & Bukhari, S. S. Can urine dipstick testing for urinary tract infection at point of care reduce laboratory workload? *J. Clin. Pathol.* **58**, 951–954 (2005).

28. PHE. *UK Standards for Microbiology Investigations: Investigation of urine*. (2019).

29. Phe. *UK Standards for Microbiology Investigations-standards-for-microbiology-investigations-smi-quality-and-consistency-in-clinical-laboratories PHE Publications gateway number: 2015306 UK Standards for Microbiology Investigations are produced in association with*. (2019).

30. National Institute for Health and Care Excellence. *Urinary tract infection (lower): antimicrobial prescribing*. www.nice.org.uk/guidance/ng109 (2018).

31. Joint Formulary Committee. *British National Formulary*. (London: BMJ Group and Pharmaceutical Press, 2021).

32. Jepson, R. G., Williams, G. & Craig, J. C. Cranberries for preventing urinary tract infections. *Cochrane Database Syst. Rev.* (2012) doi:10.1002/14651858.CD001321.pub5.

33. Cooper, T. E. *et al.* D-mannose for preventing and treating urinary tract infections.

*Cochrane Database Syst. Rev.* **2020**, (2020).

34. Crellin, E. *et al.* Trimethoprim use for urinary tract infection and risk of adverse outcomes in older patients: Cohort study. *BMJ* **360**, 341 (2018).

35. Public Health England. *English surveillance programme for antimicrobial utilisation and resistance (ESPAUR) from 2010 to 2014: 2015 Report*. *Public Health England Publications* www.gov.uk/phe%0Awww.facebook.com/PublicHealthEngland (2019).

36. National Health Service England. *Technical Guidance Annex B Information on Quality Premium*. (2018).

37. Sanchez, G. V., Master, R. N., Karlowsky, J. A. & Bordon, J. M. In vitro antimicrobial resistance of urinary Escherichia coli isolates among U.S. outpatients from 2000 to 2010. *Antimicrob. Agents Chemother.* **56**, 2181–2183 (2012).

38. NICE. National Institute for Health and Care Excellence: Urinary tract infections in adults act infections in adults. (2015).

39. National Institute for Health and Care Excellence. *UTI (recurrent): antimicrobial prescribing*. (2018).

40. KASS, E. H. Bacteriuria and the diagnosis of infections of the urinary tract; with observations on the use of methionine as a urinary antiseptic. *AMA. Arch. Intern. Med.* **100**, 709–714 (1957).

41. Dacheva, T. & Malone-Lee, J. The problems affecting the diagnosis of urinary tract infection. *Aging health* **8**, 537–545 (2012).

42. Hay, A. D. *et al.* Microbiological diagnosis of urinary tract infection by NHS and research laboratories. (2016).

43. Kupelian, A. S. *et al.* Discrediting microscopic pyuria and leucocyte esterase as diagnostic surrogates for infection in patients with lower urinary tract symptoms: Results from a clinical and laboratory evaluation. *BJU Int.* **112**, 231–238 (2013).

44. Pels, R. J., Bor, D. H., Woolhandler, S., Himmelstein, D. U. & Lawrence, R. S. Dipstick Urinalysis Screening of Asymptomatic Adults for Urinary Tract Disorders: II.

Bacteriuria. *JAMA J. Am. Med. Assoc.* **262**, (1989).

45. Khasriya, R. *et al.* The Inadequacy of Urinary Dipstick and Microscopy as Surrogate Markers of Urinary Tract Infection in Urological Outpatients With Lower Urinary Tract Symptoms Without Acute Frequency and Dysuria. *J. Urol.* **183**, 1843–1847 (2010).

46. Siegman-Igra, Y., Kulka, T., Schwartz, D. & Konforti, N. The significance of polymicrobial growth in urine: Contamination or true infection. *Scand. J. Infect. Dis.* **25**, 85–91 (1993).

47. Byrd, B. A. L., Segre, J. A. & Koch, R. Adapting Koch's postulates. (2016).

48. Kline, K. A. & Lewis, A. L. Gram-Positive Uropathogens, Polymicrobial Urinary Tract Infection, and the Emerging Microbiota of the Urinary Tract. *Microbiol. Spectr.* **4**, (2016).

49. Armbruster, C. E. *et al.* The Pathogenic Potential of Proteus mirabilis Is Enhanced by Other Urinary Tract Infection. *Infect. Immun.* **85**, e00808-16 (2017).

50. Drage, L. K. L. A longitudinal study of long-term Escherichia coli colonisation of the elderly bladder. (Newcastle University, 2017).

51. Peters, B. M., Jabra-Rizk, M. A., O'May, G. A., William Costerton, J. & Shirtliff, M. E. Polymicrobial interactions: Impact on pathogenesis and human disease. *Clin. Microbiol. Rev.* **25**, 193–213 (2012).

52. Nelson, A., De Soyza, A., Perry, J. D., Sutcliffe, I. C. & Cummings, S. P. Polymicrobial challenges to Koch's postulates: Ecological lessons from the bacterial vaginosis and cystic fibrosis microbiomes. *Innate Immun.* **18**, 774–783 (2012).

53. Graham, J. C. & Galloway, A. The Laboratory Diagnosis of Urinary Tract Infection. *J Clin Pathol* **54**, 911–919 (2001).

54. SK, A., T, N., N, R. & K, N. Laboratory diagnosis of urinary tract infections using diagnostics tests in adult patients. *Int. J. Res. Med. Sci.* **2**, 415 (2014).

55. Wilson M, G. L. Laboratory diagnosis of urinary tract infections in adult patients. *Med Microbiol* **38**, 8 (2004).

56. Sakai, Y. Low-count organisms concealed by dominant uropathogenic orginisms in urine of patients with asymptomatic bacteriuria. *Int. J. Urol.* **2**, 96–99 (1995).

57. Morgan, M. G. & McKenzie, H. Controversies in the laboratory diagnosis of community-acquired urinary tract infection. *Eur. J. Clin. Microbiol. Infect. Dis.* **12**, 491–504 (1993).

58. Hilt, E. E. *et al.* Urine is not sterile: Use of enhanced urine culture techniques to detect resident bacterial flora in the adult female bladder. *J. Clin. Microbiol.* **52**, 871–876 (2014).

59. Wolfe, A. J. *et al.* Evidence of uncultivated bacteria in the adult female bladder. *J. Clin. Microbiol.* **50**, 1376–1383 (2012).

60. Wagenlehner, F., Wullt, B., Ballarini, S., Zingg, D. & Naber, K. G. Social and economic burden of recurrent urinary tract infections and quality of life: a patient web-based study (GESPRIT). *Expert Rev. Pharmacoecon. Outcomes Res.* **18**, 1–11 (2017).

61. Renard, J. *et al.* Recurrent Lower Urinary Tract Infections Have a Detrimental Effect on Patient Quality of Life: a Prospective, Observational Study. *Infect. Dis. Ther.* **4**, 125–135 (2015).

62. Malone-Lee, J. Urinary infections are complex and hard to treat. *Bmj* **5766**, j5766 (2017).

63. Hannan, T. J. *et al.* Host-pathogen checkpoints and population bottlenecks in persistent and intracellular uropathogenic Escherichia coli bladder infection. *FEMS Microbiol. Rev.* **36**, 616–648 (2012).

64. Moore, K. L., Aguar, A. M. R. & Dalley, A. F. *Moore Essential Clinical Anatomy*. (Wolters Kluwer Health, 2015).

65. Marieb, E. N. & Hoehn, K. *Human Anatomy & Physiology*. (Pearson Education Inc., 2013).

66. Oncology Pro. Tumours of the Urinary System. https://oncologypro.esmo.org/education-library/essentials-for-clinicians/genitourinary-

tract-tumours/tumours-of-the-urinary-system.

67. Pfau, A. & Sacks, T. The bacterial flora of the vaginal vestibule, urethra and vagina in premenopausal women with recurrent urinary tract infections. *J. Urol.* **126**, 630–634 (1981).

68. Haya, J., García, A., López-Manzanara, C., Balawi, M. & Haya, L. Importance of Lactic Acid in Maintaining Vaginal Health: A Review of Vaginitis and Vaginosis Etiopathogenic Bases and a Proposal for a New Treat-ment. *Open J. Obstet. Gynecol.* **4**, 787–799 (2014).

69. Hawes, S. E. *et al.* Hydrogen peroxide-producing lactobacilli and acquisition of vaginal infections. *J. Infect. Dis.* **174**, 1058–1063 (1996).

70. Klebanoff, S. J., Hillier, S. L., Eschenbach, D. A. & Waltersdorph, A. M. Control of the microbial flora of the vagina by H2O2-generating lactobacilli. *J. Infect. Dis.* **164**, 94–100 (1991).

71. Stapleton, A. E. The Vaginal Microbiota and Urinary Tract Infection. *Microbiol. Spectr.* **4**, 1–6 (2016).

72. Perrotta, C., Aznar, M., Mejia, R., Albert, X. & Ng, C. W. Oestrogens for preventing recurrent urinary tract infection in postmenopausal women. *Cochrane Database of Systematic Reviews* (2008) doi:10.1002/14651858.CD005131.pub2.

73. Invivo Healthcare. Female EcologiX: Your New Clinical Tool. https://invivohealthcare.com/education/articles/female-ecologix-your-new-clinical-tool/.

74. Wiswell, T. E. The prepuce, urinary tract infections, and the consequences. *Pediatrics* vol. 105 860–862 (2000).

75. Drenth, J. J. The prepuce. *Ned. Tijdschr. Geneeskd.* **124**, 299–301 (1980).

76. Neubert, U. & Lentze, I. Bacterial flora of the preputial space [German]. *Hautarzt* **30**, 149–153 (1979).

77. Schoen, E. J., Colby, C. J. & Ray, G. T. Newborn circumcision decreases incidence

and costs of urinary tract infections during the first year of life. *Pediatrics* **105**, 789–793 (2000).

78. Kunin, C. M., Polyak, F. & Postel, E. Periurethral Bacterial Flora in Women: Prolonged Intermittent Colonization With Escherichia coli. *JAMA J. Am. Med. Assoc.* **243**, 134–139 (1980).

79. Yamamoto, S. *et al.* Genetic evidence supporting the fecal-perineal-urethral hypothesis in cystitis caused by Escherichia coli. *J. Urol.* **157**, 1127–1129 (1997).

80. Glover, M., Moreira, C. G., Sperandio, V. & Zimmern, P. Recurrent urinary tract infections in healthy and nonpregnant women. *Urol. Sci.* **25**, 1–8 (2014).

81. Cass, A. S. & Ireland, G. W. Antibacterial perineal washing for prevention of recurrent urinary tract infections. *Urology* **25**, 492–494 (1985).

82. Khandelwal, P., Abraham, S. N. & Apodaca, G. Cell biology and physiology of the uroepithelium. *American Journal of Physiology - Renal Physiology* vol. 297 1477–1501 (2009).

83. Carlile, A., Davies, I., Faragher, E. & Brocklehurst, J. C. The epithelium in the female urethra: A quantitative study. *J. Urol.* **138**, 775–777 (1987).

84. Stoddard, N. & Leslie, S. W. *Histology, Male Urethra*. *StatPearls* (StatPearls Publishing, 2020).

85. Jost, S. P. Cell cycle of normal bladder urothelium in developing and adult mice. *Virchows Arch. B Cell Pathol. Incl. Mol. Pathol.* **57**, 27–36 (1989).

86. Hicks, R. M. The mammalian urinary bladder: an accommodating organ. *Biological Reviews of the Cambridge Philosophical Society* vol. 50 215–246 (1975).

87. Kreft, M. E., Sterle, M., Veranič, P. & Jezernik, K. Urothelial injuries and the early wound healing response: Tight junctions and urothelial cytodifferentiation. *Histochem. Cell Biol.* **123**, 529–539 (2005).

88. Rajasekaran, M., Stein, P. & Parsons, C. L. Toxic factors in human urine that injure urothelium. *Int. J. Urol.* **13**, 409–414 (2006).

89. Jackson, A. R., Ching, C. B., McHugh, K. M. & Becknell, B. Roles for urothelium in normal and aberrant urinary tract development. *Nature Reviews Urology* vol. 17 459–468 (2020).

90. Jost, S. P., Gosling, J. A. & Dixon, J. S. The morphology of normal human bladder urothelium. *J. Anat.* **167**, 103–15 (1989).

91. Khandelwal, P., Abraham, S. N. & Apodaca, G. Cell biology and physiology of the uroepithelium. *American Journal of Physiology - Renal Physiology* vol. 297 F1477 (2009).

92. Apodaca, G. The uroepithelium: Not just a passive barrier. *Traffic* vol. 5 117–128 (2004).

93. Wu, X. R., Kong, X. P., Pellicer, A., Kreibich, G. & Sun, T. T. Uroplakins in urothelial biology, function, and disease. *Kidney International* vol. 75 1153–1165 (2009).

94. Parsons, C. L. The Role of the Urinary Epithelium in the Pathogenesis of Interstitial Cystitis/Prostatitis/Urethritis. *Urology* **69**, S9–S16 (2007).

95. Wu, X. R., Sun, T. T. & Medina, J. J. In vitro binding of type 1-fimbriated Escherichia coli to uroplakins Ia and Ib: Relation to urinary tract infections. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 9630–9635 (1996).

96. Pollen, J. J., Anwar, H., Stauffer, C. & Schmidt, J. D. Antibacterial activity of bladder surface mucin duplicated in the rabbit bladder by exogenous glycosaminoglycan (sodium pentosanpolysulfate). *Infect. Immun.* **27**, 876–881 (1980).

97. Cornish, J., Lecamwasam, J. P., Harrison, G., Vanderwee, M. A. & Miller, T. E. Host defence mechanisms in the bladder. II. Disruption of the layer of mucus. *Br. J. Exp. Pathol.* **69**, 759–770 (1988).

98. Behzadi, E. & Behzadi, P. The role of Toll-Like receptors (TLRs) in urinary tract infections (UTIs). *Cent. Eur. J. Urol.* **69**, 404–410 (2016).

99. Kawasaki, T. & Kawai, T. Toll-Like Receptor Signaling Pathways. *Front. Immunol.* **0**, 461 (2014).

100. Rehli, M. Of mice and men: species variations of Toll-like receptor expression. *Trends in immunology* vol. 23 375–378 (2002).

101. Huhta, H. *et al.* The Expression of Toll-like Receptors in Normal Human and Murine Gastrointestinal Organs and the Effect of Microbiome and Cancer. *J. Histochem. Cytochem.* **64**, 470–482 (2016).

102. Carey, A. J. *et al.* Urinary tract infection of mice to model human disease: Practicalities, implications and limitations. *Crit. Rev. Microbiol.* **42**, 780–799 (2016).

103. Murphy, K. & Weaver, C. *Janeway's Immunobiology*. (Garland Science, 2017).

104. Di Gioia, M. & Zanoni, I. Toll-like receptor co-receptors as master regulators of the immune response. *Mol. Immunol.* **63**, 143–152 (2015).

105. Lee, C. C., Avalos, A. M. & Ploegh, H. L. Accessory molecules for Toll-like receptors and their function. *Nat. Rev. Immunol.* **12**, 168–179 (2012).

106. Godfroy, J. I., Roostan, M., Moroz, Y. S., Korendovych, I. V. & Yin, H. Isolated Toll-like Receptor Transmembrane Domains Are Capable of Oligomerization. *PLoS One* **7**, (2012).

107. El-Zayat, S. R., Sibaii, H. & Mannaa, F. A. Toll-like receptors activation, signaling, and targeting: an overview. *Bull. Natl. Res. Cent.* **43**, 1–12 (2019).

108. Gao, W., Xiong, Y., Li, Q. & Yang, H. Inhibition of toll-like receptor signaling as a promising therapy for inflammatory diseases: A journey from molecular to nano therapeutics. *Frontiers in Physiology* vol. 8 (2017).

109. Parham, P. & Murphy, K. *The immune system*. (Garland Science).

110. Chan, C. Y., St. John, A. L. & Abraham, S. N. Mast Cell IL-10 Drives Localized Tolerance in Chronic Bladder Infection. *Immunity* **38**, 349–359 (2013).

111. O'Neill, L. A. J., Golenbock, D. & Bowie, A. G. The history of Toll-like receptors-redefining innate immunity. *Nat. Rev. Immunol.* **13**, 453–460 (2013).

112. Zasloff, M. Antimicrobial peptides, innate immunity, and the normally sterile urinary tract. *Journal of the American Society of Nephrology* vol. 18 2810–2816 (2007).

113. Nielsen, K. L. *et al.* Role of urinary cathelicidin LL-37 and human β-defensin 1 in uncomplicated escherichia coli urinary tract infections. *Infect. Immun.* **82**, 1572–1578 (2014).

114. Zasloff, M. Antimicrobial peptides of multicellular organisms. *Nature* vol. 415 389–395 (2002).

115. De Yang, B. *et al.* LL-37, the neutrophil granule- and epithelial cell-derived cathelicidin, utilizes formyl peptide receptor-like 1 (FPRL1) as a receptor to chemoattract human peripheral blood neutrophils, monocytes, and T cells. *J. Exp. Med.* **192**, 1069–1074 (2000).

116. Lacerda Mariano, L. & Ingersoll, M. A. The immune response to infection in the bladder. *Nat. Rev. Urol.* **17**, 439–458 (2020).

117. Gomes, A. C., Moreira, A. C., Mesquita, G. & Gomes, M. S. Modulation of iron metabolism in response to infection: Twists for all tastes. *Pharmaceuticals* vol. 11 (2018).

118. Johnson, E. E. & Wessling-Resnick, M. Iron metabolism and the innate immune response to infection. *Microbes and Infection* vol. 14 207–216 (2012).

119. Flo, T. H. *et al.* Lipocalin 2 mediates an innate immune response to bacterial infection by sequestrating iron. *Nature* **432**, 917–921 (2004).

120. Patras, K. A. *et al.* Augmentation of Urinary Lactoferrin Enhances Host Innate Immune Clearance of Uropathogenic Escherichia coli. *J. Innate Immun.* **11**, (2019).

121. Abraham, S. N. & Miao, Y. The nature of immune responses to urinary tract infections. *Nat. Rev. Immunol.* **15**, 655–663 (2015).

122. Mysorekar, I. U., Isaacson-Schmid, M., Walker, J. N., Mills, J. C. & Hultgren, S. J. Bone Morphogenetic Protein 4 Signaling Regulates Epithelial Renewal in the Urinary Tract in Response to Uropathogenic Infection. *Cell Host Microbe* **5**, 463–475 (2009).

123. Mulvey, M. A., Schilling, J. D. & Hultgren, S. J. Establishment of a persistent Escherichia coli reservoir during the acute phase of a bladder infection. *Infect. Immun.* **69**, 4572–4579 (2001).

124. Nagamatsu, K. *et al.* Dysregulation of escherichia coli α-hemolysin expression alters the course of acute and persistent urinary tract infection. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E871–E880 (2015).

125. Mysorekar, I. U. & Hultgren, S. J. Mechanisms of uropathogenic Escherichia coli persistence and eradication from the urinary tract. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 14170–14175 (2006).

126. Hedges, S. *et al.* Uroepithelial cells are part of a mucosal cytokine network. *Infect. Immun.* **62**, 2315–2321 (1994).

127. SVANBORG, C., AGACE, W., HEDGES, S., LINDSTEDT, R. & SVENSSON, M. L. Bacterial Adherence and Mucosal Cytokine Production. *Ann. N. Y. Acad. Sci.* **730**, 162–181 (1994).

128. Haraoka, M. *et al.* Neutrophil recruitment and resistance to urinary tract infection. *J. Infect. Dis.* **180**, 1220–1229 (1999).

129. Agace, W. W., Patarroyo, M., Svensson, M., Carlemalm, E. & Svanborg, C. Escherichia coli induces transuroepithelial neutrophil migration by an intercellular adhesion molecule-1-dependent mechanism. *Infect. Immun.* **63**, 4054–4062 (1995).

130. Shanin, R., Engberd, I., Hagbers, L. & Eden, C. Neutrophil recruitment and resistance to urinary tract infection. *J. Infect. Dis.* **180**, 1220–1229 (1999).

131. Winterbourn, C. C., Kettle, A. J. & Hampton, M. B. Reactive Oxygen Species and Neutrophil Function. *Annu. Rev. Biochem.* **85**, 765–792 (2016).

132. Hannan, T. J. *et al.* Inhibition of cyclooxygenase-2 prevents chronic and recurrent cystitis. *EBioMedicine* **1**, 46–57 (2014).

133. Hannan, T. J., Mysorekar, I. U., Hung, C. S., Isaacson-Schmid, M. L. & Hultgren, S. J. Early Severe Inflammatory Responses to Uropathogenic E. coli Predispose to Chronic and Recurrent Urinary Tract Infection. *PLOS Pathog.* **6**, e1001042 (2010).

134. O'Brien, V. P., Hannan, T. J., Schaeffer, A. J. & Hultgren, S. J. Are you experienced? Understanding bladder innate immunity in the context of recurrent urinary tract infection. *Curr. Opin. Infect. Dis.* **28**, (2015).

135. Schiwon, M. *et al.* Crosstalk between sentinel and helper macrophages permits neutrophil migration into infected uroepithelium. *Cell* **156**, 456–468 (2014).

136. Engel, D. R. *et al.* CCR2 Mediates Homeostatic and Inflammatory Release of Gr1 high Monocytes from the Bone Marrow, but Is Dispensable for Bladder Infiltration in Bacterial Urinary Tract Infection . *J. Immunol.* **181**, 5579–5586 (2008).

137. Nathan, C. Neutrophils and immunity: Challenges and opportunities. *Nature Reviews Immunology* vol. 6 173–182 (2006).

138. Prame Kumar, K., Nicholls, A. J. & Wong, C. H. Y. Partners in crime: neutrophils and monocytes/macrophages in inflammation and disease. *Cell and Tissue Research* vol. 371 551–565 (2018).

139. Abraham, S. N. & St. John, A. L. Mast cell-orchestrated immunity to pathogens. *Nature Reviews Immunology* vol. 10 440–452 (2010).

140. Abraham, S. N., Shin, J. S. & Malaviya, R. Type 1 fimbriated Escherichia coli-mast cell interactions in cystitis. *J. Infect. Dis.* **183**, S51–S55 (2001).

141. Zhu, K., Hill, W. G., Li, F., Shi, B. & Chai, T. C. Early Increased Urinary IL-2 and IL-10 Levels Were Associated With Development of Chronic UTI in a Murine Model. *Urology* **141**, 188.e1-188.e6 (2020).

142. Drage, L. K. L. *et al.* Elevated urine IL-10 concentrations associate with Escherichia coli persistence in older patients susceptible to recurrent urinary tract infections. doi:10.1186/s12979-019-0156-9.

143. Mora-Bau, G. *et al.* Macrophages Subvert Adaptive Immunity to Urinary Tract Infection. *PLoS Pathog.* **11**, e1005044 (2015).

144. Ragnarsdóttir, B., Lutay, N., Grönberg-Hernandez, J., Köves, B. & Svanborg, C. Genetics of innate immunity and UTI susceptibility. *Nat. Rev. Urol.* **8**, 449–468 (2011).

145. Skevaki, C., Pararas, M., Kostelidou, K., Tsakris, A. & Routsias, J. G. Single nucleotide polymorphisms of Toll-like receptors and susceptibility to infectious diseases. *Clin. Exp. Immunol.* **180**, 165–177 (2015).

146. Hawn, T. R. *et al.* A common human TLR1 polymorphism regulates the innate immune response to lipopeptides. *Eur. J. Immunol.* **37**, 2280–2289 (2007).

147. Johnson, C. M. *et al.* Cutting Edge: A Common Polymorphism Impairs Cell Surface Trafficking and Functional Responses of TLR1 but Protects against Leprosy. *J. Immunol.* **178**, 7520–7524 (2007).

148. Wujcicka, W., Paradowska, E., Studzińska, M., Wilczyński, J. & Nowakowska, D. TLR2 2258 G>A single nucleotide polymorphism and the risk of congenital infection with human cytomegalovirus. *Virol. J.* **14**, 1–12 (2017).

149. Xiong, Y., Song, C., Snyder, G. A., Sundberg, E. J. & Medvedev, A. E. R753Q polymorphism inhibits toll-like receptor (TLR) 2 tyrosine phosphorylation, dimerization with TLR6, and recruitment of myeloid differentiation primary response protein 88. *J. Biol. Chem.* **287**, 38327–38337 (2012).

150. Kormann, M. S. D. *et al.* Rare TLR2 mutations reduce TLR2 receptor function and can increase atopy risk. *Allergy Eur. J. Allergy Clin. Immunol.* **64**, 636–642 (2009).

151. Hawn, T. R. *et al.* Genetic variation of the human urinary tract innate immune response and asymptomatic bacteriuria in women. *PLoS One* **4**, (2009).

152. Tabel, Y., Berdeli, A. & Mir, S. Association of TLR2 gene Arg753Gln polymorphism with urinary tract infection in children. *Int. J. Immunogenet.* **34**, 399–405 (2007).

153. Medvedev, A. E. Toll-Like Receptor Polymorphisms, Inflammatory and Infectious Diseases, Allergies, and Cancer. *J. Interf. Cytokine Res.* **33**, 467 (2013).

154. Hawn, T. R. *et al.* Toll-like receptor 4 polymorphisms are associated with resistance to Legionnaires' disease. *Proc. Natl. Acad. Sci.* **102**, 2487–2489 (2005).

155. Arbour, N. C. *et al.* TLR4 mutations are associated with endotoxin hyporesponsiveness in humans. *Nat. Genet.* **25**, 187–191 (2000).

156. Samuelsson, P. *et al.* Toll-Like Receptor 4 Expression and Cytokine Responses in the Human Urinary Tract Mucosa. *Infect. Immun.* **72**, 3179–3186 (2004).

157. Abdi, J., Mutis, T., Garssen, J. & Redegeld, F. Characterization of the Toll-like

Receptor Expression Profile in Human Multiple Myeloma Cells. *PLoS One* **8**, e60671 (2013).

158. Smith, N. J. *et al.* Toll-like receptor responses of normal human urothelial cells to bacterial flagellin and lipopolysaccharide. *J. Urol.* **186**, 1084–1092 (2011).

159. Hawn, T. R. *et al.* Toll-like receptor polymorphisms and susceptibility to urinary tract infections in adult women. *PLoS One* **4**, e5990 (2009).

160. Yin, X. *et al.* Association of Toll-like receptor 4 gene polymorphism and expression with urinary tract infection types in adults. *PLoS One* **5**, 1–7 (2010).

161. Karoly, E. *et al.* Heat shock protein 72 (HSPA1B) gene polymorphism and toll-like receptor (TLR) 4 mutation are associated with increased risk of urinary tract infection in children. *Pediatr. Res.* **61**, 371–374 (2007).

162. Hawn, T. R. *et al.* A stop codon polymorphism of Toll-like receptor 5 is associated with resistance to systemic lupus erythematosus. *Proc. Natl. Acad. Sci.* **102**, 10593–10597 (2005).

163. Ali, A. S. M. *et al.* Targeting Deficiencies in the TLR5 Mediated Vaginal Response to Treat Female Recurrent Urinary Tract Infection. *Sci. Rep.* **7**, 1–14 (2017).

164. Andersen-Nissen, E. *et al.* Cutting Edge: Tlr5 −/− Mice Are More Susceptible to Escherichia coli Urinary Tract Infection . *J. Immunol.* **178**, 4717–4720 (2007).

165. Lane, M. C., Alteri, C. J., Smith, S. N. & Mobley, H. L. T. Expression of flagella is coincident with uropathogenic Escherichia coli ascension to the upper urinary tract. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 16669–16674 (2007).

166. Flores-Mireles, A. L., Walker, J. N., Caparon, M. & Hultgren, S. J. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Micro* **13**, 269–284 (2015).

167. Hossain, M. *et al.* Genotype–phenotype correlation of β-lactamase-producing uropathogenic Escherichia coli (UPEC) strains from Bangladesh. *Sci. Rep.* **10**, 14549 (2020).

168. Bahrani-Mougeot, F., Gunther, N. W., Donnenberg, M. S. & Mobley, H. L. T. Uropathogenic Escherichia coli. in *Escherichia Coli* 239–268 (Elsevier, 2002). doi:10.1016/B978-012220751-8/50009-4.

169. Chaudhuri, R. R. & Henderson, I. R. The evolution of the Escherichia coli phylogeny. *Infect. Genet. Evol.* **12**, 214–226 (2012).

170. Ejrnaes, K. Bacterial Characteristics of Importance for Recurrent Urinary Tract Infections Caused by Escherichia coli. (University Copenhagen, 2009).

171. Ejrnæs, K. *et al.* Characteristics of Escherichia coli causing persistence or relapse of urinary tract infections: Phylogenetic groups, virulence factors and biofilm formation. *Virulence* **2**, 528–537 (2011).

172. Tenaillon, O., Skurnik, D., Picard, B. & Denamur, E. The population genetics of commensal Escherichia coli. *Nat. Rev. Microbiol.* **8**, 207–217 (2010).

173. Denamur, E., Clermont, O., Bonacorsi, S. & Gordon, D. The population genetics of pathogenic Escherichia coli. *Nat. Rev. Microbiol.* doi:10.1038/s41579-020-0416-x.

174. Gibreel, T. M. *et al.* Population structure, virulence potential and antibiotic susceptibility of uropathogenic Escherichia coli from Northwest England. doi:10.1093/jac/dkr451.

175. Nicolas-Chanoine, M.-H., Bertrand, X. & Madec, J.-Y. Escherichia coli ST131, an Intriguing Clonal Group. (2014) doi:10.1128/CMR.00125-13.

176. Petty, N. K. *et al.* Global dissemination of a multidrug resistant Escherichia coli clone. doi:10.1073/pnas.1322678111.

177. Mcnally, A. *et al.* Diversification of Colonization Factors in a Multidrug-Resistant Escherichia coli Lineage Evolving under Negative Frequency-Dependent Selection. (2019) doi:10.1128/mBio.00644-19.

178. Lehtinen, S. *et al.* Evolution of antibiotic resistance is linked to any genetic mechanism affecting bacterial duration of carriage. *Proc. Natl. Acad. Sci.* **114**, 1075–1080 (2017).

179. Katouli, M. Population structure of gut Escherichia coli and its role in development of extra-intestinal infections. *Iranian Journal of Microbiology* vol. 2 59–72 (2010).

180. Spaulding, C. N. *et al.* Selective depletion of uropathogenic E. coli from the gut by a FimH antagonist. *Nature* **546**, 528–532 (2017).

181. Kaper, J. B., Nataro, J. P. & Mobley, H. L. T. PATHOGENIC ESCHERICHIA COLI. *Nat. Rev.* | *Microbiol.* **2**, 3 (2004).

182. Sauer, M. M. *et al.* Binding of the Bacterial Adhesin FimH to Its Natural, Multivalent High-Mannose Type Glycan Targets. *J. Am. Chem. Soc* **141**, 936–944 (2019).

183. Köves, B. & Wullt, B. The Roles of the Host and the Pathogens in Urinary Tract Infections. *Eur. Urol. Suppl.* **15**, 88–94 (2016).

184. Scott, V. C. S. S., Haake, D. A., Churchill, B. M., Justice, S. S. & Kim, J.-H. H. Intracellular Bacterial Communities: A Potential Etiology for Chronic Lower Urinary Tract Symptoms. *Urology* **86**, 425–431 (2015).

185. Ali, A. S. M. *et al.* Targeting Deficiencies in the TLR5 Mediated Vaginal Response to Treat Female Recurrent Urinary Tract Infection. *Sci. Rep.* **7**, 11039 (2017).

186. Omar, M., Abdulwahab-Ahmed, A., Chaparala, H. & Monga, M. Does Stone Removal Help Patients with Recurrent Urinary Tract Infections? *J. Urol.* **194**, 997–1001 (2015).

187. Borghi, L., Nouvenne, A. & Meschi, T. Nephrolithiasis and urinary tract infections: 'The chicken or the egg' dilemma? *Nephrol. Dial. Transplant.* **27**, 3982–3985 (2012).

188. Cooper, J., Brumfitt, W., Hamilton-Miller, J. M. & Reynolds, A. V. The role of periurethral colonization in the aetiology of recurrent urinary infection in women. *Br. J. Obstet. Gynaecol.* **87**, 1145–1151 (1980).

189. Brumfitt, W., Gargan, R. A. & Hamilton-Miller, J. M. Periurethral enterobacterial carriage preceding urinary infection. *Lancet (London, England)* **1**, 824–826 (1987).

190. Connell, H. *et al.* Type 1 fimbrial expression enhances Escherichia coli virulence for the urinary tract. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 9827–9832 (1996).

191. Gunther IV, N. W., Lockatell, V., Johnson, D. E. & Mobley, H. L. T. In vivo dynamics

of type 1 fimbria regulation in uropathogenic Escherichia coli during experimental urinary tract infection. *Infect. Immun.* **69**, 2838–2846 (2001).

192. Korhonen, T. K., Virkola, R. & Holthofer, H. Localization of binding sites for purified Escherichia coli P fimbriae in the human kidney. *Infect. Immun.* **54**, 328–332 (1986).

193. Nielubowicz, G. R. & Mobley, H. L. T. Host-pathogen interactions in urinary tract infection. *Nat. Rev. Urol.* **7**, 430–441 (2010).

194. Hagan, E. C., Donnenberg, M. S. & Mobley, H. L. T. Uropathogenic Escherichia coli. *EcoSal Plus* **3**, (2009).

195. Lillington, J., Geibel, S. & Waksman, G. Biogenesis and adhesion of type 1 and P pili. *Biochimica et Biophysica Acta - General Subjects* vol. 1840 2783–2793 (2014).

196. Zhou, G. *et al.* Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: evidence from in vitro FimH binding. *J. Cell Sci.* **114**, 4095–4103 (2001).

197. Chen, S. L. *et al.* Positive selection identifies an in vivo role for FimH during urinary tract infection in addition to mannose binding. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 22439–22444 (2009).

198. Melican, K. *et al.* Uropathogenic escherichia coli P and type 1 fimbriae act in synergy in a living host to facilitate renal colonization leading to nephron obstruction. *PLoS Pathog.* **7**, 1001298 (2011).

199. Simms, A. N. & Mobley, H. L. T. PapX, a P fimbrial operon-encoded inhibitor of motility in uropathogenic Escherichia coli. *Infect. Immun.* **76**, 4833–4841 (2008).

200. Chevance, F. F. V. & Hughes, K. T. Coordinating assembly of a bacterial macromolecular machine. *Nat. Rev. Microbiol. 2008 66* **6**, 455–465 (2008).

201. Andersen-Nissen, E. *et al.* Evasion of Toll-like receptor 5 by flagellated bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 9247–9252 (2005).

202. Smith, K. D. *et al.* Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nat. Immunol.* **4**, (2003).

203. Frick-Cheng, A. E. *et al.* The gene expression profile of uropathogenic escherichia

coli in women with uncomplicated urinary tract infections is recapitulated in the mouse model. *MBio* **11**, 1–16 (2020).

204. Lane, M. C., Simms, A. N. & Mobley, H. L. T. Complex interplay between type 1 fimbrial expression and flagellum-mediated motility of uropathogenic Escherichia coli. *J. Bacteriol.* **189**, 5523–5533 (2007).

205. Lane, M. C. *et al.* Role of Motility in the Colonization of Uropathogenic Escherichia coli in the Urinary Tract Role of Motility in the Colonization of Uropathogenic Escherichia coli in the Urinary Tract. *Infect. Immun.* **73**, 7644–7656 (2005).

206. Wright, K. J., Seed, P. C. & Hultgren, S. J. Uropathogenic Escherichia coli Flagella Aid in Efficient Urinary Tract Colonization. *Infect. Immun.* **73**, 7657–7668 (2005).

207. Lane, M. C. *et al.* Role of motility in the colonization of uropathogenic Escherichia coli in the urinary tract. *Infect. Immun.* **73**, 7644–7656 (2005).

208. Hagan, E. C., Lloyd, A. L., Rasko, D. A., Faerber, G. J. & Mobley, H. L. T. Escherichia coli Global Gene Expression in Urine from Women with Urinary Tract Infection. *PLoS Pathog* **6**, 1001187 (2010).

209. Schaffer, J. N. & Pearson, M. M. Proteus mirabilis and Urinary Tract Infections. *Microbiol. Spectr.* **3**, (2015).

210. Bonnin, R. A. *et al.* A single Proteus mirabilis lineage from human and animal sources: a hidden reservoir of OXA-23 or OXA-58 carbapenemases in Enterobacterales. *Sci. Rep.* **10**, 1–9 (2020).

211. Papazafiropoulou, A. *et al.* Prevalence of asymptomatic bacteriuria in type 2 diabetic subjects with and without microalbuminuria. *BMC Res. Notes* **3**, 169 (2010).

212. Matthews, S. J. & Lancaster, J. W. Urinary tract infections in the elderly population. *American Journal Geriatric Pharmacotherapy* vol. 9 286–309 (2011).

213. CM, O., FW, B. & JM, M. Classification, identification, and clinical significance of Proteus, Providencia, and Morganella. *Clin. Microbiol. Rev.* **13**, 534–546 (2000).

214. Schaffer, J. N. & Pearson, M. M. Proteus mirabilis and Urinary Tract Infections.

*Microbiol Spectr.* **3**, 1–39 (2015).

215. SM, J., DJ, S., HL, M. & ME, S. Complicated catheter-associated urinary tract infections due to Escherichia coli and Proteus mirabilis. *Clin. Microbiol. Rev.* **21**, 26–59 (2008).

216. HS, S., RK, F. & RN, J. Frequency of occurrence and antimicrobial susceptibility of Gram-negative bacteremia isolates in patients with urinary tract infection: results from United States and European hospitals (2009-2011). *J. Chemother.* **26**, 133–138 (2014).

217. E, L. *et al.* Bacteremia in a multilevel geriatric hospital. *J. Am. Med. Dir. Assoc.* **12**, 204–207 (2011).

218. Armbruster, C. E., Mobley, H. L. T. & Pearson, M. M. Pathogenesis of Proteus mirabilis Infection. *EcoSal Plus* **8**, (2018).

219. MM, G. *et al.* Traditional Foley drainage systems--do they drain the bladder? *J. Urol.* **177**, 203–207 (2007).

220. PA, W. & S, C. Current trends in the management of difficult urinary catheterizations. *West. J. Emerg. Med.* **13**, 472–478 (2012).

221. Bolla, S. R., Odeluga, N. & Jetti, R. Histology, Bladder. *StatPearls* (2021).

222. Flores-Mireles, A. L. *et al.* Fibrinogen Release and Deposition on Urinary Catheters Placed during Urological Procedures. *J. Urol.* **196**, 416–421 (2016).

223. Lee, K.-H. *et al.* The Influence of Urinary Catheter Materials on Forming Biofilms of Microorganisms. *J. Bacteriol. Virol.* **47**, 32–40 (2017).

224. HL, M. & JW, W. Urease-positive bacteriuria and obstruction of long-term urinary catheters. *J. Clin. Microbiol.* **25**, 2216–2217 (1987).

225. Torzewska, A. & Rozalski, A. Inhibition of crystallization caused by Proteus mirabilis during the development of infectious urolithiasis by various phenolic substances. *Microbiol. Res.* **169**, 579–584 (2014).

226. Chew, R. *et al. Large urate cystolith associated with Proteus urinary tract infection*.

*Kidney international* vol. 81 (2012).

227. Li, X. *et al.* Visualization of Proteus mirabilis within the matrix of urease-induced bladder stones during experimental urinary tract infection. *Infect. Immun.* **70**, 389–394 (2002).

228. Mobley, H. L. & Belas, R. Swarming and pathogenicity of Proteus mirabilis in the urinary tract. *Trends Microbiol.* **3**, 280–284 (1995).

229. Jones, B. D., Lockatell, C. V, Johnson, D. E., Warren, J. W. & Mobley, H. L. Construction of a urease-negative mutant of Proteus mirabilis: analysis of virulence in a mouse model of ascending urinary tract infection. *Infect. Immun.* **58**, 1120–1123 (1990).

230. Dattelbaum, J. D., Lockatell, C. V., Johnson, D. E. & Mobley, H. L. T. UreR, the transcriptional activator of the Proteus mirabilis urease gene cluster, is required for urease activity and virulence in experimental urinary tract infections. *Infect. Immun.* **71**, 1026–1030 (2003).

231. Schaffer, J. N., Norsworthy, A. N., Sun, T.-T. & Pearson, M. M. *Proteus mirabilis* fimbriae- and urease-dependent clusters assemble in an extracellular niche to initiate bladder stone formation. *Proc. Natl. Acad. Sci.* **113**, 4494–4499 (2016).

232. JOHNSON, D. E. *et al.* Contribution of Proteus-Mirabilis Urease To Persistence, Urolithiasis,and Acute Pyelonephritis in a Mouse Model of Ascending Urinary-TractInfection. *Infect. Immun.* **61**, 2748–2754 (1993).

233. Murphy, C. A. & Belas, R. Genomic rearrangements in the flagellin genes of Proteus mirabilis. *Mol. Microbiol.* **31**, 679–690 (1999).

234. Tahoun, A. *et al.* Host species adaptation of TLR5 signalling and flagellin recognition. *Sci. Rep.* **7**, 1–13 (2017).

235. Belas, R. & Flaherty, D. Sequence and genetic analysis of multiple flagellin-encoding genes from Proteus mirabilis. *Gene* **148**, 33–41 (1994).

236. Rather, P. N. Swarmer cell differentiation in Proteus mirabilis. *Environmental Microbiology* vol. 7 1065–1073 (2005).

237. C, A., N, C., PL, J. & C, H. Ability of Proteus mirabilis to invade human urothelial cells is coupled to motility and swarming differentiation. *Infect. Immun.* **60**, 4740–4746 (1992).

238. Peerbooms, P. G. H., Marian, A., Verweij, J. J. & Maclaren, D. M. Vero Cell Invasiveness of Proteus mirabilis. *Infect. Immun.* **43**, 1068–1071 (1984).

239. X, L., DA, R., CV, L., DE, J. & HL, M. Repression of bacterial motility by a novel fimbrial gene product. *EMBO J.* **20**, 4854–4862 (2001).

240. Zunino, P., Piccini, C. & Legnani-Fajardo, C. Flagellate and non-flagellate Proteus mirabilis in the development of experimental urinary tract infection. *Microb. Pathog.* **16**, 379–385 (1994).

241. Jansen, A. M., Lockatell, V., Johnson, D. E. & Mobley, H. L. T. Mannose-Resistant Proteus-Like Fimbriae Are Produced by Most Proteus mirabilis Strains Infecting the Urinary Tract, Dictate the In Vivo Localization of Bacteria, and Contribute to Biofilm Formation. *Infect. Immun.* **72**, 7294 (2004).

242. Lai, H. C. *et al.* Association between urine pH and common uropathogens in children with urinary tract infections. *J. Microbiol. Immunol. Infect.* **54**, 290–298 (2021).

243. Armbruster, C. E. & Mobley, H. L. T. Merging mythology and morphology: The multifaceted lifestyle of Proteus mirabilis. *Nature Reviews Microbiology* vol. 10 743–754 (2012).

244. M, T. *et al.* A FimH inhibitor prevents acute bladder infection and treats chronic cystitis caused by multidrug-resistant uropathogenic Escherichia coli ST131. *J. Infect. Dis.* **208**, 921–928 (2013).

245. Feneley, R. C. L., Hopley, I. B. & Wells, P. N. T. Urinary catheters: history, current status, adverse events and research agenda. *J. Med. Eng. Technol.* **39**, 459 (2015).

246. Carey, A. J. *et al.* Urinary tract infection of mice to model human disease: Practicalities, implications and limitations. *Crit. Rev. Microbiol.* **42**, 780–799 (2016).

247. Yang, P. J., Pham, J., Choo, J. & Hu, D. L. Duration of urination does not change with body size. *Proc. Natl. Acad. Sci.* **111**, 11932–11937 (2014).

248. Lee, G., Romih, R. & Zupančič, D. Cystitis: From urothelial cell biology to clinical applications. *Biomed Res. Int.* **2014**, (2014).

249. Romih, R., Jezernik, K. & Mašera, A. Uroplakins and cytokeratins in the regenerating rat urothelium after sodium saccharin treatment. *Histochem. Cell Biol. 1998 1093* **109**, 263–269 (1998).

250. D, Z. *et al.* A toll-like receptor that prevents infection by uropathogenic bacteria. *Science* **303**, 1522–1526 (2004).

251. Ito, H. *et al.* Muro-Neuro-Urodynamics; a Review of the Functional Assessment of Mouse Lower Urinary Tract Function. *Front. Physiol.* **8**, 49 (2017).

252. Nielsen, T. K. *et al.* A Porcine Model for Urinary Tract Infection. *Front. Microbiol.* **10**, 2564 (2019).

253. Niazi, S., Vishnupad, K. S. & Veng-Pedersen, P. Absorption and disposition characteristics of nitrofurantoin in dogs. *Biopharmaceutics & Drug Disposition* vol. 4 213–223 (1983).

254. Hammond, A., Stuijfzand, B., Avison, M. B. & Hay, A. D. Antimicrobial resistance associations with national primary care antibiotic stewardship policy: Primary care-based, multilevel analytic study. *PLoS One* **15**, e0232903 (2020).

255. Masters, P. A., O'Bryan, T. A., Zurlo, J., Miller, D. Q. & Joshi, N. Trimethoprim-Sulfamethoxazole Revisited. *Arch. Intern. Med.* **163**, 402–410 (2003).

256. Sköld, O. Sulfonamides and trimethoprim. *http://dx.doi.org/10.1586/eri.09.117* **8**, 1–6 (2014).

257. Huttner, A. *et al.* Nitrofurantoin revisited: a systematic review and meta-analysis of controlled trials. *J. Antimicrob. Chemother.* **70**, 2456–2464 (2015).

258. Singh, N. *et al.* Kidney function and the use of nitrofurantoin to treat urinary tract infections in older women. *CMAJ* **187**, 648–656 (2015).

259. Mc Osker, C. C. & Fitzpatrick, P. M. Nitrofurantoin: Mechanism of action and implications for resistance development in common uropathogens. *J. Antimicrob.*

*Chemother.* **33**, 23–30 (1994).

260. Conklin, J. D. The pharmacokinetics of nitrofurantoin and its related bioavailability. *Antibiot. Chemother.* **25**, 233–252 (1978).

261. FDA. Macrobid® (nitrofurantoin monohydrate/macrocrystals) Capsules product information. *Https://Www.Accessdata.Fda.Gov/Drugsatfda_Docs/Label/2003/20064Slr014_Macrobid_Lbl.Pdf* vol. 2009 1–11 (2009).

262. Joint Formulary Committee. *British National Formulary*. (BMJ Group and Pharmaceutical Press, 2021).

263. ten Doesschate, T., Groenwold, R. H. H., Bonten, M. J. M. & van Werkhoven, C. H. Effectiveness of extended- versus normal-release nitrofurantoin for cystitis: an instrumental variable analysis. *J. Antimicrob. Chemother.* **74**, 3337–3343 (2019).

264. Wijma, R. A., Huttner, A., Koch, B. C. P. P., Mouton, J. W. & Muller, A. E. Review of the pharmacokinetic properties of nitrofurantoin and nitroxoline. *J. Antimicrob. Chemother.* **73**, 2916–2926 (2018).

265. Watari, N., Aizawa, K. & Kaneniwa, N. Dose- and time-dependent kinetics of the renal excretion of nitrofurantoin in the rabbit. *J. Pharm. Sci.* **74**, 165–170 (1985).

266. Watari, N., Funaki, T., Aizawa, K. & Kaneniwa, N. Nonlinear assessment of nitrofurantoin bioavailability in rabbits. *J. Pharmacokinet. Biopharm.* **11**, 529–545 (1983).

267. Huttner, A. *et al.* The pharmacokinetics of nitrofurantoin in healthy female volunteers: a randomized crossover study. *J. Antimicrob. Chemother.* **74**, 1656–1661 (2019).

268. Fransen, F., Melchers, M. J. B., Meletiadis, J. & Mouton, J. W. Pharmacodynamics and differential activity of nitrofurantoin against ESBL-positive pathogens involved in urinary tract infections. *J. Antimicrob. Chemother.* **71**, 2883–2889 (2016).

269. Komp Lindgren, P., Klockars, O., Malmberg, C. & Cars, O. Pharmacodynamic studies of nitrofurantoin against common uropathogens. *J. Antimicrob. Chemother.* **70**, 1076–1082 (2014).

270. Fransen, F., Melchers, M. J. B., Lagarde, C. M. C., Meletiadis, J. & Mouton, J. W. Pharmacodynamics of nitrofurantoin at different pH levels against pathogens involved in urinary tract infections. doi:10.1093/jac/dkx313.

271. Sandegren, L., Lindqvist, A., Kahlmeter, G. & Andersson, D. I. Nitrofurantoin resistance mechanism and fitness cost in Escherichia coli. doi:10.1093/jac/dkn222.

272. Mccalla, D. R., Kaiser, C. & Green2, M. H. L. *Genetics of Nitrofurazone Resistance in Escherichia coli*. vol. 133 http://jb.asm.org/ (1978).

273. Whiteway, J. *et al. Oxygen-Insensitive Nitroreductases: Analysis of the Roles of nfsA and nfsB in Development of Resistance to 5-Nitrofuran Derivatives in Escherichia coli*. *JOURNAL OF BACTERIOLOGY* vol. 180 http://jb.asm.org/ (1998).

274. Zhang, X. *et al.* Unravelling mechanisms of nitrofurantoin resistance and epidemiological characteristics among Escherichia coli clinical isolates. *Int. J. Antimicrob. Agents* **52**, 226–232 (2018).

275. Li, J. *et al.* The nature and epidemiology of OqxAB, a multidrug efflux pump. *Antimicrobial Resistance and Infection Control* vol. 8 1–13 (2019).

276. Ho, P.-L. *et al.* Plasmid-Mediated OqxAB Is an Important Mechanism for Nitrofurantoin Resistance in Escherichia coli. (2015) doi:10.1128/AAC.02156-15.

277. Mörtl, S. *et al.* Biosynthesis of riboflavin: Lumazine synthase of Escherichia coli. *J. Biol. Chem.* **271**, 33201–33207 (1996).

278. Fischer, M. & Bacher, A. Biosynthesis of vitamin B2: Structure and mechanism of riboflavin synthase. *Archives of Biochemistry and Biophysics* vol. 474 252–265 (2008).

279. Vogl, C. *et al.* Characterization of riboflavin (vitamin B2) transport proteins from Bacillus subtilis and Corynebacterium glutamicum. *J. Bacteriol.* **189**, 7367–7375 (2007).

280. Vervoort, J. *et al.* An In Vitro Deletion in ribE Encoding Lumazine Synthase Contributes to Nitrofurantoin Resistance in Escherichia coli. (2014) doi:10.1128/AAC.03952-14.

281. Vervoort, J. *et al.* An In Vitro Deletion in ribE Encoding Lumazine Synthase Contributes to Nitrofurantoin Resistance in Escherichia coli. (2014) doi:10.1128/AAC.03952-14.

282. Glannakopoulos, X. *et al.* Human bladder urine oxygen content: Implications for urinary tract diseases. *Int. Urol. Nephrol.* **29**, 393–401 (1997).

283. Cantón, R., González-Alba, J. M. & Galán, J. C. CTX-M enzymes: Origin and diffusion. *Frontiers in Microbiology* vol. 3 (2012).

284. Edowik, Y., Caspari, T. & Williams, H. M. The Amino Acid Changes T55A, A273P and R277C in the Beta-Lactamase CTX-M-14 Render E. coli Resistant to the Antibiotic Nitrofurantoin, a First-Line Treatment of Urinary Tract Infections. *Microorganisms* **8**, 1983 (2020).

285. Thai, T., Salisbury, B. H. & Zito, P. M. Ciprofloxacin. *StatPearls* (2020).

286. Hawkey, P. M. Mechanisms of quinolone action and microbial response. *J. Antimicrob. Chemother.* **51**, 29–35 (2003).

287. Varughese, L. R. *et al.* Analytical profiling of mutations in quinolone resistance determining region of gyrA gene among UPEC. (2018) doi:10.1371/journal.pone.0190729.

288. Bagel, S., Hu¨llen, V., Hu¨llen, H., Wiedemann, B. & Heisig, P. *Impact of gyrA and parC Mutations on Quinolone Resistance, Doubling Time, and Supercoiling Degree of Escherichia coli*. vol. 43 http://aac.asm.org/ (1999).

289. Heisig, P. *Genetic Evidence for a Role of parC Mutations in Development of High-Level Fluoroquinolone Resistance in Escherichia coli Downloaded from*. *ANTIMICROBIAL AGENTS AND CHEMOTHERAPY* vol. 40 http://aac.asm.org/ (1996).

290. Bui, T. & Preuss, C. V. Cephalosporins - StatPearls. *NCBI* https://www.ncbi.nlm.nih.gov/books/NBK551517/ (2020).

291. Rice, L. B. Mechanisms of Resistance and Clinical Relevance of Resistance to β-Lactams, Glycopeptides, and Fluoroquinolones. *Mayo Clin. Proc.* **87**, 198 (2012).

292. P, B., JR, T., C, W. & M, F. Cephalexin and cephaloglycin activity in vitro and absorption and urinary excretion of single oral doses in normal young adults. *Appl. Microbiol.* **16**, 1684–1694 (1968).

293. Merck. Sanger Sequencing Steps & Method. https://www.sigmaaldrich.com/GB/en/technical-documents/protocol/genomics/sequencing/sanger-sequencing.

294. Liao, X. *et al.* Current challenges and solutions of de novo assembly. *Quantitative Biology* vol. 7 90–109 (2019).

295. Mukherjee, S. *et al.* 1,003 reference genomes of bacterial and archaeal isolates expand coverage of the tree of life. *Nat. Biotechnol. 2017 357* **35**, 676–683 (2017).

296. Minoche, A. E., Dohm, J. C. & Himmelbauer, H. Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and Genome Analyzer systems. *Genome Biol.* **12**, 1–15 (2011).

297. Rhoads, A. & Au, K. F. PacBio Sequencing and Its Applications. *Genomics, Proteomics and Bioinformatics* vol. 13 278–289 (2015).

298. Jain, M., Olsen, H. E., Paten, B. & Akeson, M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. *Genome Biol.* **17**, 1–11 (2016).

299. Ardui, S., Ameur, A., Vermeesch, J. R. & Hestand, M. S. Single molecule real-time (SMRT) sequencing comes of age: Applications and utilities for medical diagnostics. *Nucleic Acids Research* vol. 46 2159–2168 (2018).

300. Rice, E. S. & Green, R. E. New Approaches for Genome Assembly and Scaffolding. *Annual Review of Animal Biosciences* vol. 7 17–40 (2019).

301. Nagarajan, N. & Pop, M. Sequence assembly demystified. *Nature Reviews Genetics* vol. 14 157–167 (2013).

302. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595 (2017).

303. Zimin, A. V. *et al.* The MaSuRCA genome assembler. *Bioinformatics* **29**, 2669–2677 (2013).

304. Chen, Z., Erickson, D. L. & Meng, J. Benchmarking hybrid assembly approaches for genomic analyses of bacterial pathogens using Illumina and Oxford Nanopore sequencing. *BMC Genomics* **21**, 631 (2020).

305. Ricker, N., Qian, H. & Fulthorpe, R. R. The limitations of draft assemblies for understanding prokaryotic adaptation and evolution. *Genomics* **100**, 167–175 (2012).

306. Kolmogorov, M., Raney, B., Paten, B. & Pham, S. Ragout--a reference-assisted assembly tool for bacterial genomes. *Bioinformatics* **30**, i302–i309 (2014).

307. Seemann, T. Prokka: Rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).

308. Jia, B. *et al.* CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. **45**, (2017).

309. Selander, R. K. *et al.* Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Applied and Environmental Microbiology* vol. 51 873–884 (1986).

310. Maiden, M. C. J. *et al.* Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 3140–3145 (1998).

311. Abram, K. *et al.* Mash-based analyses of Escherichia coli genomes reveal 14 distinct phylogroups. *Commun. Biol. 2021 41* **4**, 1–12 (2021).

312. Applied Maths. *Multi-Locus Sequence Typing (MLST)*. http://download.applied-maths.com/sites/default/files/applications/pdf/MLST.pdf (2018).

313. Clermont, O., Christenson, J. K., Denamur, E. & Gordon, D. M. The Clermont Escherichia coli phylo-typing method revisited: Improvement of specificity and detection of new phylo-groups. *Environ. Microbiol. Rep.* **5**, 58–65 (2013).

314. Sim, M. *et al.* Growth rate control of flagellar assembly in Escherichia coli strain

RP437 OPEN. *Nat. Publ. Gr.* (2017) doi:10.1038/srep41189.

315. Jiang, Y. *et al.* Multigene Editing in the Escherichia coli Genome via the CRISPR-Cas9 System. (2015) doi:10.1128/AEM.04023-14.

316. Hall, B. G., Acar, H., Nandipati, A. & Barlow, M. Growth Rates Made Easy. *Mol. Biol. Evol.* **31**, 232–238 (2014).

317. Mcnally, A., Cheng, L., Harris, S. R. & Corander, J. The Evolutionary Path to Extraintestinal Pathogenic, Drug-Resistant Escherichia coli Is Marked by Drastic Reduction in Detectable Recombination within the Core Genome. doi:10.1093/gbe/evt038.

318. Silva, M. *et al.* chewBBACA: A complete suite for gene-by-gene schema creation and strain identification. *Microb. genomics* **4**, (2018).

319. Saint, S. *et al.* A Program to Prevent Catheter-Associated Urinary Tract Infection in Acute Care. *N. Engl. J. Med.* **374**, 2111–2119 (2016).

320. Brennand, C. *et al.* Antibiotic treatment for intermittent bladder catheterisation with once daily prophylaxis (the AnTIC study): Study protocol for a randomised controlled trial. *Trials* **17**, 276 (2016).

321. Thomas-White, K., Brady, M., Wolfe, A. J. & Mueller, E. R. The Bladder Is Not Sterile: History and Current Discoveries on the Urinary Microbiome. *Current Bladder Dysfunction Reports* vol. 11 18–24 (2016).

322. Langdon, A., Crook, N. & Dantas, G. The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome Med. 2016 81* **8**, 1–16 (2016).

323. Hsieh, J. J. Ecological fallacy | epidemiology | Britannica. https://www.britannica.com/science/ecological-fallacy.

324. Magiorakos, A.-P. *et al.* Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. *Clin. Microbiol. Infect.* **18**, 268–281 (2012).

325. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

326. Tørresen, O. K. *et al.* Tandem repeats lead to sequence assembly errors and impose multi-level challenges for genome and protein databases. *Nucleic acids research* vol. 47 10994–11006 (2019).

327. Bosi, E. *et al.* MeDuSa: A multi-draft based scaffolder. *Bioinformatics* **31**, 2443–2451 (2015).

328. Wirth, T. *et al.* Sex and virulence in Escherichia coli: An evolutionary perspective. *Mol. Microbiol.* **60**, 1136–1151 (2006).

329. EnteroBase. https://enterobase.warwick.ac.uk/species/index/ecoli.

330. Kumar, S., Stecher, G., Li, M., Knyaz, C. & Tamura, K. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **35**, 1547–1549 (2018).

331. Poirel, L. *et al.* Antimicrobial Resistance in Escherichia coli. *Microbiol. Spectr.* **6**, 289–316 (2018).

332. Osei Sekyere, J. Genomic insights into nitrofurantoin resistance mechanisms and epidemiology in clinical Enterobacteriaceae. *Futur. Sci. OA* **4**, (2018).

333. Dolk, F. C. K., Pouwels, K. B., Smith, D. R. M., Robotham, J. V & Smieszek, T. Antibiotics in primary care in England: which antibiotics are prescribed and for which conditions? *J. Antimicrob. Chemother.* **73**, ii2–ii10 (2018).

334. Public Health England. *AMR local indicators: Antimicrobial Resistance*. https://fingertips.phe.org.uk/ (2021).

335. Fasugba, O., Gardner, A., Mitchell, B. G. & Mnatzaganian, G. Ciprofloxacin resistance in community-and hospital-acquired Escherichia coli urinary tract infections: a systematic review and meta-analysis of observational studies. (2015) doi:10.1186/s12879-015-1282-4.

336. Kahlmeter, G., Åhman, J. & Matuschek, E. Antimicrobial Resistance of Escherichia

coli Causing Uncomplicated Urinary Tract Infections: A European Update for 2014 and Comparison with 2000 and 2008. *Infect. Dis. Ther.* **4**, 417–423 (2015).

337. Shackley, D. C. *et al.* Variation in the prevalence of urinary catheters: A profile of National Health Service patients in England. *BMJ Open* **7**, e013842 (2017).

338. Cortese, Y. J., Wagner, V. E., Tierney, M., Devine, D. & Fogarty, A. Review of catheter-associated urinary tract infections and in vitro urinary tract models. *Journal of Healthcare Engineering* vol. 2018 (2018).

339. Holland, A. Benefits of the Insertion Tip and Closed-System Sleeve for Intermittent Catheterization. in (2012).

340. Goetz, L. L., Droste, L., Klausner, A. P. & Newman, D. K. Catheters used for intermittent catheterization. in *Clinical Application of Urologic Catheters, Devices and Products* 47–77 (Springer International Publishing, 2017). doi:10.1007/978-3-319-14821-2_2.

341. Benton, J., Chawla, J., Parry, S. & Stickler, D. Virulence factors in Escherichia coli from urinary tract infections in patients with spinal injuries. *J. Hosp. Infect.* **22**, 117–127 (1992).

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

343. Thomas-White, K. *et al.* Culturing of female bladder bacteria reveals an interconnected urogenital microbiota. *Nat. Commun.* **9**, 1–7 (2018).

344. Nienhouse, V. *et al.* Interplay between Bladder Microbiota and Urinary Antimicrobial Peptides: Mechanisms for Human Urinary Tract Infection Risk and Symptom Severity. *PLoS One* **9**, e114185 (2014).

345. Frost, I. *et al.* Cooperation, competition and antibiotic resistance in bacterial colonies. *ISME J.* **12**, 1582–1593 (2018).

346. Gomes, A. L. C., Galagan, J. E. & Segrè, D. Resource Competition May Lead to Effective Treatment of Antibiotic Resistant Infections. *PLoS One* **8**, e80775 (2013).

347. Melnyk, A. H., Wong, A. & Kassen, R. The fitness costs of antibiotic resistance mutations. *Evol. Appl.* **8**, 273–283 (2015).

348. McNulty, C. & Allison, R. Urinary tract infections: when is it appropriate to prescribe an antibiotic? | Implementing guidelines | Guidelines in Practice. https://www.guidelinesinpractice.co.uk/urology/urinary-tract-infections-when-is-it-appropriate-to-prescribe-an-antibiotic/454498.article (2019).

349. Marcusson, L. L., Frimodt-Møller, N. & Hughes, D. Interplay in the selection of fluoroquinolone resistance and bacterial fitness. *PLoS Pathog.* **5**, 1000541 (2009).

350. Meyhoff, H. H., Nordling, J., Gammelgaard, P. A. & Vejlsgaard, R. Does Antibacterial Ointment Applied to Urethral Meatus in Women Prevent Recurrent Cystitis. *Scand. J. Urol. Nephrol.* **15**, 81–83 (1981).

351. Lio, P. A. & Kaye, E. T. Topical antibacterial agents. *Infect. Dis. Clin. North Am.* **18**, 717–733 (2004).

352. Humphrey, B. *et al.* Fitness of Escherichia coli strains carrying expressed and partially silent IncN and IncP1 plasmids. *BMC Microbiol.* **12**, 1–9 (2012).

353. Zhao, G. *et al.* Structure and function of Escherichia coli RimK, an ATP-grasp fold, l-glutamyl ligase enzyme. *Proteins Struct. Funct. Bioinforma.* **81**, 1847–1854 (2013).

354. Sim, M. *et al.* Growth rate control of flagellar assembly in Escherichia coli strain RP437. *Sci. Rep.* **7**, 1–11 (2017).

355. Liu, A. *et al.* Selective Advantage of Resistant Strains at Trace Levels of Antibiotics: a Simple and Ultrasensitive Color Test for Detection of Antibiotics and Genotoxic Agents. *Antimicrob. Agents Chemother.* **55**, 1204–1210 (2011).

356. CLSI. *M100* | *Performance Standards for Antimicrobial Susceptibility Testing*. (2921).

357. Rodloff, A., Bauer, T., Ewig, S., Kujath, P. & Müller, E. Übersichtsarbeit: Sensibel, intermediär und resistent - Wirkintensität von antibiotika. *Deutsches Arzteblatt* vol. 105 657–662 (2008).

358. Wiser, M. J. & Lenski, R. E. A comparison of methods to measure fitness in

Escherichia coli. *PLoS One* **10**, (2015).

359. Andersson, D. I. & Hughes, D. Antibiotic resistance and its cost: Is it possible to reverse resistance? *Nature Reviews Microbiology* vol. 8 260–271 (2010).

360. Smith, K. P. & Kirby, J. E. The inoculum effect in the era of multidrug resistance: Minor differences in inoculum have dramatic effect on MIC Determination. *Antimicrob. Agents Chemother.* **62**, (2018).

361. Andersson, D. I. & Hughes, D. Evolution of antibiotic resistance at non-lethal drug concentrations. *Drug Resist. Updat.* **15**, 162–172 (2012).

362. Gullberg, E. *et al.* Selection of Resistant Bacteria at Very Low Antibiotic Concentrations. *PLoS Pathog.* **7**, e1002158 (2011).

363. Andersson, D. I. & Hughes, D. *Microbiological effects of sublethal levels of antibiotics*. vol. 12 465–478 (Nature Publishing Group, 2014).

364. Sandegren, L. *Selection of antibiotic resistance at very low antibiotic concentrations*. *Upsala Journal of Medical Sciences* vol. 119 103–107 (Informa Healthcare, 2014).

365. Hawn, T. R. *et al.* Toll-like receptor polymorphisms and susceptibility to urinary tract infections in adult women. *PLoS One* **4**, (2009).

366. Randhawa, A. K. *et al.* Association of Human TLR1 and TLR6 Deficiency with Altered Immune Responses to BCG Vaccination in South African Infants. *PLoS Pathog.* **7**, e1002174 (2011).

367. Hawn, T. R. *et al.* A Common Dominant TLR5 Stop Codon Polymorphism Abolishes Flagellin Signaling and Is Associated with Susceptibility to Legionnaires' Disease. *J. Exp. Med.* **198**, 1563–1572 (2003).

368. Uhlén, M. *et al.* Tissue-based map of the human proteome (TLR1). *Science (80-. ).* **347**, (2015).

369. Sundac, L. *et al.* Protein-based profiling of the immune response to uropathogenic Escherichia coli in adult patients immediately following hospital admission for acute cystitis. *Pathog. Dis.* **74**, 62 (2016).

370. Gonzalez, E. J., Arms, L. & Vizzard, M. A. The role(s) of cytokines/chemokines in urinary bladder inflammation and dysfunction. *Biomed Res. Int.* **2014**, (2014).

371. Umpiérrez, A. *et al.* Innate immune responses to Proteus mirabilis flagellin in the urinary tract. *Microbes Infect.* **15**, 688–696 (2013).

372. Smith, K. D. *et al.* Toll-like receptor 5 recognizes a conserved site on flagellin required for protofilament formation and bacterial motility. *Nat. Immunol.* **4**, 1247–1253 (2003).

373. Snyder, J. A. *et al.* Transcriptome of Uropathogenic Escherichia coli during Urinary Tract Infection. *J. Infect. Immun.* **72**, 6373–6381 (2004).

374. Schenk, M., Belisle, J. T. & Modlin, R. L. TLR2 Looks at Lipoproteins. *Immunity* **31**, 847–849 (2009).

375. Tan, R. S. T., Ho, B., Leung, B. P. & Ding, J. L. TLR Cross-talk Confers Specificity to Innate Immunity. *Int. Rev. Immunol.* **33**, 443–453 (2014).

376. Puigbò, P., Garcia-Vallvé, S. & McInerney, J. O. TOPD/FMTS: A new software to compare phylogenetic trees. *Bioinformatics* **23**, 1556–1558 (2007).

377. Annunziato, F., Romagnani, C. & Romagnani, S. The 3 major types of innate and adaptive cell-mediated effector immunity. *J. Allergy Clin. Immunol.* **135**, 626–635 (2015).

378. Bal, S. M., Golebski, K. & Spits, H. Plasticity of innate lymphoid cell subsets. *Nat. Rev. Immunol. 2020 209* **20**, 552–565 (2020).

379. Kaiko, G. E., Horvat, J. C., Beagley, K. W. & Hansbro, P. M. Immunological decision-making: how does the immune system decide to mount a helper T-cell response? *Immunology* **123**, 326–338 (2008).

380. Netea, M. G., Van Der Meer, J. W. M., Sutmuller, R. P., Adema, G. J. & Kullberg, B. J. From the Th1/Th2 paradigm towards a toll-like receptor/T-helper bias. *Antimicrobial Agents and Chemotherapy* vol. 49 3991–3996 (2005).

381. Vaughan, M. H. *et al.* The Urinary Microbiome in Postmenopausal Women with

Recurrent Urinary Tract Infections. *bioRxiv* 2020.12.21.423901 (2020) doi:10.1101/2020.12.21.423901.

382. McKinnell, J. A., Stollenwerk, N. S., Jung, C. W. & Miller, L. G. Nitrofurantoin Compares Favorably to Recommended Agents as Empirical Treatment of Uncomplicated Urinary Tract Infections in a Decision and Cost Analysis. *Mayo Clin. Proc.* **86**, 480 (2011).

383. Public Health England. *English Surveillance Programme for Antimicrobial Utilisation and Resistance (ESPAUR), Report 2018-2019*. https://webarchive.nationalarchives.gov.uk/20200806045257/https://www.gov.uk/government/publications/english-surveillance-programme-antimicrobial-utilisation-and-resistance-espaur-report (2019).

# Chapter 8 Appendices

| | | Doubling Time | vs 1646 BASE | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA T37M ΔnfsB | vs 1646 6 | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 μg/ml | 1646 BASE | 21.9 ± 0.9 | | 0.122673576 | 0.295828825 | 0.695055674 | 0.017228224 | 0.997790985 | 0.169542953 | 0.064218721 | 0.254093641 | 0.877463929 | 0.954412847 |
| | ΔnfsA | 24.2 ± 0.6 | 0.122673576 | | 0.692302026 | 0.202611486 | 0.097796129 | 0.289223067 | 0.994388341 | 0.327761849 | 0.953484397 | 0.206442909 | 0.302880895 |
| | ΔnfsA-rimK | 23.7 ± 0.9 | 0.295828825 | 0.692302026 | | 0.457678968 | 0.124716193 | 0.443226848 | 0.727454966 | 0.265391168 | 0.81059115 | 0.40482999 | 0.444109685 |
| | nfsA T37M | 22.5 ± 0.8 | 0.695055674 | 0.202611486 | 0.457678968 | | 0.023518472 | 0.781808017 | 0.268446022 | 0.093547058 | 0.377554646 | 0.841424957 | 0.755312942 |
| | ΔnfsA ΔnfsB | 25.7 ± 0.4 | 0.017228224 | 0.097796129 | 0.124716193 | 0.023518472 | | 0.090104149 | 0.200592864 | 0.985817564 | 0.28909813 | 0.037036894 | 0.105575481 |
| | ΔnfsA-rimK ΔnfsB | 21.9 ± 1.6 | 0.997790985 | 0.289223067 | 0.443226848 | 0.781808017 | 0.090104149 | | 0.316204066 | 0.141105085 | 0.375885689 | 0.906728161 | 0.96347127 |
| | nfsA T37M ΔnfsB | 24.2 ± 0.8 | 0.169542953 | 0.994388341 | 0.727454966 | 0.268446022 | 0.200592864 | 0.316204066 | | 0.381013137 | 0.954226272 | 0.253101806 | 0.325734175 |
| | 1646 6 | 25.7 ± 1.1 | 0.064218721 | 0.327761849 | 0.265391168 | 0.093547058 | 0.985817564 | 0.141105085 | 0.381013137 | | 0.415631641 | 0.096394204 | 0.152392701 |
| | ΔrimK | 24.1 ± 1.2 | 0.254093641 | 0.953484397 | 0.81059115 | 0.377554646 | 0.28909813 | 0.375885689 | 0.954226272 | 0.415631641 | | 0.338890758 | 0.379065639 |
| | ΔnfsB | 22.2 ± 1.1 | 0.877463929 | 0.206442909 | 0.40482999 | 0.841424957 | 0.037036894 | 0.906728161 | 0.253101806 | 0.096394204 | 0.338890758 | | 0.871349899 |
| | ΔrimK ΔnfsB | 21.8 ± 1.8 | 0.954412847 | 0.302880895 | 0.444109685 | 0.755312942 | 0.105575481 | 0.96347127 | 0.325734175 | 0.152392701 | 0.379065639 | 0.871349899 | |

| | | Doubling Time | vs 1646 BASE | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA T37M ΔnfsB | vs 1646 6 | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 μg/ml | 1646 BASE | 24.7 ± 0.7 | | 0.734299063 | 0.272913092 | 0.685579842 | 0.494451891 | 0.297671867 | 0.813125141 | 0.679125344 | 0.524994937 | 0.29596761 | 0.696829444 |
| | ΔnfsA | 24.3 ± 0.9 | 0.734299063 | | 0.204514293 | 0.95107419 | 0.363076681 | 0.531076564 | 0.621491467 | 0.478434261 | 0.363958133 | 0.217649041 | 0.535592366 |
| | ΔnfsA-rimK | 26.2 ± 1 | 0.272913092 | 0.204514293 | | 0.188410418 | 0.66069529 | 0.069801136 | 0.462434894 | 0.398511493 | 0.431956823 | 0.829371904 | 0.583623787 |
| | nfsA T37M | 24.2 ± 0.9 | 0.685579842 | 0.95107419 | 0.188410418 | | 0.335592435 | 0.575052955 | 0.584495672 | 0.439086191 | 0.329492328 | 0.198555282 | 0.503772466 |
| | ΔnfsA ΔnfsB | 25.6 ± 0.9 | 0.494451891 | 0.363076681 | 0.66069529 | 0.335592435 | | 0.129968301 | 0.725563201 | 0.705825415 | 0.7837265 | 0.785282386 | 0.86309003 |
| | ΔnfsA-rimK ΔnfsB | 23.4 ± 0.8 | 0.297671867 | 0.531076564 | 0.069801136 | 0.575052955 | 0.129968301 | | 0.288791491 | 0.14795183 | 0.086961126 | 0.061820153 | 0.251422608 |
| | nfsA T37M ΔnfsB | 25 ± 1.1 | 0.813125141 | 0.621491467 | 0.462434894 | 0.584495672 | 0.725563201 | 0.288791491 | | 0.947105235 | 0.847644599 | 0.536795606 | 0.882238638 |
| | 1646 6 | 25.1 ± 0.6 | 0.679125344 | 0.478434261 | 0.398511493 | 0.439086191 | 0.705825415 | 0.14795183 | 0.947105235 | | 0.847971029 | 0.455657623 | 0.906889296 |
| | ΔrimK | 25.3 ± 0.4 | 0.524994937 | 0.363958133 | 0.431956823 | 0.329492328 | 0.7837265 | 0.086961126 | 0.847644599 | 0.847971029 | | 0.497358016 | 0.98925464 |
| | ΔnfsB | 25.9 ± 0.8 | 0.29596761 | 0.217649041 | 0.829371904 | 0.198555282 | 0.785282386 | 0.061820153 | 0.536795606 | 0.455657623 | 0.497358016 | | 0.680315392 |
| | ΔrimK ΔnfsB | 25.3 ± 1.2 | 0.696829444 | 0.535592366 | 0.583623787 | 0.503772466 | 0.86309003 | 0.251422608 | 0.882238638 | 0.906889296 | 0.98925464 | 0.680315392 | |

| | | Doubling Time | vs 1646 BASE | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA T37M ΔnfsB | vs 1646 6 | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 μg/ml | 1646 BASE | 29.3 ± 1.1 | | 0.002741594 | 0.045426795 | 0.004937734 | 0.004716447 | 0.000203469 | 0.002803552 | 0.012699364 | 0.557917767 | 0.887155135 | 0.445260805 |
| | ΔnfsA | 23.6 ± 1 | 0.002741594 | | 0.308728556 | 0.710479013 | 0.264459612 | 0.461696916 | 0.351556143 | 0.601193119 | 0.016364982 | 0.001499447 | 0.00235424 |
| | ΔnfsA-rimK | 25.4 ± 1.3 | 0.045426795 | 0.308728556 | | 0.471437178 | 0.785128043 | 0.086910969 | 0.6505194 | 0.614987946 | 0.155878582 | 0.030781937 | 0.007666249 |
| | nfsA T37M | 24.1 ± 1 | 0.004937734 | 0.710479013 | 0.471437178 | | 0.48263316 | 0.238021821 | 0.619350615 | 0.848758453 | 0.028857752 | 0.002709862 | 0.00040914 |
| | ΔnfsA ΔnfsB | 25 ± 0.5 | 0.004716447 | 0.264459612 | 0.785128043 | 0.48263316 | | 0.018967596 | 0.738577958 | 0.70047188 | 0.041169153 | 0.002073664 | 0.000149969 |
| | ΔnfsA-rimK ΔnfsB | 22.6 ± 0.7 | 0.000203469 | 0.461696916 | 0.086910969 | 0.238021821 | 0.018967596 | | 0.028711967 | 0.214994594 | 0.002117574 | 8.47572E-05 | 7.26047E-06 |
| | nfsA T37M ΔnfsB | 24.7 ± 0.5 | 0.002803552 | 0.351556143 | 0.6505194 | 0.619350615 | 0.738577958 | 0.028711967 | | 0.841780951 | 0.027460448 | 0.001168919 | 7.59927E-05 |
| | 1646 6 | 24.5 ± 1.2 | 0.012699364 | 0.601193119 | 0.614987946 | 0.848758453 | 0.70047188 | 0.214994594 | 0.841780951 | | 0.056423386 | 0.007837685 | 0.001591455 |
| | ΔrimK | 28.3 ± 1.3 | 0.557917767 | 0.016364982 | 0.155878582 | 0.028857752 | 0.041169153 | 0.002117574 | 0.027460448 | 0.056423386 | | 0.463448006 | 0.19133198 |
| | ΔnfsB | 29.6 ± 1 | 0.887155135 | 0.001499447 | 0.030781937 | 0.002709862 | 0.002073664 | 8.47572E-05 | 0.001168919 | 0.007837685 | 0.463448006 | | 0.518317497 |
| | ΔrimK ΔnfsB | 30.5 ± 0.8 | 0.445260805 | 0.000235424 | 0.007666249 | 0.00040914 | 0.000149969 | 7.26047E-06 | 7.59927E-05 | 0.001591455 | 0.19133198 | 0.518317497 | |

| | | Doubling Time | vs 1646 BASE | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA T37M ΔnfsB | vs 1646 6 | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 μg/ml | 1646 BASE | 43.2 ± 1.8 | | 0.000876408 | 0.001543998 | 0.000652694 | 6.35228E-07 | 7.76752E-07 | 5.55643E-06 | 1.32271E-06 | 0.282389302 | 0.430255306 | 0.856799838 |
| | ΔnfsA | 31 ± 2 | 0.000876408 | | 0.701079663 | 0.831646762 | 0.015308103 | 0.010180429 | 0.185441115 | 0.034553222 | 5.42698E-05 | 4.1817E-05 | 0.000413868 |
| | ΔnfsA-rimK | 32.2 ± 1.9 | 0.001543998 | 0.701079663 | | 0.553374198 | 0.004448172 | 0.003148866 | 0.070229955 | 0.010754058 | 8.48999E-05 | 6.46211E-05 | 0.000715687 |
| | nfsA T37M | 30.3 ± 2.1 | 0.000652694 | 0.831646762 | 0.553374198 | | 0.029739255 | 0.019342184 | 0.296500693 | 0.06376137 | 4.35315E-05 | 3.39439E-05 | 0.000312714 |
| | ΔnfsA ΔnfsB | 24.3 ± 1 | 6.35228E-07 | 0.015308103 | 0.004448172 | 0.029739255 | | 0.61615918 | 0.046055872 | 0.545452305 | 1.07517E-08 | 1.26495E-09 | 1.58399E-07 |
| | ΔnfsA-rimK ΔnfsB | 23.4 ± 1.3 | 7.76752E-07 | 0.010180429 | 0.003148866 | 0.019342184 | 0.61615918 | | 0.028071249 | 0.308988843 | 2.04317E-08 | 4.09191E-09 | 2.27655E-07 |
| | nfsA T37M ΔnfsB | 27.7 ± 1 | 5.55643E-06 | 0.185441115 | 0.070229955 | 0.296500693 | 0.046055872 | 0.028071249 | | 0.150906778 | 7.99286E-08 | 1.04208E-08 | 1.3943E-06 |
| | 1646 6 | 25.3 ± 1.1 | 1.32271E-06 | 0.034553222 | 0.010754058 | 0.06376137 | 0.545452305 | 0.308988843 | 0.150906778 | | 2.30886E-08 | 3.12088E-09 | 3.40098E-07 |
| | ΔrimK | 45.9 ± 1.4 | 0.282389302 | 5.42698E-05 | 8.48999E-05 | 4.35315E-05 | 1.07517E-08 | 2.04317E-08 | 7.99286E-08 | 2.30886E-08 | | 0.602480633 | 0.337119786 |
| | ΔnfsB | 45 ± 0.9 | 0.430255306 | 4.1817E-05 | 6.46211E-05 | 3.39439E-05 | 1.26495E-09 | 4.09191E-09 | 1.04208E-08 | 3.12088E-09 | 0.602480633 | | 0.52421171 |
| | ΔrimK ΔnfsB | 43.7 ± 1.6 | 0.856799838 | 0.000413868 | 0.000715687 | 0.000312714 | 1.58399E-07 | 2.27655E-07 | 1.3943E-06 | 3.40098E-07 | 0.337119786 | 0.52421171 | |

| | | Doubling Time | vs 1646 BASE | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA T37M ΔnfsB | vs 1646 6 | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 16 μg/ml | 1646 BASE | | | | | | | | | | | | |
| | ΔnfsA | 60.1 ± 1.6 | | | 0.151620043 | 0.123103515 | 1.53902E-11 | 1.11302E-10 | 6.26291E-11 | 1.51803E-10 | | | |
| | ΔnfsA-rimK | 64.7 ± 2.4 | | 0.151620043 | | 0.016764125 | 5.2478E-10 | 1.60228E-09 | 1.41603E-09 | 1.92391E-09 | | | |
| | nfsA T37M | 55.4 ± 2.2 | | 0.123103515 | 0.016764125 | | 7.43976E-09 | 2.78682E-08 | 2.58996E-08 | 3.45247E-08 | | | |
| | ΔnfsA ΔnfsB | 26.6 ± 0.3 | | 1.53902E-11 | 5.2478E-10 | 7.43976E-09 | | 0.261663315 | 0.006554513 | 0.225709811 | | | |
| | ΔnfsA-rimK ΔnfsB | 27.7 ± 0.9 | | 1.11302E-10 | 1.60228E-09 | 2.78682E-08 | 0.261663315 | | 0.381548648 | 0.89665954 | | | |
| | nfsA T37M ΔnfsB | 28.7 ± 0.5 | | 6.26291E-11 | 1.41603E-09 | 2.58996E-08 | 0.006554513 | 0.381548648 | | 0.494671857 | | | |
| | 1646 6 | 27.9 ± 0.9 | | 1.51803E-10 | 1.92391E-09 | 3.45247E-08 | 0.225709811 | 0.89665954 | 0.494671857 | | | | |
| | ΔrimK | | | | | | | | | | | | |
| | ΔnfsB | | | | | | | | | | | | |
| | ΔrimK ΔnfsB | | | | | | | | | | | | |

| | | Doubling Time | vs 1646 BASE | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA T37M ΔnfsB | vs 1646 6 | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 32 μg/ml | 1646 BASE | | | | | | | | | | | | |
| | ΔnfsA | | | | | | | | | | | | |
| | ΔnfsA-rimK | | | | | | | | | | | | |
| | nfsA T37M | | | | | | | | | | | | |
| | ΔnfsA ΔnfsB | 34.2 ± 1.3 | | | | | | 0.78078804 | 0.058634033 | 0.577540826 | | | |
| | ΔnfsA-rimK ΔnfsB | 33.7 ± 0.8 | | | | | 0.78078804 | | 0.010432082 | 0.312831199 | | | |
| | nfsA T37M ΔnfsB | 37.6 ± 0.9 | | | | | 0.058634033 | 0.010432082 | | 0.103209639 | | | |
| | 1646 6 | 35.2 ± 1 | | | | | 0.577540826 | 0.312831199 | 0.103209639 | | | | |
| | ΔrimK | | | | | | | | | | | | |
| | ΔnfsB | | | | | | | | | | | | |
| | ΔrimK ΔnfsB | | | | | | | | | | | | |

| | | Doubling Time | vs 1646 BASE | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA T37M ΔnfsB | vs 1646 6 | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 64 μg/ml | 1646 BASE | | | | | | | | | | | | |
| | ΔnfsA | | | | | | | | | | | | |
| | ΔnfsA-rimK | | | | | | | | | | | | |
| | nfsA T37M | | | | | | | | | | | | |
| | ΔnfsA ΔnfsB | 90.5 ± 5.9 | | | | | | 0.739931013 | 0.007375169 | 0.002021133 | | | |
| | ΔnfsA-rimK ΔnfsB | 88 ± 3.4 | | | | | 0.739931013 | | 0.002904266 | 0.000773134 | | | |
| | nfsA T37M ΔnfsB | 131.6 ± 10.8 | | | | | 0.007375169 | 0.002904266 | | 0.544604407 | | | |
| | 1646 6 | 142 ± 11.3 | | | | | 0.002021133 | 0.000773134 | 0.544604407 | | | | |
| | ΔrimK | | | | | | | | | | | | |
| | ΔnfsB | | | | | | | | | | | | |
| | ΔrimK ΔnfsB | | | | | | | | | | | | |

**Appendix A Table showing the doubling time of each tested isolate (1646 BASE and its *nfsA/nfsB* mutants) and the relevant significance data when the isolates were compared against other isolates at concentrations of 0, 2, 4, 8, 16, 32, and 64 μg/ml nitrofurantoin.** Adjusted p-value= 0.0009 (9E-04).

| | | Doubling Time | vs W3110 | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA_T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA_T37M ΔnfsB | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0 μg/ml | W3110 | 34.9 ± 0.8 | | 0.351212208 | 0.349061866 | 0.079649323 | 0.142612222 | 0.157329587 | 0.68663919 | 0.512968709 | 0.388460425 | 0.739823526 |
| | ΔnfsA | 32.7 ± 1.7 | 0.351212208 | | 0.879411643 | 0.311806125 | 0.582684291 | 0.950406238 | 0.428076193 | 0.665633582 | 0.777734119 | 0.35626195 |
| | ΔnfsA-rimK | 33.1 ± 1.3 | 0.349061866 | 0.879411643 | | 0.23277371 | 0.456383187 | 0.784990354 | 0.438115888 | 0.74606664 | 0.693772055 | 0.377009691 |
| | nfsA_T37M | 29 ± 2.3 | 0.079649323 | 0.311806125 | 0.23277371 | | 0.586027963 | 0.268806588 | 0.091009579 | 0.159482433 | 0.602330106 | 0.105621076 |
| | ΔnfsA ΔnfsB | 31 ± 1.8 | 0.142612222 | 0.582684291 | 0.456383187 | 0.586027963 | | 0.546873764 | 0.168248892 | 0.310836115 | 0.910226189 | 0.183776826 |
| | ΔnfsA-rimK ΔnfsB | 32.5 ± 1 | 0.157329587 | 0.950406238 | 0.784990354 | 0.268806588 | 0.546873764 | | 0.186010257 | 0.508115956 | 0.787262475 | 0.264209103 |
| | nfsA_T37M ΔnfsB | 34.4 ± 0.5 | 0.68663919 | 0.428076193 | 0.438115888 | 0.091009579 | 0.168248892 | 0.186010257 | | 0.660442663 | 0.442522982 | 0.603643278 |
| | ΔrimK | 33.8 ± 1.2 | 0.512968709 | 0.665633582 | 0.74606664 | 0.159482433 | 0.310836115 | 0.508115956 | 0.660442663 | | 0.568493287 | 0.489102962 |
| | ΔnfsB | 31.5 ± 3.1 | 0.388460425 | 0.777734119 | 0.693772055 | 0.602330106 | 0.910226189 | 0.787262475 | 0.442522982 | 0.568493287 | | 0.353789908 |
| | ΔrimK ΔnfsB | 35.8 ± 2.1 | 0.739823526 | 0.35626195 | 0.377009691 | 0.105621076 | 0.183776826 | 0.264209103 | 0.603643278 | 0.489102962 | 0.353789908 | |

| | | Doubling Time | vs W3110 | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA_T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA_T37M ΔnfsB | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 μg/ml | W3110 | 37.8 ± 1 | | 0.184249862 | 0.145621484 | 0.235849878 | 0.016411644 | 0.010714833 | 0.945796741 | 0.307617548 | 0.193983566 | 0.778142781 |
| | ΔnfsA | 35.4 ± 1.2 | 0.184249862 | | 0.963163448 | 0.820650999 | 0.356529237 | 0.242024045 | 0.257885487 | 0.0653479 | 0.82138641 | 0.357752747 |
| | ΔnfsA-rimK | 35.3 ± 1.1 | 0.145621484 | 0.963163448 | | 0.771510632 | 0.345978331 | 0.228242493 | 0.219510699 | 0.054706145 | 0.843714817 | 0.314967677 |
| | nfsA_T37M | 35.8 ± 1.1 | 0.235849878 | 0.820650999 | 0.771510632 | | 0.212790129 | 0.137201717 | 0.324902271 | 0.079422212 | 0.677121749 | 0.4468072 |
| | ΔnfsA ΔnfsB | 33.9 ± 0.8 | 0.016411644 | 0.356529237 | 0.345978331 | 0.212790129 | | 0.71591314 | 0.041069193 | 0.013302828 | 0.633456812 | 0.070274531 |
| | ΔnfsA-rimK ΔnfsB | 33.4 ± 0.9 | 0.010714833 | 0.242024045 | 0.228242493 | 0.137201717 | 0.71591314 | | 0.027268378 | 0.009608784 | 0.486672148 | 0.046872635 |
| | nfsA_T37M ΔnfsB | 37.7 ± 1.3 | 0.945796741 | 0.257885487 | 0.219510699 | 0.324902271 | 0.041069193 | 0.027268378 | | 0.3146931 | 0.247057737 | 0.84556859 |
| | ΔrimK | 40.3 ± 1.9 | 0.307617548 | 0.0653479 | 0.054706145 | 0.079422212 | 0.013302828 | 0.009608784 | 0.3146931 | | 0.06992795 | 0.25337646 |
| | ΔnfsB | 34.9 ± 1.7 | 0.193983566 | 0.82138641 | 0.843714817 | 0.677121749 | 0.633456812 | 0.486672148 | 0.247057737 | 0.06992795 | | 0.325670997 |
| | ΔrimK ΔnfsB | 37.3 ± 1.4 | 0.778142781 | 0.357752747 | 0.314967677 | 0.4468072 | 0.070274531 | 0.046872635 | 0.84556859 | 0.25337646 | 0.325670997 | |

| | | Doubling Time | vs W3110 | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA_T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA_T37M ΔnfsB | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 4 μg/ml | W3110 | 37.7 ± 1.2 | | 0.961514181 | 0.233065711 | 0.567594552 | 0.554730291 | 0.163061251 | 0.838394644 | 0.017524601 | 0.079168626 | 0.398383489 |
| | ΔnfsA | 37.5 ± 1.8 | 0.961514181 | | 0.294751272 | 0.612693468 | 0.629175848 | 0.23682252 | 0.82832652 | 0.032327579 | 0.134202897 | 0.435937218 |
| | ΔnfsA-rimK | 40.4 ± 1.6 | 0.233065711 | 0.294751272 | | 0.479025606 | 0.153320314 | 0.034217508 | 0.438611583 | 0.167208306 | 0.648327995 | 0.868505086 |
| | nfsA_T37M | 38.8 ± 1.3 | 0.567594552 | 0.612693468 | 0.479025606 | | 0.324308136 | 0.077230485 | 0.819181534 | 0.043113026 | 0.208499054 | 0.670118543 |
| | ΔnfsA ΔnfsB | 36 ± 2.1 | 0.554730291 | 0.629175848 | 0.153320314 | 0.324308136 | | 0.516999643 | 0.508949622 | 0.018071896 | 0.067977629 | 0.245858442 |
| | ΔnfsA-rimK ΔnfsB | 34 ± 2 | 0.163061251 | 0.23682252 | 0.034217508 | 0.077230485 | 0.516999643 | | 0.185312145 | 0.003609932 | 0.011798864 | 0.073164637 |
| | nfsA_T37M ΔnfsB | 38.2 ± 2 | 0.838394644 | 0.82832652 | 0.438611583 | 0.819181534 | 0.508949622 | 0.185312145 | | 0.05913669 | 0.23224234 | 0.584585943 |
| | ΔrimK | 44.4 ± 2 | 0.017524601 | 0.032327579 | 0.167208306 | 0.043113026 | 0.018071896 | 0.003609932 | 0.05913669 | | 0.279716112 | 0.167563866 |
| | ΔnfsB | 41.5 ± 1.4 | 0.079168626 | 0.134202897 | 0.648327995 | 0.208499054 | 0.067977629 | 0.011798864 | 0.23224234 | 0.279716112 | | 0.573147969 |
| | ΔrimK ΔnfsB | 39.9 ± 2.1 | 0.398383489 | 0.435937218 | 0.868505086 | 0.670118543 | 0.245858442 | 0.073164637 | 0.584585943 | 0.167563866 | 0.573147969 | |

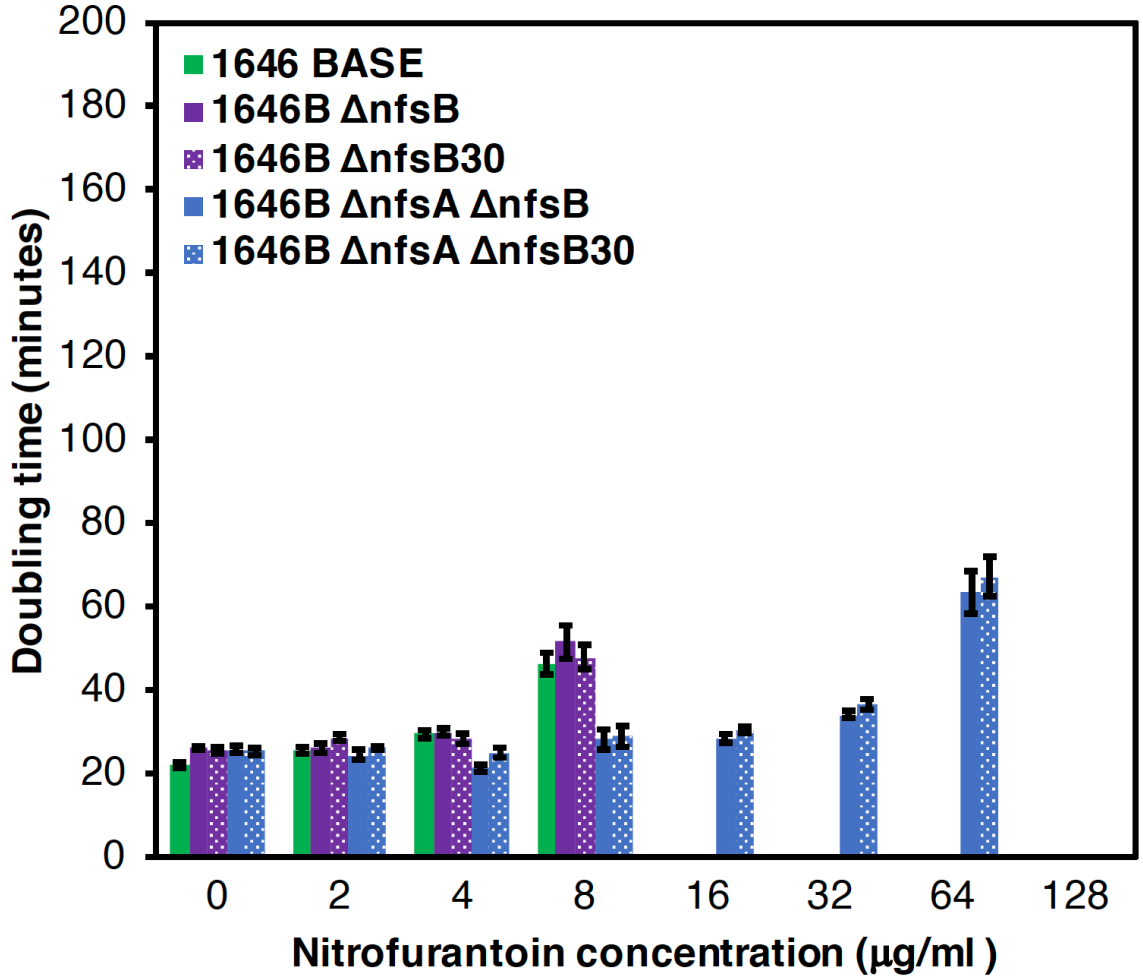
| | | Doubling Time | vs W3110 | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA_T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA_T37M ΔnfsB | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 8 μg/ml | W3110 | 74.7 ± 2.9 | | 4.48468E-05 | 1.60773E-05 | 5.55432E-05 | 5.63083E-07 | 3.78524E-08 | 1.42153E-07 | 0.487292761 | 0.816852493 | 0.622781063 |
| | ΔnfsA | 50.9 ± 2.6 | 4.48468E-05 | | 0.904274947 | 0.780770182 | 0.072370279 | 0.00088898 | 0.034669076 | 0.000665407 | 0.000139467 | 0.001429618 |
| | ΔnfsA-rimK | 50.4 ± 2.1 | 1.60773E-05 | 0.904274947 | | 0.667304134 | 0.056992605 | 0.000360307 | 0.020139839 | 0.000453967 | 5.90164E-05 | 0.00084012 |
| | nfsA_T37M | 51.9 ± 2.4 | 5.55432E-05 | 0.780770182 | 0.667304134 | | 0.032777771 | 0.00032587 | 0.012690302 | 0.000844163 | 0.000179296 | 0.001906296 |
| | ΔnfsA ΔnfsB | 44.6 ± 1.6 | 5.63083E-07 | 0.072370279 | 0.056992605 | 0.032777771 | | 0.009153396 | 0.821237145 | 6.40318E-05 | 2.55517E-06 | 6.3837E-05 |
| | ΔnfsA-rimK ΔnfsB | 37.4 ± 1.6 | 3.78524E-08 | 0.00088898 | 0.000360307 | 0.00032587 | 0.009153396 | | 0.002789196 | 8.83209E-06 | 1.74844E-07 | 5.22513E-06 |
| | nfsA_T37M ΔnfsB | 44.1 ± 0.8 | 1.42153E-07 | 0.034669076 | 0.020139839 | 0.012690302 | 0.821237145 | 0.002789196 | | 4.20881E-05 | 8.3836E-07 | 3.32539E-05 |
| | ΔrimK | 79.5 ± 5.6 | 0.487292761 | 0.000665407 | 0.000453967 | 0.000844163 | 6.40318E-05 | 8.83209E-06 | 4.20881E-05 | | 0.408803873 | 0.332733881 |
| | ΔnfsB | 73.7 ± 3.2 | 0.816852493 | 0.000139467 | 5.90164E-05 | 0.000179296 | 2.55517E-06 | 1.74844E-07 | 8.3836E-07 | 0.408803873 | | 0.773205168 |
| | ΔrimK ΔnfsB | 72 ± 4.3 | 0.622781063 | 0.001429618 | 0.00084012 | 0.001906296 | 6.3837E-05 | 5.22513E-06 | 3.32539E-05 | 0.332733881 | 0.773205168 | |

| | | Doubling Time | vs W3110 | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA_T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA_T37M ΔnfsB | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 16 μg/ml | W3110 | | | | | | | | | | | |
| | ΔnfsA | | | | | | | | | | | |
| | ΔnfsA-rimK | | | | | | | | | | | |
| | nfsA_T37M | | | | | | | | | | | |
| | ΔnfsA ΔnfsB | 39 ± 2.7 | | | | | | 0.106454407 | 0.705003298 | | | |
| | ΔnfsA-rimK ΔnfsB | 32.8 ± 2.1 | | | | | 0.106454407 | | 0.012919103 | | | |
| | nfsA_T37M ΔnfsB | 40.3 ± 1.3 | | | | | 0.705003298 | 0.012919103 | | | | |
| | ΔrimK | | | | | | | | | | | |
| | ΔnfsB | | | | | | | | | | | |
| | ΔrimK ΔnfsB | | | | | | | | | | | |

| | | Doubling Time | vs W3110 | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA_T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA_T37M ΔnfsB | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 32 μg/ml | W3110 | | | | | | | | | | | |
| | ΔnfsA | | | | | | | | | | | |
| | ΔnfsA-rimK | | | | | | | | | | | |
| | nfsA_T37M | | | | | | | | | | | |
| | ΔnfsA ΔnfsB | 54.4 ± 1.9 | | | | | | 0.18103934 | 0.396558017 | | | |
| | ΔnfsA-rimK ΔnfsB | 50.8 ± 1.4 | | | | | 0.18103934 | | 0.0196895 | | | |
| | nfsA_T37M ΔnfsB | 56.7 ± 1.5 | | | | | 0.396558017 | 0.0196895 | | | | |
| | ΔrimK | | | | | | | | | | | |
| | ΔnfsB | | | | | | | | | | | |
| | ΔrimK ΔnfsB | | | | | | | | | | | |

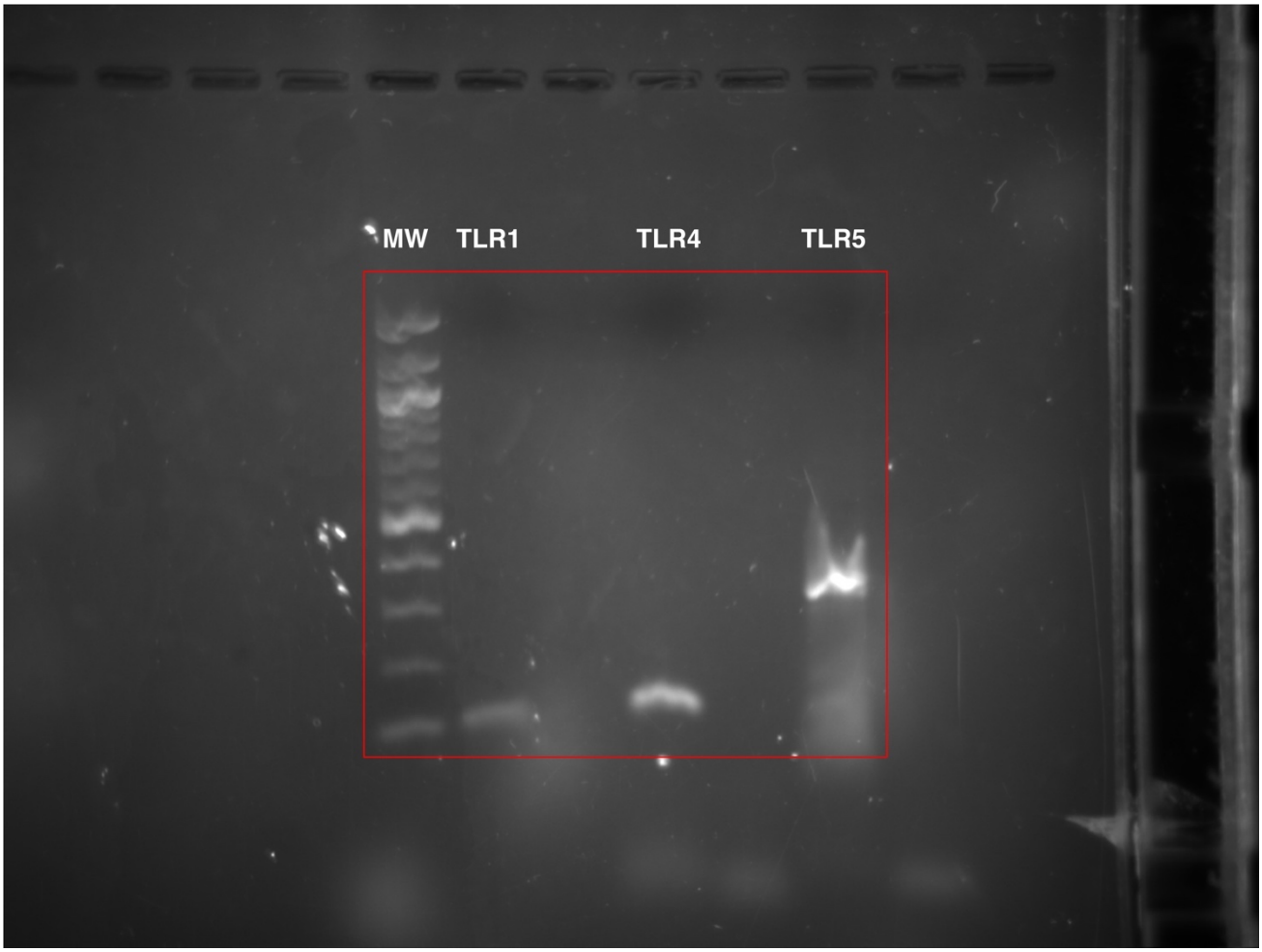
| | | Doubling Time | vs W3110 | vs ΔnfsA | vs ΔnfsA-rimK | vs nfsA_T37M | vs ΔnfsA ΔnfsB | vs ΔnfsA-rimK ΔnfsB | vs nfsA_T37M ΔnfsB | vs ΔrimK | vs ΔnfsB | vs ΔrimK ΔnfsB |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 64 μg/ml | W3110 | | | | | | | | | | | |
| | ΔnfsA | | | | | | | | | | | |
| | ΔnfsA-rimK | | | | | | | | | | | |
| | nfsA_T37M | | | | | | | | | | | |
| | ΔnfsA ΔnfsB | 117.2 ± 5.8 | | | | | | 0.068217953 | | | | |
| | ΔnfsA-rimK ΔnfsB | 97 ± 7.7 | | | | | 0.068217953 | | | | | |
| | nfsA_T37M ΔnfsB | | | | | | | | | | | |
| | ΔrimK | | | | | | | | | | | |
| | ΔnfsB | | | | | | | | | | | |
| | ΔrimK ΔnfsB | | | | | | | | | | | |

**Appendix B Table showing the doubling time of each tested isolate (W3110 and its *nfsA/nfsB* mutants) and the relevant significance data when the isolates were compared against other isolates at concentrations of 0,2,4,8,16,32, and 64 μg/ml nitrofurantoin.** Adjusted p-value= 0.001 (1E-03).

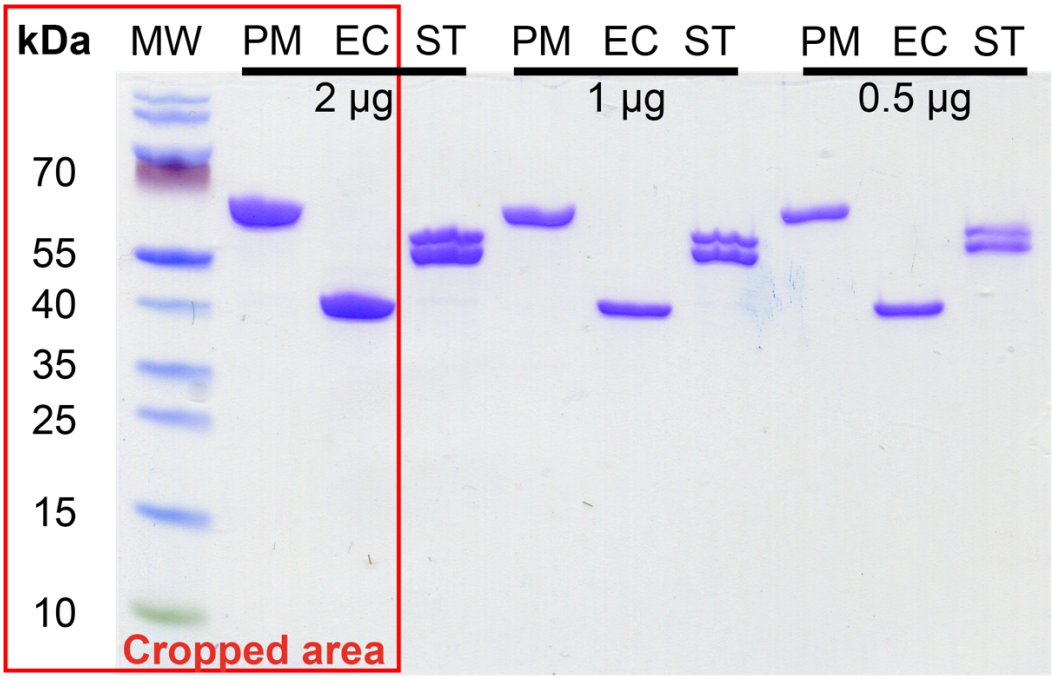



**Appendix C Doubling times of Δ*nfsB and* Δ*nfsB30* mutants.** There was no significant difference in doubling times between Δ*nfsB* & Δ*nfsB30* or Δ*nfsA* Δ*nfsB* & Δ*nfsA* Δ*nfsB30*.


**Appendix D Unaltered, annotated agarose gel picture showing the TLR1 and TLR6 mRNA.** MW: Molecular weight. The area highlighted in red shows the cropped region in Figure 63.




**Appendix E Unaltered, annotated agarose gel picture showing the TLR2, TLR4, and TLR5 mRNA.** MW: Molecular weight. The area highlighted in red shows the cropped region in Figure 63.




**Appendix F Unaltered, annotated SDS-PAGE gel picture of flagellin preparations.** MW: Molecular weight, PM: *Proteus mirabilis* isolate UTI 100, EC: *Escherichia coli* isolate TPA 3408, ST: *Salmonella typhimurium* isolate TPA 1. The area highlighted in red shows the cropped region in Figure 66

| Gene | Frequency | *P. mirabilis* homologues |
|---|---|---|
| *gdh* | 7 | gdhA |
| *pgm* | 6 | *pgm* |
| *adh* | 5 | *tdh* |
| *est* | 5 | N/A |
| *glp* | 5 | *glpK* |
| *thi* | 5 | *fadA* |
| *trp* | 5 | *trpE, trpB* |
| *alp* | 4 | *phoA* |
| *glpK* | 4 | N/A |
| *aro* | 3 | *aroE* |
| *cpn60* | 3 | *groL* |
| *gltA* | 3 | *gltA* |
| *mut* | 3 | *mutL, mutS* |
| *recA* | 3 | *recA* |
| *fusA* | 2 | *fusA* |
| *gua* | 2 | *guaA* |
| *gyrB* | 2 | *gyrB* |
| *icd* | 2 | N/A |
| *nuo* | 2 | *nuoH* |
| *pps* | 2 | *ppsA* |
| *pyrG* | 2 | *pyrG* |
| *rplB* | 2 | *rplB* |
| *rpoB* | 2 | *rpoB* |
| *rpoD* | 2 | N/A |

**Appendix G Overview of the genes used in at least 2 MLST schemes for gram-negative bacteria.** Genes highlighted in red were not present in *P. mirabilis*.

| Strain | fadA | guaA | tdh | trpE | PM | Database |
|---|---|---|---|---|---|---|
| HI4320 | 1 | 1 | 1 | 1 | **1** | NCBI |
| NLAE-zI-C285 | 1 | 3 | 24 | 23 | **2** | NCBI |
| AR 0155 | 1 | 31 | 9 | 11 | **3** | NCBI |
| PM185 | 1 | 31 | 9 | 11 | **3** | NCBI |
| 292 PMIR | 2 | 3 | 6 | 1 | **4** | NCBI |
| FDAARGOS 67 | 2 | 16 | 12 | 14 | **5** | NCBI |
| NLAE-zI-G534 | 3 | 12 | 29 | 10 | **6** | NCBI |
| 1091 | 4 | 9 | 3 | 5 | **7** | NCBI |
| 1326 PMIR | 4 | 17 | 15 | 23 | **8** | NCBI |
| 1330 PMIR | 4 | 17 | 15 | 23 | **8** | NCBI |
| AOUC-001 | 4 | 17 | 17 | 23 | **9** | NCBI |
| C05028 | 4 | 17 | 17 | 23 | **9** | NCBI |
| CYPM1 | 4 | 17 | 17 | 23 | **9** | NCBI |
| CYPV1 | 4 | 17 | 17 | 23 | **9** | NCBI |
| NO-051 03 | 4 | 17 | 17 | 23 | **9** | NCBI |
| PM 178 | 4 | 17 | 29 | 23 | **10** | NCBI |
| GB08 | 4 | 17 | 32 | 23 | **11** | NCBI |
| UTI139 | 4 | 20 | 26 | 8 | **12** | BUTI |
| 646 PMIR | 4 | 20 | 26 | 8 | **12** | NCBI |
| FDAARGOS 81 | 4 | 20 | 26 | 8 | **12** | NCBI |
| FDAARGOS 85 | 4 | 20 | 26 | 8 | **12** | NCBI |
| WGLW4 | 4 | 20 | 26 | 8 | **12** | NCBI |
| C02011 | 4 | 29 | 14 | 2 | **13** | NCBI |
| 1166 PMIR | 5 | 22 | 21 | 13 | **14** | NCBI |
| 1293 PMIR | 5 | 22 | 21 | 13 | **14** | NCBI |
| 1313 PMIR | 5 | 22 | 21 | 13 | **14** | NCBI |
| GB11 | 5 | 28 | 11 | 17 | **15** | NCBI |
| PM 125 | 6 | 17 | 29 | 23 | **16** | NCBI |
| UTI755 | 7 | 15 | 18 | 23 | **17** | BUTI |
| AR 0159 | 7 | 15 | 18 | 23 | **17** | NCBI |
| ATCC 29906 | 8 | 10 | 5 | 3 | **18** | NCBI |
| 373 PMIR | 9 | 27 | 23 | 16 | **19** | NCBI |
| BB2000 | 10 | 4 | 22 | 16 | **20** | NCBI |
| FDAARGOS 80 | 11 | 1 | 18 | 4 | **21** | NCBI |
| GED7834 | 11 | 2 | 16 | 2 | **22** | NCBI |
| 1114 PMIR | 11 | 2 | 16 | 19 | **23** | NCBI |
| ATCC 7002 | 11 | 2 | 16 | 19 | **23** | NCBI |
| FDAARGOS 60 | 11 | 2 | 16 | 19 | **23** | NCBI |
| UMB0315 | 11 | 2 | 16 | 19 | **23** | NCBI |
| 1230 SSON | 11 | 6 | 5 | 7 | **24** | NCBI |
| 672 PMIR | 11 | 7 | 28 | 9 | **25** | NCBI |
| 418 PMIR | 11 | 18 | 4 | 6 | **26** | NCBI |
| 232 PMIR | 12 | 6 | 10 | 1 | **27** | NCBI |
| AR 0059 | 12 | 6 | 10 | 1 | **27** | NCBI |
| 25933 | 13 | 25 | 8 | 24 | **28** | NCBI |
| PM655 | 14 | 24 | 2 | 26 | **29** | NCBI |
| AR 0156 | 15 | 11 | 30 | 18 | **30** | NCBI |
| T1C | 16 | 30 | 6 | 21 | **31** | NCBI |
| PM005 | 17 | 8 | 13 | 6 | **32** | NCBI |
| PM593 | 17 | 21 | 7 | 25 | **33** | NCBI |
| 127 PMIR | 17 | 23 | 8 | 5 | **34** | NCBI |
| 25 PMIR | 18 | 13 | 34 | 23 | **35** | NCBI |
| 360 PMIR | 18 | 13 | 34 | 23 | **35** | NCBI |
| 47 PMIR | 19 | 19 | 20 | 1 | **36** | NCBI |
| 51 PMIR | 19 | 19 | 20 | 1 | **36** | NCBI |
| 68 PMIR | 19 | 19 | 20 | 1 | **36** | NCBI |
| WGLW6 | 20 | 12 | 31 | 12 | **37** | NCBI |
| M16 | 20 | 12 | 27 | 20 | **38** | NCBI |
| PM187 | 21 | 32 | 33 | 22 | **39** | NCBI |
| Pm-Oxa48 | 22 | 26 | 6 | 15 | **40** | NCBI |
| 1134 PMIR | 23 | 20 | 9 | 11 | **41** | NCBI |
| 1150 PMIR | 23 | 20 | 9 | 11 | **41** | NCBI |
| PR03 | 24 | 5 | 25 | 8 | **42** | NCBI |
| 429 PMIR | 24 | 14 | 19 | 1 | **43** | NCBI |
| 430 PMIR | 24 | 14 | 19 | 1 | **43** | NCBI |
| 50664164 | 25 | 33 | 35 | 27 | **44** | NCBI |
| UTI100 | 26 | 34 | 36 | 28 | **45** | BUTI |
| UTI376 | 14 | 35 | 37 | 29 | 46 | BUTI |
| COW5395 | 29 | 38 | 31 | 32 | **47** | COW |
| COW5416 | 30 | 37 | 9 | 21 | **48** | COW |
| COW5273 | 28 | 36 | 39 | 31 | **49** | COW |

**Appendix H Strain ID, allele number (*fadA, guaA, tdh, trpE)* and PM (Sequence type) number of the 64 NCBI strains, 3 BUTI isolates, and 3 bovine isolates.** Allele numbers were assigned by temporal order, i.e. strains that were genotyped earlier had smaller allele numbers assigned to their *fadA, guaA, tdh, trpE* genes. ST numbers were assigned base on the unique combination of allele numbers for each isolate (Section 1.8.3.3) and like allele numbers, were assigned by temporal order.