1. Introduction

1.1 The global burden of infectious diseases including *Salmonella*

In economically developed regions of the world, infectious diseases are sometimes considered a problem of the past. Clearly infections such as the common cold, influenza, travellers-associated and healthcare-associated occur but the burden and impact are somewhat restricted compared to 50 years ago. However, infectious diseases remain a significant cause of morbidity and mortality across the globe, particularly in resource restricted settings such as Lower and Middle Income Countries (LMICs). According to the Global Burden of Disease Study 2017, the global incidence of communicable diseases, including TB, malaria, and HIV/AIDS, is on the order of 25 billion cases[1]. Two other leading causes of infectious disease associated with significant health impact are lower respiratory infections and diarrhoeal episodes. There are on the order of two million deaths per year from bacterial infections[2].

Of those, approximately 1.4 million global cases are attributed to diarrhoeal disease, still a significant problem in much of the world, particularly in children[3]. It is estimated that there were an estimated 6.29 billion diarrhoeal episodes in 2017, with an approximated 1.16 billion cases in children under 5 years, although this estimate is difficult to accurately assess[1,4].

Many of these infections occur in settings where healthcare and access to clean water is limited, and there are severe repercussions on longer term health and economic productivity[1]. It is estimated that in LMICs, the 842,000 deaths that are estimated to occur annually from diarrhoeal disease are attributable to a lack of clean water, sanitation and hygiene[5]. Of diarrhoeal diseases, in 2015 the leading cause of mortality from diarrhoeal disease was estimated to be rotavirus, followed by *Shigella* spp and *Salmonella* spp[1,3]. The 2015 Global Burden of Disease Study found diarrhoeal disease to be the ninth leading cause of death globally, and the fourth leading cause of death in children under 5 years, accounting for 499,000 child deaths[3].

1.1.1 Definition of *Salmonella enterica* by classical phenotyping and genotyping using MultiLocus Sequence Typing (MLST)

Salmonella species are a major contributor to systemic and diarrhoeal disease, with approximately 15 million systemic cases and 95.1 million diarrhoeal cases each year[1,3]. *Salmonella* form a group of Gram-negative, foodborne, rod-shaped bacteria that can invade the small intestine to cause limited or, more rarely, systemic infections. They were first described by the scientist Daniel E. Salmon in 1855. *Salmonella* are classified into two species: *Salmonella bongori* (*S. bongori*), which comprises 22 serovars and *Salmonella enterica* (*S. enterica*), which comprises six subspecies and at least 2650 serovars[6–8]. Serovars are traditionally defined using specific antisera. Isolates of serovars within both of these species can infect humans, but serovars within *S. enterica* remain more significant in causing human disease; *S. bongori* is primarily associated with the colonisation of reptiles[9].

Classically, the identification of *Salmonella* serotypes has been done using slide agglutination with specific antisera to distinguish between key surface antigens present on the surfaces of the different serovars. This typing of *Salmonella* can distinguish a serovar predominantly on the basis of the antigenicity of lipopolysaccharide (LPS) O antigen and flagellar antigens (phases 1 and 2 of H antigen, respectively), as described by the White-Kauffman-Le Minor scheme first developed in 1934[6]. There are at least 64 and 114 variants of the O and H antigens, respectively. Several O antigen types may be present on the cell surface of an isolate in conjunction, whereas only one flagellar variant is usually expressed at a time. Agglutination reactions are conducted by mixing the isolate being tested with the specific antisera against the O (encoded by genes including *rfb*) or H (flagella subunit encoded by *fliC* or *fljB*) antigens. In a surveillance setting or a reference laboratory, a large array of O and H antisera is normally used to type the *Salmonella* after culturing. For example, *Salmonella enterica* subspecies *enterica* Typhimurium (*S*. Typhimurium) is reported as 1,4,[5],12;i;1,2, which means that it is in subspecies I of *S. enterica* and it expresses O antigens 4, 12, and sometimes 5[6]. It is "i" for Phase 1 Flagellar (H) antigen and 1,2 for Phase 2 H antigen. The White-Kauffman-Le Minor scheme also includes capsular typing (K), which only applies to a small subset of serovars that harbour this type of surface antigen. *S*. Paratyphi C, *S*. Dublin, and *S*. Typhi isolates can express the Vi polysaccharide, a capsular antigen associated with immunogenicity and virulence[10–13].

More recently, the system of MLST has been advocated as a replacement of serological and biochemical characterizations to distinguish between *S. enterica* on the basis of DNA sequence and the associated evolutionary relatedness[14]. This method groups isolates within a given ST if they share identical alleles for a set of housekeeping genes, and isolates are placed within an ST-based clonal complex if they differ in one or two alleles[6,14,15].

Within *S. enterica*, the over 2500 serovars fit within six subspecies: subspecies I (ssp. I) *enterica*, ssp. II *salamae*, ssp. IIIa *arizonae*, ssp. IIIb *diarizonae*, ssp. IV *houtenae*, ssp. VI *indica*. The vast majority of serovars (approximately 1500) are classified within ssp. I *enterica*, causing over 99% of human and animal infections. Subspecies were previously distinguished from each other on the basis of general phylogeny and biochemical characteristics. O antigens are a key saccharide component of LPS on the surface of the *Salmonella* cell. While the composition differs between serovars, the O-antigen chain is comprised of a main branch of repeating units of sugars, which may contain branching sugars[16–18]. *Salmonella* can express LPS of differing composition and length on their surface, which can impact immune detection. This is an important factor for the development of *Salmonella* vaccines, as the diversity of LPS lengths and accessibility at the bacterial cell surface may compromise the utility of LPS as a suitable antibody target. There are additional characteristics that may be used to distinguish within serovars, including the phage types that may infect and lyse a given strain and other physical and biochemical properties, but these will not be considered further here.

S. enterica serovars and isolates can be arbitrarily further classified into two groupings; typhoidal and non-typhoidal on the basis of disease presentation. Typhoidal *Salmonella* are mainly the serovars Typhi and Paratyphi (A, B, C). Isolates of these serovars are associated with invasive systemic infections classically referred to as typhoid or enteric fever. Isolates of *S*. Typhi and *S*. Paratyphi A only cause a serious typhoid-like infection in humans and are consequently referred to as human-restricted. Isolates of *S*. Typhi and *S*. Paratyphi C can express Vi antigen, whereas *S*. Paratyphi A and B isolates normally do not. Further, *S*. Paratyphi B is actually a complex serovar and, like *S*. Paratyphi C, is now infrequently isolated globally compared to *S*. Typhi and *S*. Paratyphi[19–22].

Other typhoidal serovars can cause a typhoid-like disease in other animals. For instance, *Salmonella enterica* subspecies *enterica* serovar Gallinarum is host-restricted to chickens and causes a chicken typhoidal systemic disease. Interestingly, many host-restricted serovars

have genetically degraded genomes in that they have accumulated inactivated genes known as pseudogenes, and this factor may contribute to their host-restricted phenotype[23,24].

The majority of other serovars fall into the non-typhoidal classification. These include the classical serovars *S*. Typhimurium and *S*. Enteritidis. Isolates of these serovars are predominantly associated with localised gastroenteric disease, but this is not a strict definition as most disease-associated serovars can cause fully invasive infections in some circumstances. Also, in sub-Saharan Africa, isolates of *S*. Typhimurium and *S*. Enteritidis are a common cause of invasive disease[25]. Thus, the typhoidal/non-typhoidal classification is useful but not absolute.

1.1.2 Burden of all *Salmonella* serovars

As a single group of bacteria, *Salmonella* are a leading cause of bacterial infections globally, with approximately 167 million cases each year. The United States Centers for Disease Control estimates that 1.35 million cases arise annually in the United States, causing 26,500 hospitalizations and 420 deaths[26]. *S*. Typhi and *S*. Paratyphi, the leading causes of enteric fever, cause approximately 14.3 million cases and 135,900 deaths annually and are responsible for approximately 76% of enteric fever cases[27].

Together, *S*. Typhi and *S*. Paratyphi are referred to as typhoidal *Salmonella*, and they disproportionately affect people in LMICs, where prevalence is high. This burden is regarded as being highest in South and Southeast Asia, where the incidence rate may be between 200 and 700 cases per 100,000. However, as more studies are undertaken on typhoid in Africa there is an increasing recognition that the disease is very common in some parts of this continent[28,29]. In comparison, in higher income countries, the incidence rate is fewer than 15 cases per 100,000. Globally, there has been a considerable decline in typhoidal disease from 25.9 million cases in 1990 to 14.3 million cases in 2017; however, the burden is still significant. This is particularly true in South Asia, which accounts for 10.3 million (72%) of the 14.3 million cases[27].

Of the global cases of typhoidal *Salmonella*, it is estimated that 76.3% (10.9 million) were infections of *S*. Typhi, and 3.4 million were infections of *S*. Paratyphi. Demographically,

the incidence is highest amongst children between 5 and 9 years in high-incidence regions, followed by children between 1 and 4 years[27]. The fatality rate varies, though is highest in children and the elderly at up to 1.6%, with 17.2% of total deaths in children under 5 years, and 59.3% of total deaths in children under 15 years. The fatality rate is highest in South Asia, followed by sub-Saharan Africa[27]. However, data are sparse from many regions of the world, particularly Oceania and central sub-Saharan Africa, so the true burden of typhoid and paratyphoid fever may be higher than currently estimated[27]. There are also significant problems with the diagnosis of typhoid, which requires well founded laboratories for the gold standard approach of blood culture[19,30,31]. Serological assays are largely unreliable, in part due to the marketing of kits with poor specificity and sensitivity[30,32].

In contrast to typhoidal *Salmonella*, non-typhoidal *Salmonella* cause approximately 153 million cases each year[26]. These cases can be divided into non-typhoidal gastroenteritis and non-typhoidal invasive disease, which have strikingly different distribution, incidence, and fatality rates. CDC figures estimate that the highest rates of non-typhoidal *Salmonella* in US travellers abroad occur in Africa, with an incidence of 25.8 cases per 100,000 air travellers. Invasive non-typhoidal *Salmonella* infections caused by a variety of serovars are responsible for conservatively 534,600 cases and 77,500 deaths annually, or less conservatively 3.4 million cases and 618,316 deaths[33,34]. Because of insufficient surveillance in regions with the highest burdens, it is difficult to assess the true number of cases. Finally, infections localized to the gastrointestinal tract are responsible for nearly 94 million cases and 155,000 deaths annually[4].

1.1.3 Brief description of *S*. Typhi and *S*. Paratyphi

While the focus of this thesis is on non-typhoidal *Salmonella* (NTS) and specifically *S*. Typhimurium, it is important to distinguish NTS from *S*. Typhi and *S*. Paratyphi, the leading causes of enteric fever. Unlike most non-typhoidal infections, *S*. Typhi and Paratyphi cause systemic disease, characterized by lengthy high fevers, headaches, and general malaise, and if left untreated, can be debilitating or deadly. The two infections are difficult to distinguish from each other clinically but more insight has recently been acquired through human challenge studies[35]. These infections are highly associated with poor water supply and sanitation and are found predominantly in LMICs in South and Southeast Asia and sub-Saharan Africa.

Typhoid fever is largely managed using antimicrobials, and the three first-line antimicrobials to treat *S*. Typhi were chloramphenicol, ampicillin, and trimethoprim-sulfamethoxazole, followed by fluoroquinolones, third-generation cephalosporins, and azithromycin[36]. However, the choice of antimicrobial usage shows significant regional variation and in Asia the fluoroquinolones have until recently been the predominant drug of choice[37,38].

South and East Asia have a significant burden of Paratyphi A cases, with nearly 30% of all typhoidal *Salmonella* cases in India and Nepal and above 60% in China being of this serovar[27,35,39]. The number of cases of *S*. Typhi and *S*. Paratyphi is proportionally much lower than that of non-typhoidal *Salmonella* (NTS) globally; however, mortality from systemic infections is much higher when compared with overall NTS. Antimicrobial resistance (AMR) is becoming a key influencer in both typhoidal and non-typhoidal *Salmonella* disease. The incidence rates vary between serovars and by region, but the challenge of AMR is increasing[40–42].

Genomically, *S*. Typhi and *S*. Paratyphi are distinct from their non-typhoidal relatives and each other. They belong to *Salmonella* enterica subsp. I. *Salmonella* Typhi has O-antigen type O9-12, phase 1 flagellin type H:d, and is normally positive for the capsular Vi antigen. *S.* Typhi and *S.* Paratyphi have undergone extensive pseudogenization when compared with other broader host-range *Salmonella* serovars, thought to be a sign of host-adaption and restriction that have occurred more recently[43]. They share approximately 90% of their genomes with non-typhoidal *Salmonella* serovars, and *S*. Typhi contains on average around 200 pseudogenes, many due to early stop codons or frame-shift mutations. The reference *S*. Typhi CT18 genome is comprised of approximately 4599 coding sequences and encodes the Vi antigen, which is responsible for the production of capsule[44]. Phylogenetically, *S*. Typhi and *S*. Paratyphi sit separately from other serovars of *S*. enterica, with long branches between each other as well. *S*. Typhi is distinguished from *S*. Paratyphi by its Vi capsular profile; dominant serovars of *S*. Paratyphi do not show a capsular phenotype[45]. The severity of typhoid and paratyphoid fever justify the attention they have received; however, non-typhoidal *Salmonella* infections also present a significant burden to many parts of the world. Recent human challenge studies have been conducted using both *S*. Typhi and *S*. Paratyphi[20,46,47]. These have provided significant information on the in-volunteer evolution of these diseases and aspects of their pathogenesis, immunology and vaccinology[48,49]. Clearly, the two diseases have different temporal and bacteriological signatures, for example the patterns of bacteraemia differ significantly[47,50]. However, such studies were conducted in healthy

volunteers in England and with specific bacterial isolates, so they may not be completely representative of disease in the field.

1.1.4 Burden of non-typhoidal *Salmonella*

There is still a considerable burden of disease caused by non-typhoidal *Salmonella* (NTS) serovars, and it is a major cause of bacterial diarrhoea[4]. Of these, serovars *S*. Typhimurium and *S*. Enteritidis are responsible for over 50% of *Salmonella* infections worldwide[51]. Globally, non-typhoidal *Salmonella* is responsible for approximately 94 million cases of gastroenteritis (self-limiting diarrhoea) and 535,000 cases of invasive disease[4,33]. Estimates suggest that 86% of non-typhoidal infections are foodborne[4]. In the developed world, transmission of non-typhoidal *Salmonella* occurs by consumption of food or water contaminated with animal faeces. Undercooked eggs are a common source of infection by *S*. Enteritidis in particular[52–54]. Direct interaction with infected animals, and in some cases, direct contact with infected humans are also involved[55]. In addition, because the two predominant serovars of human illness *S*. Typhimurium and *S*. Enteritidis are broadly host-generalists, they can be transmitted from other animals, causing isolated cases.

Outbreaks in the developed world largely occur through food contamination in the supply chain and, while serious, these are typically self-resolving infections. Consequently, as the vast majority of gastrointestinal *Salmonella* infections are foodborne, they do not disproportionately affect a given demographic. Non-typhoidal serovars can cause invasive disease in the developed world. In the US, it is understood that approximately 7% of non-typhoidal *Salmonella* cases are invasive although these are primarily in immunocompromised hosts, and are responsible for only a small proportion of deaths[56].

In contrast, in parts of the developing world, non-typhoidal *Salmonella* serovars causes considerable invasive in addition to non-invasive disease. According to the GBD 2017 Non-Typhoidal *Salmonella* Invasive Disease Collaborators, the highest burden of invasive non-typhoidal *Salmonella* (iNTS) occurs in sub-Saharan Africa, where there are endemic serovars, predominantly *S*. Typhimurium, *S*. Enteritidis, *S*. Dublin, and *S*. Isangi[33]. There is also a considerable burden of iNTS in Southeast Asia, particularly amongst immunocompromised patients, e.g. in Vietnam[33,57]. Estimates of the incidence rate vary considerably, with a recent

analysis from the GBD 2017 Non-Typhoidal *Salmonella* Invasive Disease Collaborators estimating 535,000 cases and 77,500 deaths of iNTS in 2017, while an earlier analysis from Ao *et al.* in 2010 estimated 3.4 million cases and 618,316 deaths in 2010[33,34]. These infections disproportionately affect children under 5 years, similar to typhoidal *Salmonella*, and immunocompromised and malnourished populations. The burden is particularly high in those with co-morbidities of HIV, malaria, and sickle-cell disease[34,58–62].

1.1.5 Host distribution and transmission of *S.* Typhimurium

Of the many non-typhoidal serovars, *S.* Typhimurium is one of the most reported, causing a large percentage of gastrointestinal and invasive illnesses; hence the focus of this work on *S.* Typhimurium. *S.* Typhimurium is classically regarded as a host-generalist, meaning that it can infect a number of different species. The name "Typhimurium" refers to the typhoid-like systemic illness it causes in mice. Transmission, as for all other *Salmonella* species, is predominantly via the faecal-oral route, and the bacteria can be transferred between species through the ingestion of contaminated food and water and poor hygiene. *S.* Typhimurium foodborne infection is particularly associated with consumption of contaminated pork, beef, fruits and vegetables[63–65]. In the developed world, another mode of transmission is from domestic pets to humans. In humans, it is predicted that approximately 100 bacteria must be ingested to cause disease after passage through the acidic stomach environment[66,67]. In animals other than humans, *S.* Typhimurium readily transmits faecal-orally and causes infection, thus enhancing soil contamination[68]. While it is largely a gastrointestinal organism in humans, it may cause more widespread illness in other species, such as pneumonia in calves[69].

Interestingly, while the canonical understanding of *S.* Typhimurium transmission is that it occurs mostly through zoonoses, genomic studies of animal and human isolates especially of invasive *S.* Typhimurium from the same geographic areas in LMICs have shown that the animal samples can cluster separately from human ones[57,70,71]. This suggests that, in fact, transmission of *S.* Typhimurium is occurring more commonly between humans. The role of human-to-human transmission has been investigated in sub-Saharan Africa[72,73]. However, such research has not yet been able to definitively show what the reservoir for human-specific

S. Typhimurium infection is, what the transmission chain looks like, or what the contribution is of factors including food, water, or direct contact with infected animals and humans[74,75].

1.1.6 Phylogeny of *S*. Typhimurium

S. Typhimurium is a relatively diverse serovar of *Salmonella*, consisting of many sequence types that are distributed globally across many species. On the phylogenetic tree of *S*. enterica subspecies I, *S*. Typhimurium sits apart from other serovars, with approximately 40,000 – 60,000 SNPs divergence from its most recent common ancestor. Indeed, the serovar shows substantially more strain diversity at the isolate level than do many other serovars. While there is considerable diversity within *S*. Typhimurium, isolates have on average only acquired 400-600 SNPs, which accrues to an average nucleotide identity of 99.99%[76,77]. While some STs exhibit a broad host range, such as ST19 including DT104 (definitive type 104), others such as ST313 and ST34 look to have adapted to infecting certain species. Due to the prevalence and diversity of *S*. Typhimurium isolates, there has not been a recent comprehensive phylogenetic analysis of all *S*. Typhimurium STs, and differentiation between STs is largely assigned using MLST classification. The core genome of *Salmonella* is regarded to comprise approximately 3496 genes[78]. However, there is also additional diversity in the pan genome, and this is generally ignored when calculating SNP differences.

S. Typhimurium currently harbours 12 known STs, of which 9 form a single clonal complex. The central ST of the main clonal complex is ST19, and the other eight STs in this complex each have one allelic difference from ST19. ST19 is the most frequent ST isolated globally, and with the other eight STs in the complex, comprises the vast majority of *S*. Typhimurium isolates[7,79,80]. The three STs that sit outside of the clonal complex are ST36, ST40, and ST207.

While many of these sequence types, including ST19, are host-generalists, there are some that are more host-restricted. This includes pigeon-specific STs like ST98 and the human-associated ST34 and ST313[57,81,82]. Disease caused by ST19 isolates are reported globally, but there appears to be less geographic distribution of some of the other STs. For instance, ST36 is found widely in India, but it is less prevalent elsewhere; in contrast, ST34 appears to have found a distinct niche in humans in Southeast Asia[83,57]. The predominant clade of

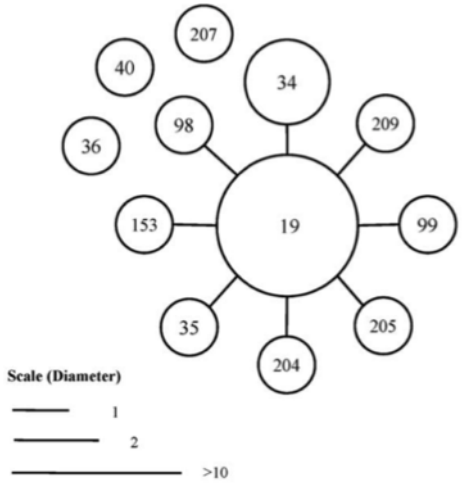

Figure 1.1 Radial phylogeny of *S*. Typhimurium. Phylogenetic organization of *S*. Typhimurium sequence types, adapted from Lan *et al.*, 2009[7].

ST313 has diverged substantially from ST19 and is found primarily in sub-Saharan Africa, although there have been isolates from Brazil and the UK, which predominantly cause gastroenteritis. Interestingly, isolates from the different geographic regions have distinct genotypes and phenotypes[84,85]. Clearly, there may be significant sampling bias in some of these associations and phenotypic links, and further work is required to obtain a fuller picture.

1.1.7 Disease presentation and treatment of *S.* Typhimurium in humans

S. Typhimurium infections in healthy humans typically presents as gastroenteritis, characterized by diarrhoea, stomach cramps, fever, and sometimes vomiting, with 5% of otherwise-healthy patients developing secondary bacteraemia[55]. These symptoms generally begin between six hours and six days after exposure to contaminated food or water and can last between four and seven days[55]. If there are no complications, the infection is self-limiting and resolves without intervention within a week. However, complications can arise in those with weakened immune systems—children under five, elderly, and immunocompromised individuals. This may occur in between 2 and 45% of patients with *Salmonella* gastroenteritis. In these cases, the bacteria can spread beyond the gastrointestinal tract, causing more systemic illness, including meningitis and septic arthritis, requiring intervention[86].

In severe cases, CDC guidelines recommend antimicrobial treatment, with patients administered broad-spectrum antimicrobials until any AMR data is obtained. For antimicrobial susceptible invasive *S*. Typhimurium infection, standard treatment includes fluoroquinolones, azithromycin, and third generation cephalosporins[87].

Invasive lineages of *S*. Typhimurium often cause more systemic disease in immunocompromised hosts, resulting in symptoms with some resemblance to those seen in typhoid and paratyphoid—high fever, general malaise, and headaches. Importantly, cases of invasive *Salmonella* are not always characterized by diarrhoea. As a result, iNTS is often confused with other diseases causing similar symptoms including malaria and typhoid, and cases may be misdiagnosed and mistreated as a result[88,59].

1.1.8 Human infection by *S*. Typhimurium

Depending on the *Salmonella* serovar and severity of infection, the distribution and diffusion of bacteria may vary. The focus here will be on *S*. Typhimurium, though most clinically relevant non-typhoidal *Salmonella* serovars may have the broadly similar initial interactions with the human host[89]. Bacteria are ingested orally and pass through the acidic stomach environment by deploying an acid tolerance response to enter the small intestine[90]. Once established on the surface of epithelial cells in the gut, which is dependent on the level of colonization restriction from the microbiome and the ability to pass through the mucous layer, they employ a type three secretion system (T3SS) encoded on the *Salmonella* Pathogenicity Island I (SPI-1). This initial interaction can trigger a gut inflammation, modulate host cytoskeleton arrangement, and the bacteria can enter cells[89,91–93]. Work is still ongoing to understand the mechanisms *Salmonella* employ to breach the lumen, but it is known that they can readily enter gut enterocytes, microfold (M) cells, and roving dendritic cells.

In the case of enterocytes, the bacteria can actively trigger host cytoskeleton remodelling, creating a membrane ruffle that enables bacterial entry[94,95]. In enterocytes, *Salmonella* are shunted by the cell into a membrane-enclosed vacuole, called a *Salmonella*-containing vacuole (SCV). Intracellular survival and replication are facilitated by virulence-associated factors encoded by genes within *Salmonella* Pathogenicity Island-2 (SPI-2) and other chromosomal loci[96–98]. This compartment is co-opted by the bacteria as a niche to hide and

potentially replicate in, and both the host and bacteria battle to exert influence on the fate of the SCV.

Salmonella can modulate host endocytic trafficking to limit SCV fusion with the lysosomal compartment, thus avoiding degradation[99–101]. In some cases, the SCV transcytoses to the basolateral membrane of the enterocyte, where it may be phagocytosed by dendritic cells, neutrophils, or macrophages. Enterocytes containing a high burden of *Salmonella* may also undergo apoptosis, resulting in their uptake by phagocytic cells and the dissemination of the bacteria[102,103]. In the case of direct entry into M cells, a type of specialized gut cell that *Salmonella* can invade at significant levels, the mechanism of entry is less clear. It appears that bacteria are able to penetrate these antigen-sampling cells without the aid of the T3SS, although they still modulate the host cytoskeleton, and the process may involve dynamin[104,105].

A further cell type exploited for entry are CD18+ dendritic cells that periodically traverse the lumen to monitor and sample antigens in the mucosa. Bacteria are phagocytosed by the dendritic cells, which then re-enter lamina propria, and the bacteria can then replicate within the dendritic cells[106,107]. Once in phagocytes, bacteria may trigger rapid or delayed pyroptosis, an inflammatory programmed cell death, thus facilitating further bacterial dissemination within and outside of the intestinal environment[108]. Notably, all of the interactions that occur between the host and *Salmonella* are enabled by a diverse set of virulence factors, many of them effector proteins produced by the T3SS[97,109–112]. Additionally, key host factors are adapted to control *Salmonella* infections[113–117].

1.2 Invasive non-typhoidal *Salmonella* compared to other non-typhoidal strains, an overview

Over the past 60 years, there has been an emergence of sub-lineages of non-typhoidal *Salmonella* that more frequently cause invasive disease (iNTS), with some similarities to infections caused by *S*. Typhi or Paratyphi[119]. These invasive *Salmonella* infections have been observed extensively in sub-Saharan Africa and to a more limited extent in Southeast Asia[57,82]. Due to high levels of malaria, HIV, and malnutrition that cause an immuno-

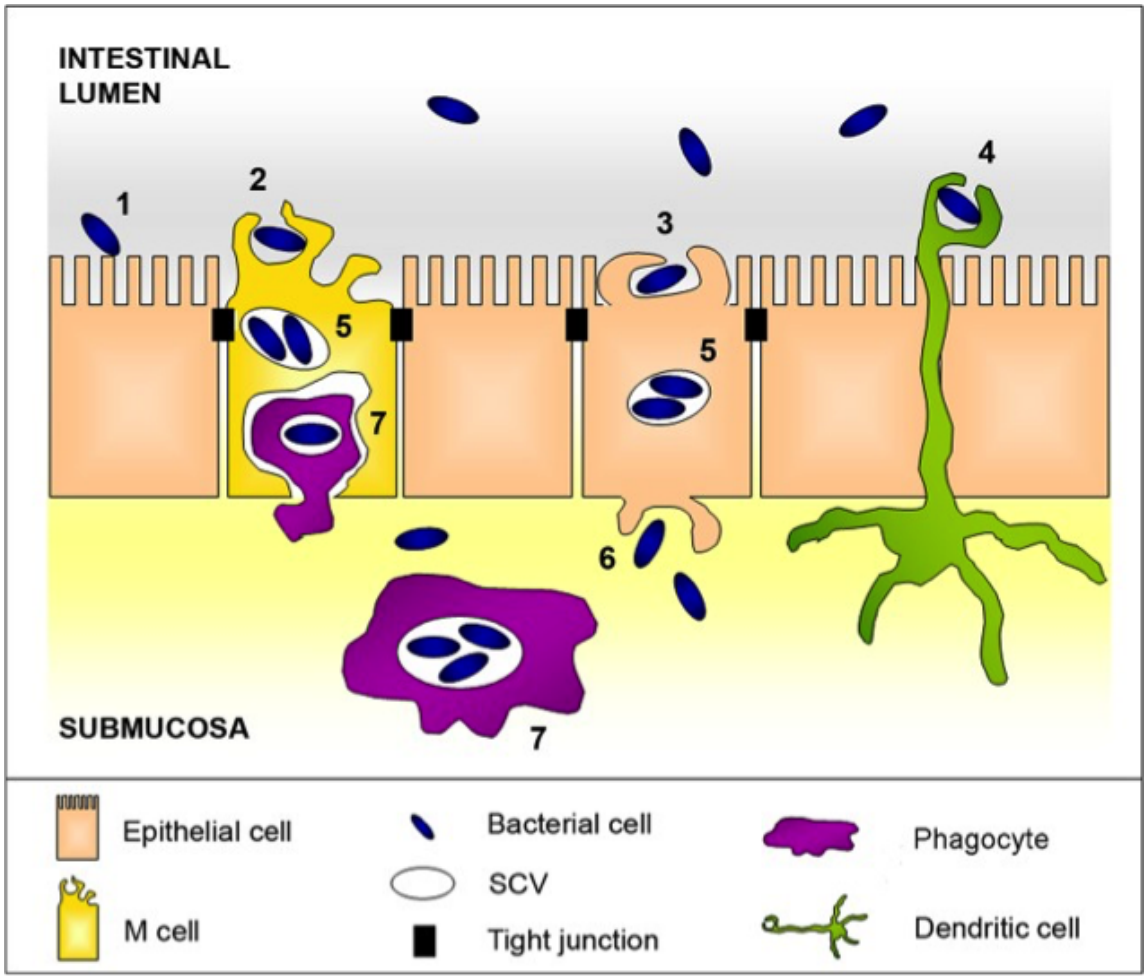

Figure 1.2 *Salmonella* invasion strategies of the intestine. *Salmonella* bacteria are able to enter the intestinal epithelium through multiple mechanisms, with and without uptake by immune cells. Adapted from Fabrega and Vila, 2013[118].

compromised state, infection with iNTS poses a serious risk with high mortality in those parts of the world where these conditions predominate, specifically in sub-Saharan Africa where *Salmonella* is one of the leading causes of mortality from bacterial disease [120,36,121]. While NTS can be invasive, this typically happens in only 8% of global cases [55]. However, certain serotypes, namely S. Dublin, Cholerasuis, Enteritidis, and Typhimurium are broadly also associated more frequently with invasive disease compared to other non-typhoidal serovars. Within these 'more invasive' serovars, specific sub-lineages that are genotypically and phenotypically distinct are potentially more likely to be invasion-associated.

1.2.1 Global distribution of iNTS

While iNTS occurs in the developed world, it is a more serious and pervasive problem in sub-Saharan Africa and parts of Asia. The burden of iNTS has been difficult to ascertain due to a general lack of effective surveillance, resources, and misdiagnosis. In particular, Latin America is severely lacking in surveillance of bacteraemia including iNTS, and it is unknown how severe a problem it is there [119]. iNTS is often mistaken for malaria, typhoid, or other bacterial infections, as they may all co-present to healthcare facilities in the same geographical regions [122–126]. The highest burden of iNTS occurs in sub-Saharan Africa and is predominantly caused by *S*. Typhimurium, which is responsible for 48% of cases across 33 countries [127]. Within Africa, surveillance has historically been better in Eastern and Central Africa, where there is a high level of invasive *S*. Typhimurium. In sub-Saharan Africa, iNTS is at least as great a problem as typhoid fever, and it is a persisting problem that is highly correlated with HIV infection, malaria, and malnutrition [88,128,129].

Worryingly, there is a rise in drug resistance associated with iNTS infection in sub-Saharan Africa, exacerbating the problem. Current and historic surveillance indicates that there is a spectrum of disease incidence across sub-Saharan Africa, with 8% and 45% of community-acquired bacteraemia due to iNTS in South Africa and the Democratic Republic of the Congo (DRC), respectively. The incidence of iNTS disease in sub-Saharan Africa was estimated to be two million cases in 2010. The region with the next highest incidence is Europe at 763,191 cases, largely due to higher prevalence in Eastern Europe [34].

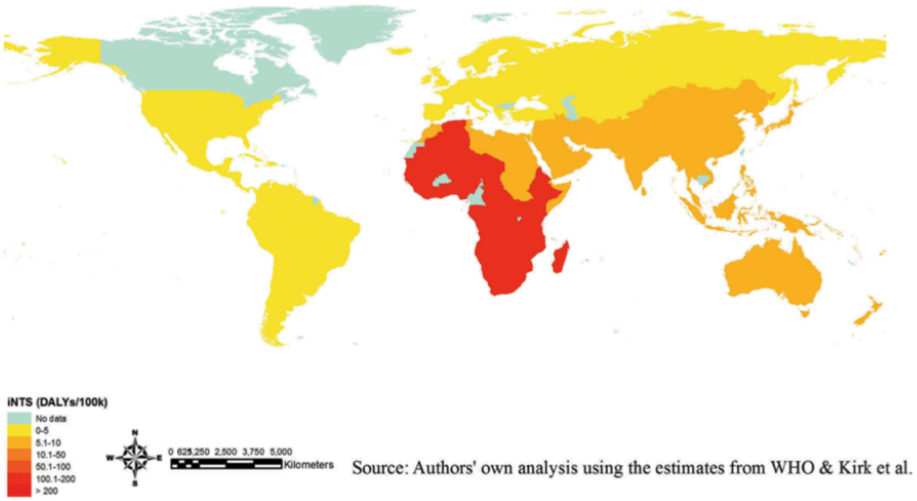

Figure 1.3 iNTS disability-adjusted life years (DALYs)/100k. iNTS infections have a distributed impact on DALYs globally, with the greatest burden in sub-Saharan Africa. Adapted from Balasubramanian *et al.*, 2019[130].

In South and Southeast Asia, regions with high levels of diarrheal disease, iNTS has not been found to be as common, particularly when compared to the incidence of typhoid and paratyphoid fever. There are high levels of *Salmonella* found in Southeast Asia, and a recent study on sepsis in Indonesia, Vietnam, and Thailand has documented that 2.7% of sepsis patients are positive for iNTS[131]. *S.* Typhimurium is the most common serovar associated with iNTS in South and Southeast Asia, and ST19 and ST34 are the two major *S.* Typhimurium STs responsible. Many of the cases of iNTS in Southeast Asia have been associated with HIV infection, as in Africa[57,131]. While surveillance for iNTS is not as widespread in some parts of Asia as in Africa, the burden of iNTS disease, despite a similar case fatality rate, seems to be lower in this region. In South Asia, there is a similar lack of data, though there are some reports concerning cases of invasive disease in India, suggesting that there may be higher levels of infection than reported in at least some regions[132].

In Latin America, a region that has particularly sparse epidemiological data on bacteraemia, there is likely to be some iNTS, though lower than in sub-Saharan Africa. It may be that data is available at a local and country level on this continent but this has so far not been broadly reported. A single surveillance study in Colombia identified 4,010 *S. enterica* samples isolates from blood and faeces over a six-year period, a far lower number of cases than that found in Africa[133]. ST313 found in Brazil appear to be genetically distinct from African ST313 and cause mostly gastroenteritis[119,130,85].

1.2.2 Disease presentation and treatment of iNTS

iNTS is acquired through faecal-oral contamination. Linked to the invasive nature of the infection, common clinical symptoms include non-specific markers of febrile illness such as fever, malaise, and headaches. However, beyond these clinical features, disease presentation may vary. A study of Malawian adults with iNTS found many to have bacteraemia, fever, and splenomegaly[134]. Another study has found that one-third of adults with iNTS present with respiratory symptoms[88]. It is less common for patients to present with diarrhoea compared to normal non-typhoidal *Salmonella* infections; diarrhoea occurs in approximately 35% of children and 46% of adults, based on multiple studies from Eastern and Central Africa[59,75,88,135–137,121]. In some cases, young children, typically younger than those with nonfocal disease, may present with meningitis, and the mortality rate from this complication is high, at approximately 50%, and subsequent neurological sequelae are found in approximately half of recovered patients[60,135,138,139].

As iNTS is most prevalent in low-resource settings, patients may not have access to healthcare facilities, and even if they do, there is a chance of misdiagnosis and inability to access the appropriate medication. iNTS is a severe illness, and initial treatment is broad-spectrum antimicrobials, such as the fluoroquinolone ciprofloxacin, until any drug resistance is confirmed. In response to high levels of multidrug resistance (MDR), ceftriaxone is now the standard treatment for undifferentiated sepsis in much of Africa[140]. In low-resource settings, it is unlikely that such rigorous testing is undertaken to diagnose the infection, subsequently test for AMR, and administer the appropriate antimicrobials. Depending on the region, access to broad spectrum antimicrobials like ciprofloxacin may be minimal. Due to the complications of misdiagnosis, and the time and resources involved to culture bacteria, antimicrobials may be arbitrarily administered, if at all.

1.2.3 Risk factors for iNTS disease

In acute iNTS cases, particularly in coinfections with malaria, HIV, malnutrition, and sickle cell disease, the prognosis is poor, and there is up to a 30% mortality rate in those with malaria. Much research has been done on the effects of malaria on iNTS infection, particularly in sub-Saharan Africa where the malaria burden is high. Studies have shown that

macrophages are hampered in their oxidative burst function by the presence of Plasmodium parasites, and complement mediated killing of bacteria is also reduced by consumption of complement[141–143]. Moreover, malaria disease appears to inhibit phagocyte recruitment to sites of bacterial infection, thus facilitating bacterial dissemination. Interestingly, this may be mediated through the induction of IL-10, an anti-inflammatory cytokine, as a component of the immune response to malaria, and this is likely consequently a side effect of a suppression of the inflammatory response to the bacteria as well[25,144,145].

HIV is another common risk factor for iNTS infection, particularly amongst adults in Africa and Vietnam. In Africa, approximately 95% of adults and 20% of children with iNTS patients are positive for HIV, and a study from Vietnam reported that 73% of patients with iNTS from *S*. Typhimurium ST34 were positive for HIV[57,88]. HIV results in a loss of CD4+ T cells, which increases susceptibility to iNTS by depleting the inflammatory response to *Salmonella*[146]. This loss subsequently fails to sufficiently activate bacteria-infected phagocytic cells, allowing for greater infiltration by bacteria in the initial infection and the potential for recurrent infection[82]. In addition, mucosal Th17 cells are also depleted during early HIV infection, resulting in poor initial response to an iNTS infection thus enabling it to spread more easily[147,148].

Malnutrition is another significant risk factor for iNTS infection in Africa due to reduced immune function in malnourished individuals and a leakier gut barrier[149]. Malnourished children additionally have lower levels of complement, and neutrophils have been shown to have lower activity in clearing bacteria[150,151]. These problems are potentially further exacerbated by a weaker inflammatory response, leading to insufficient recruitment of immune cells to the site of infection, and a lower overall humoral immune response, which delays recovery from infection[152,153]. In African children in iNTS endemic areas, sickle cell disease may also be a significant risk factor, causing significant immune dysregulation[154]. It is hypothesized that complement-mediated bacterial killing is severely restricted in patients with sickle cell disease, as is neutrophil oxidative burst, limiting the clearing of infection[155].

1.2.4 Genetic factors characterizing/influencing *S.* Typhimurium iNTS

While the majority of cases of iNTS are caused by *S.* Typhimurium, a serovar that can infect multiple species, an interesting characteristic of invasive serovars is their adaptation to a more human-specific niche. The most prominent example of this is *S.* Typhimurium ST313, a lineage that is the largest contributor to iNTS. Recent data from Ghana showed that the *Salmonella* isolated from human patients were genetically distinct from those isolated from the environment[156]. Comparative analysis of multiple ST313 and ST19 genomes have shown pseudogenization of some genes in ST313 that may equip the bacteria for a more host-adapted lifestyle[157,158]. It is known that the genomes of *S.* Typhi and *S.* Paratyphi A also display this characteristic—a more 'streamlined' and smaller genome that is potentially adapted to a human-specific lifestyle.

Experiments with ST313 *S.* Typhimurium have shown that, compared to ST19, they have an enhanced ability to infect macrophages and greater resistance to killing[157,159]. The prophage BTP1 in ST313 encodes a glycosyltransferase operon (gtr), which modifies LPS length. It is hypothesized that is may contribute to the increased invasiveness of ST313[160]. Furthermore, the ST313 genomes can harbour a single nucleotide polymorphism (SNP) in the promoter of gene *pgtE*, encoding outer membrane protease PgtE. PgtE is involved in resistance to certain antimicrobial peptides and the human complement cascade. The SNP in the promoter region of *pgtE* upregulates expression of PgtE, enhancing the ability to evade the human immune response and survive extracellularly in a host[161].

In contrast, ST34, which are a cause of iNTS in Southeast Asia, do not show this marked pseudogenization profile but do have a wide set of AMR signatures and unique flagellar characteristics[57]. A significant proportion of ST34 isolates from Vietnam are monophasic in terms of their flagellar expression. *S.* Typhimurium ST19 are normally biphasic, able to independently express the genes *fliC* and *fljB* that encode flagellar subunits. Dependent on environmental conditions, the *Salmonella* can undergo phase switching to alternate which flagellar protein is expressed[162,163]. However, the ancestral ST34 clone is monophasic, whereas a more recent set of Vietnamese ST34 clones are biphasic, having reacquired the *fljBA* operon containing *fliC* and *fljB* through a transposition event[57]. *In vitro* work in human macrophages and induced human gut organoids (iHOs) has shown that there are significant

differences in invasion between monophasic and biphasic *S*. Typhimurium ST34 isolates (Lees *et al.*, unpublished). This, combined with multidrug resistance (MDR), has likely enabled these isolates to thrive in this setting.

1.2.5 Phylogeny of iNTS

As previously mentioned, the four serovars most commonly causing iNTS are *S. enterica* Typhimurium, Enteritidis, Dublin, and Cholerasuis, although there is regional and temporal variation. While the isolates that cause iNTS do fit within the phylogenetic structure of their parent clades, they often form sub-structure in the phylogenetic tree, as evidenced by their cladal genetic differences. However, as a broader category, iNTS phylogeny has not been well studied due to the fact that iNTS isolates fall within multiple serovars, and these are not typically studied as a unit. The phylogeny of the dominant serovars that are associated with invasive disease, most prominently *S*. Typhimurium, is much better understood. *S*. Typhimurium ST313 is a major cause of iNTS, particularly in sub-Saharan Africa where it accounts for approximately reported 40% of cases[119].

1.3 iNTS in sub-Saharan Africa

Sub-Saharan Africa has the highest burden of invasive non-typhoidal *Salmonella*, based on the epidemiological data that has been reported to date. iNTS has emerged as a common causes of bacteraemia and is now considered to be the leading cause of bacterial-related morbidity and mortality in many parts of the region. This is in part to a significant reduction in the burden of *Neisseria meningitidis*, thanks in part to vaccination programmes in the northern "Meningitis Belt," targeting *N. meningitidis* A using a conjugate vaccine[164,165]. iNTS have adapted to the niche of the immunocompromised individuals, primarily those who are malnourished or have HIV or malaria. Given its rise in prominence and high rates of mortality in the region, iNTS and in particular *S*. Typhimurium ST313, warrants greater attention.

1.3.1 *S.* Typhimurium ST313 as major contributors to iNTS

iNTS has risen in prominence in recent years as surveillance and awareness of blood stream infections has increased. This has led to a better understanding of the serovars and sequence types that contribute most to the iNTS burden. In sub-Saharan Africa, where the iNTS burden is the highest, *S.* Typhimurium ST313 are responsible for the many cases. Interestingly, although African ST313 can cause both gastroenteritis and invasive disease, it is associated with more iNTS than NTS[119,166]. In Africa, approximately 40% of iNTS infections are routinely caused by ST313, and the next most common serovar is *S.* Enteritidis, which appears to cause approximately 12% of the iNTS, followed by S. Dublin at 11%[36,119]. Burden of disease estimations have not been calculated exclusively for ST313, but there have been many independent studies in various parts of Africa, particularly sub-Saharan Africa, documenting the proportion of ST313 cases of all *S. enterica* cases. These studies have provided evidence that ST313 is widespread across Africa and constitutes a major percentage of iNTS cases.

Comprehensive whole genome sequence-based analysis of 129 iNTS isolates across seven countries in sub-Saharan Africa, isolated between 1988 and 2010, found that 93% were ST313[82]. A study of 29 iNTS isolates from HIV-infected adults with fever in Mozambique found that 100% analysed by MLST were ST313, which is consistent with high levels of ST313 reported from neighbouring countries[167]. Similarly, longitudinal sampling in the DRC from 2007-2011 revealed that 96% of the *S.* Typhimurium isolates (79% of total non-typhoidal *S. enterica* isolates) recovered over that period belonged to the ST313 clade[168]. In Western Kenya, a recent study reported that 57.6% of iNTS cases in children under five years of age were caused by *S.* Typhimurium ST313, and 66.7% of iNTS cases in patients above five years of age were caused by ST313[169]. It is as yet unclear why ST313 have become the predominant cause of iNTS in sub-Saharan Africa above other serovars also endemic that have followed similar evolutionary paths. However, host adaptation and multiple antimicrobial resistance likely played a role[80,82].

1.3.2 Brief description of *S.* Enteritidis

While *S.* Enteritidis as a cause of iNTS has received less attention than *S.* Typhimurium, it is responsible for a significant proportion of cases in sub-Saharan Africa[119,170]. *S.* Enteritidis

infections follow a similar trend to *S*. Typhimurium in the developed world. The serovar is regarded as a host-generalist and cases mostly gastroenteritis; in the developing world, especially in sub-Saharan Africa, it is a common cause of invasive disease. Phylogenetically, *S*. Enteritidis is distinct from *S*. Typhimurium. A 2016 study on *S*. Enteritidis phylogeny showed distinct lineages in Africa compared to other parts of the world. In Africa, there are two main lineages of *S*. Enteritidis in circulation and associated with invasive disease, and these display evidence of genome degradation and harbour the virulence-associated plasmid, potentially signals of moving toward a more human host-restricted lifestyle[170].

Based on assessment of a global collection of 675 *S*. Enteritidis genomes, these two lineages appear to segregate on a geographic basis, rather than the temporal basis that defines ST313 lineages. One *S*. Enteritidis lineage is concentrated in West Africa and is distinct from the lineage found in Central and Eastern Africa, and both of these lineages were associated with AMR to at least one antimicrobial[170]. Further analysis of these genomes revealed that the two African-specific lineages appear to have split from the global epidemic clade and expanded in Africa after 1945 based on assessment of the most recent common ancestor. This timeline matches with the expansion of *S*. Typhimurium ST313 in Africa in association with the HIV pandemic and a newly emerging HIV-positive immunosuppressed population[88,171]. Similar to African *S*. Typhimurium ST313, African *S*. Enteritidis show evidence of genome reduction associated with a greater restriction to a human-associated lifestyle and higher invasiveness[170].

1.3.3 iNTS surveillance in Africa

There remains a significant gap in surveillance of iNTS disease globally. However, in the past decade, some regions of Africa have received more attention and surveillance because of the rising awareness that iNTS is a significant problem. The Typhoid Fever Surveillance in Africa Program (TSAP), administered by the International Vaccine Institute (IVI) collected data on bacterial infections, especially typhoid and iNTS, across thirteen sites in ten sub-Saharan African countries from 2010 to 2014. The countries included were: Senegal, Guinea-Bissau, Burkina Faso, Ghana, Sudan, Ethiopia, Kenya, Tanzania, Madagascar, and South Africa. These sites and countries were chosen on the basis of previously identified cases of typhoid fever, and field sites had already been established for other purposes. Importantly, these sites

were in a mixture of high and low population density areas. Of the 568 blood bacterial isolates collected in TSAP, 17% were found to be NTS, and 40% of those were *S*. Typhimurium[128].

While the TSAP study has ended, there are ongoing follow-up studies being conducted in Africa to continue surveillance of bacterial infections. This coincides with the roll-out of the typhoid conjugate vaccine and associated data collection in some countries. The Severe Typhoid in Africa (SETA) Program, also conducted by IVI has collected additional data on typhoid fever and other bacterial infections, particularly iNTS, following on from the TSAP study. However, SETA has concentrated on six sub-Saharan African countries: Ghana, Ethiopia, the DRC, Nigeria, Madagascar, and Burkina Faso, to identify *Salmonella* infections as the cause of fever in pre-determined regions and enrol patients for samples from multiple body sites and potential longer-term follow up studies[172].

Independently, the Institute of Tropical Medicine in Antwerp, Belgium has collected longitudinal data on and isolates from bloodstream infections in the DRC and Burkina Faso. Within-country surveillance also occurs in Kenya, largely administered by the Kenya Medical Research Institute (KEMRI). Surveillance conducted at KEMRI was instrumental in recognizing iNTS as a major cause of illness in sub-Saharan Africa, especially in HIV-positive patients[173,174]. While there is still incomplete surveillance in Africa, there is a much more comprehensive understanding of the endemicity of iNTS, and further surveillance will better elucidate the picture.

1.3.4 Phylogeny of *S*. Typhimurium ST313s

Many of the *S*. Typhimurium isolates associated with invasive *Salmonella* disease in sub-Saharan Africa fall into the lineage ST313[82]. This MLST branches off from the dominant ST19 type that is a common cause of salmonellosis in other parts of the world. Phylogenetic analysis based on whole genome sequences has shown that the ST313 in Africa have evolved into two distinct lineages—lineages 1 and 2[80,82]. Isolates in both lineage 1 and 2 are normally MDR, while isolates in lineage 2 have also acquired an additional resistance to chloramphenicol, an antimicrobial that was in common use while lineage II ST313 were emerging[80,82]. ST313 reference strain D23580 differs from ST19 reference strain LT2 by a number of SNPs, and these differences include 20 genes affected by four large deletions, 77

pseudogenes that could be characterized, and significant plasmid and prophage differences[80]. There are also additional pangenome differences.

Within the ST313 clade, there is considerable ongoing evolution, with recent genomic analysis revealing further changes in the AMR profile and phylogenetic structure. ST313 isolates from the 1990s and early 2000s clustered as a single lineage, forming ST313 lineage I, and isolates within this cluster were susceptible to the antimicrobial chloramphenicol[80,175]. The *S.* Typhimurium ST313 lineage II branched off later than lineage I and harboured chloramphenicol resistance, acquired very early on by this lineage. Bayesian (BEAST) analysis of 129 iNTS isolates in the context of a global collection of *S.* Typhimurium date the emergence of lineage I and II independent of one another $\sim$60 and $\sim$43 years ago, respectively, and the majority of ST313 cases from the mid-1990s onwards have been caused by lineage II bacteria[82].

Isolates from both lineages contain two Tn*21*-like transpositional elements, which appear to have independently arisen in the two lineages. While lineage I and lineage II cluster more closely with each other than with any other *S.* Typhimurium, there are distinct differences that separate them. Epidemiological data suggest that lineage II arose in response to widespread chloramphenicol usage, as lineage II is characterized by chloramphenicol resistance, whereas lineage I isolates are susceptible. Ongoing surveillance of iNTS in sub-Saharan Africa suggest that lineage I has been outcompeted by lineage II. With the increase in samples that have been sequenced in the past five years, it is now obvious that there is considerable substructure within the phylogenetic tree of ST313, and there is a significant sub-lineage emerging in the DRC that has extensive drug resistance.

Phylogenetic analysis by Van Puyvelde *et al.* of all available sequenced African ST313 have provided greater resolution of where and how ST313 is evolving, with a particular focus on the newly-identified sub-lineage II.1[176]. This sub-lineage (ST313 II.1), which branches off from lineage II, is defined by extended spectrum beta-lactamase production, azithromycin resistance, and an IncHI2 plasmid beyond the existing MDR profile of ST313. This sub-lineage has so far been responsible for > 10% of *S.* Typhimurium cases in the central region of the DRC; however, this outbreak illustrates the ongoing changes and evolution in ST313 and the ability of this lineage to adapt to antimicrobial exposure. Further investigations in the Congo region have found additional evolution related to decreased ciprofloxacin susceptibility. These isolates appear to form yet another sub-lineage (II.2) and may be responsible for a

significant number of cases of iNTS (Van Puyvelde, personal communication). It is likely that ST313 will continue to evolve in response to new or different treatment regiments used regionally.

1.3.5 Genetic and phenotypic characteristics of ST313 reference strain D23580

The best-characterized isolate amongst ST313 lineage II is D23580, which has been sequenced to completion and is used as a reference strain. D23580 was isolated from a febrile patient in Malawi in 2004 during an MDR outbreak of iNTS in Blantyre, Malawi and has since been extensively analysed[80]. AMR in D23580 is largely plasmid-mediated, encoded within a 117 kbp plasmid pSLT-BT, conferring resistance to chloramphenicol, ampicillin, streptomycin, sulphonamide, and trimethoprim. It also carries three other smaller plasmids, pBT1, pBT2, and pBT3, none of which carries known resistance genes. pSLT-BT shares significant homology with pSLT found in SL1344[177,178]. Beyond its MDR profile, D23580 is considerably different from ST19 strains, namely reference isolate SL1344. D23580 diverges from SL1344 by 856 core SNPs across the chromosome and also possesses a distinct array of prophages, which may influence virulence and survival[179]. D23580 has lost function of the gene *sseI*, a type III-secreted effector important in virulence and *ratB*, which is implicated in intestinal persistence in mice. There are also large blocks of deletions in a set of genes of unknown functions: STM1549-1553, and a loss of allantoin metabolism, which are similarly absent in *S*. Typhi[80]. These convergent changes with *S*. Typhi, a known invasive bacterium, suggest a shift towards a more human-host adapted lifestyle. Interestingly, while lineage I ST313 isolates show some signs of genome degradation in the form of pseudogenization, they do not share some of the novel chromosomal deletions present in lineage II isolates[80,82].

As D23580 has been fully sequenced and the genome annotated, it has been used extensively in the laboratory to explore the differences between ST313 and ST19. Metabolically, D23580 has lost the ability to ferment melibiose, has a lower utilization of L-tartaric acid and dihydroxyacetone, has some preference for alternative carbon sources, can ferment inositol, can use the butylene glycol pathway to make pyruvate, is able to survive on citrate as a carbon source, and has a lower utilization of purine and pyrimidine as phosphorous sources[180,157]. With regards to stress, D23580 has a greater resistance to human serum

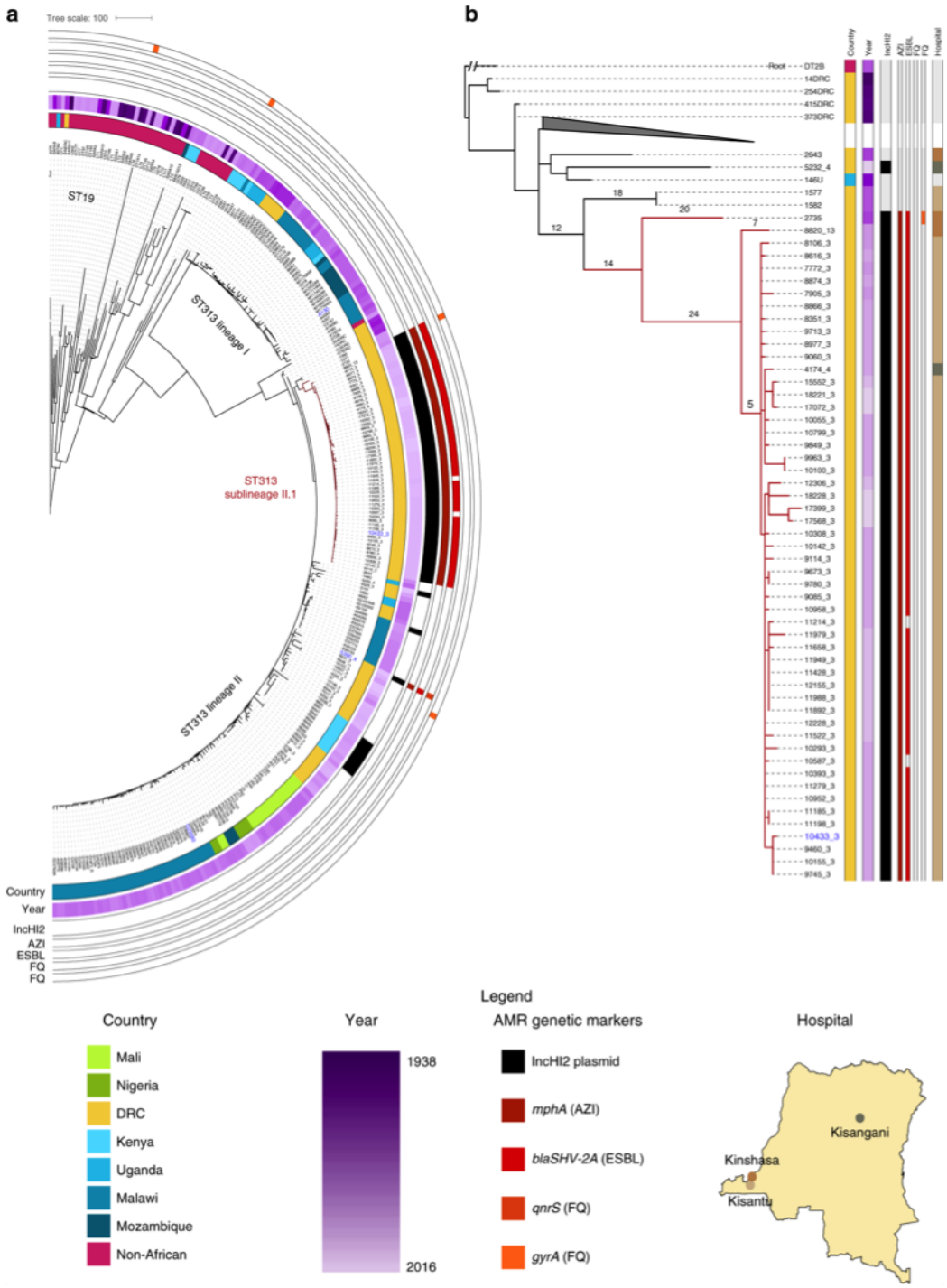

Figure 1.4 African *S.* Typhimurium ST313 sub-lineage II.1 in context of other African invasive *S.* Typhimurium (a) and independently (b). ST313 lineage II contains considerable sub-structure in its phylogeny, including recently-described lineage II.1 found in the DRC. From Van Puyvelde *et al.*, 2019[176].

killing, a higher tolerance to acid stress, a distinct and reduced biofilm morphology, and decreased survival in dry conditions[157,161,181–183]. In vitro comparisons with ST19 strains have shown that D23580 more readily invades and replicates inside macrophages; however, studies of epithelial cell invasion have produced contradictory results. In macrophages, D23580 and other ST313 lineage II isolates appear to cause less inflammation, which may account for greater intracellular survival[157,159,182,184,185].

Animal experiments have also been used to confirm and elucidate differences between ST313 and ST19 isolates: in mice, experiments have shown that D23580 disseminates from the gut more readily than SL1344, and there are higher levels of bacteraemia, alongside a less inflammatory response[158,180,184,186]. While other isolates have been subjected to long-read sequencing to the same depth as D23580 and have been experimented on, D23580 continues to be the best-described ST313 lineage II isolate.

1.4 AMR in *Salmonella*

AMR in bacterial pathogens has become a critical concern in recent years, with an increasing number of microorganisms exhibiting resistance to a spectrum of drug treatments. This is particularly concerning as there are few new antimicrobials being discovered or developed, and there is a dearth of vaccines to address the myriad of bacterial infections[2,187]. This presents a serious threat to the global community, as the rise in MDR, resistance to more than three classes of antimicrobials, may make even routine surgeries and hospital treatments an opportunity for bacterial infection. Moreover, these AMR organisms pose a great concern to immunocompromised and at-risk populations in the developing world, where there has been a sharp rise in MDR and more invasive infections, in large part due to indiscriminate use of antimicrobials[187].

In *Salmonella* species, the distribution of MDR is varied depending on the geography and serovars. In some parts of the world, notably Southeast Asia, where access to antimicrobials is easy, there are extremely high levels of MDR. In contrast, where access to antimicrobials is lower, as is the case in much of sub-Saharan Africa, the acquisition of MDR in *Salmonella* is generally slower. Similarly, in much of the developed world, particularly in parts of Europe that have rigorous antimicrobial stewardship, AMR in *Salmonella* is lower. However, there

has been a noticeable overall increase in drug resistance in *Salmonella* species globally to many of the first-line broad-spectrum antimicrobials, and it is likely only a matter of time before AMR further increases. The degree of AMR in *Salmonella* species is also dependent on the serovar. Regions with high endemicity of typhoid and paratyphoid, which require antimicrobial intervention, have much higher levels of drug resistance.

The H58 clade of *S*. Typhi, an MDR clade that has become the dominant clone globally, expanded at least in part because of its ability to outcompete other antimicrobial susceptible clones[188–190]. In regions where there is particularly high antimicrobial usage, H58 has acquired further AMR, including resistance to azithromycin, one of the last treatment options for MDR *S*. Typhi. This has resulted in an ongoing outbreak of extremely drug resistant (XDR) *S*. Typhi first reported in Hyderabad, Pakistan[191]. Similarly, there are ongoing outbreak of MDR *S*. Typhi in sub-Saharan Africa[190,192–194]. Combined with poor healthcare infrastructure and treatment, this outbreak has the potential to expand and acquire further AMR. This then has a knock-on effect on other bacteria found in the environment. Although iNTS serovars may not naturally carry as much plasmid-mediated AMR as *S*. Typhi, if they coexist in the same environment, as is the case in parts of the developing world, particularly in sub-Saharan Africa, then the potential for horizontal transfer of AMR genes increases.

1.4.1 WHO R&D 'directive' on FQR *Salmonella*

In 2017, the World Health Organization (WHO) issued a priority list of antimicrobial-resistant tuberculosis and other bacterial pathogens, listing the infections that had concerning levels of drug resistance. The 20 bacterial pathogens and 25 patterns of resistance acquisition of greatest concern were determined using ten criteria: mortality, health-care burden, community burden, prevalence of resistance, 10-year trend of resistance, transmissibility, preventability in the community setting, preventability in the health-care setting, treatability, and pipeline[187]. The pathogens and resistance mechanisms chosen were categorized based on their criticality—critical, high, and medium priorities. Notably, of these, Gram-negative bacteria were a more critical concern than Gram-positives, and fluoroquinolone resistant (FQR) *Salmonella* species fell within the "Priority 2: high" category. This categorization was made on a multitude of factors, and for non-typhoidal *Salmonella*, some of the key factors were mortality, transmissibility, and likelihood for alternative treatment or vaccine

development[187]. This directive from the WHO has helped clarify which antimicrobials need to be most actively conserved for emergency use and how to prioritize research and development into novel therapeutics.

1.4.2 How AMR is assessed in clinical settings

The gold-standard for antimicrobial susceptibility testing (AST) is broth microdilution measurements, which involves growing bacteria in a standardized set of antimicrobial concentration over a designated period of time[195,196]. However, this method is time intensive and laborious, and most clinical laboratories assess antimicrobial susceptibility using antimicrobial disk diffusion tests or minimum inhibitory concentration (MIC) test strips. Disk diffusion, or Bauer-Kirby, testing works by placing a disk containing a defined concentration of an antimicrobial on an agar plate inoculated evenly with $\sim$1-2 x 10^8 colony forming units of bacteria/millilitre and subsequently measuring the zone of bacterial growth inhibition that occurs after overnight incubation. This method is widely-used due to the ease, standardization, and reliability of the assay and relatively low cost of antimicrobial disks[195,197,198]. MIC test strips operate on the same principal of a zone of inhibition, although in this case, incrementally more concentrated antimicrobial is dotted along the underside of a plastic strip. After the strip is placed on an agar plate inoculated evenly with bacteria, the level at which bacterial growth is no longer inhibited is determined to be the MIC. Disk diffusion tests remain the cheapest option, so they are the ones most commonly used in the field, although they are considered less reliable and to generate slightly lower MICs than do the MIC test strips[199]. They are considered less reliable in part because they are more difficult to interpret. A caveat with any MIC test that involves reading the result is that results may vary depending on the stringency of the 'reader', and it becomes very important to maintain consistency of how the results are read and interpreted.

Yet another method that has gained widespread use and acceptance in recent years is the automated Vitek System, which uses reagent cards containing minute quantities of antimicrobials pre-loaded. Bacterial culture at a specified optical density are added to the plastic cards, and a machine measures the turbidity of the culture under antimicrobial exposure after a period of incubation. These data are then outputted as a report, specifying whether bacteria

are above or below the breakpoint for the given antimicrobial, and whether that corresponds to a susceptible, intermediate, or resistant phenotype[195].

There are two institutional bodies that determine MIC "breakpoints" for any given antimicrobial based on levels of susceptibility of groups of bacteria. This is updated every couple of years to reflect changes in antimicrobial susceptibility. One such organization is the Clinical and Laboratory Standards Institute (CLSI), an international voluntary organization accredited by the American National Standards Institute, which collates reported data from clinicians and laboratories to advise appropriate breakpoints for antimicrobial usage[200,201]. The other is the European Committee on Antimicrobial Susceptibility Testing (EUCAST). While these are both accepted, the breakpoints for given bacteria for specific antimicrobials may vary because they use different metrics to assess levels of resistance, which may lead to inconsistencies between how an infection is treated[202,203]. While breakpoints determined from both organizations are accepted and recommended by the WHO, they do not necessarily align, which may be problematic in determining whether a patient should be treated with a certain drug on the basis of AST[202].

1.4.3 Distribution of fluoroquinolone usage

Fluoroquinolones are one of the most widely-used classes of antimicrobials, given their ability to function on both Gram-positive and Gram-negative bacteria. Amongst fluoroquinolones, one of the most widely-used is ciprofloxacin, a second-generation fluoroquinolone. In the developed world, it is commonly prescribed to travellers for traveller's diarrhoea and is the first line empirical treatment for an invasive infection in hospitals[204,205]. A study of antimicrobial use in 2004 in the United States and European Union found that fluoroquinolones were heavily used as a first-line treatment for outpatient respiratory tract infections in the United States[206]. For ciprofloxacin, the defined daily doses per 1000 inhabitants per day (DID) was 0.97% in the United States compared to a range of 0.17% (Croatia) - 1.81% (Portugal)[206]. In the developing world, usage of fluoroquinolones is highly correlated with access, and this differs across regions. The highest use of fluoroquinolones is in South and Southeast Asia where fluoroquinolones are accessible from a pharmacy without prescription and are widely used by patients as self-medication.

A study of paediatric patients presenting with diarrhoea in Ho Chi Minh City, Vietnam between 2014 and 2016 found that over 66.7% of admitted patients were administered fluoroquinolones, and in many cases, this was prior to determination of the infection[207]. A study conducted in Singapore between 2006 and 2008 found that there was a steady increase in ciprofloxacin administration amongst hospital inpatients, from 317.87 to 448.76 defined daily doses per 1000 inpatient-days over the two-year period. Compared to the other classes of antimicrobials tracked in this study, fluoroquinolone usage increased the most[208]. A study of AMR and antimicrobial prescription in regions of South Asia and Africa for *Escherichia coli* (*E. coli*) infection found that in three South Asian study sites (Pakistan, India, and Bangladesh), 60% of the antimicrobials administered to patients presenting with dysentery were ciprofloxacin. In contrast, for cases of watery diarrhoea, ciprofloxacin was given approximately 10% of the time in Pakistan and Bangladesh, compared to 60% of the time in India[209]. In Africa, the same study found that ciprofloxacin usage was much lower, with ciprofloxacin accounting for 0-10% of prescribed antimicrobials for either diarrhoea or dysentery[209]. There is insufficient data about fluoroquinolone usage in Africa, although there is still distribution and use of fluoroquinolones, particularly for the treatment of typhoid and other bacteraemia, for which ciprofloxacin is one of the primary treatment options. There is some data showing widespread usage of antimicrobials in livestock, and this heavy usage, ranging from 77.6% in Nigeria to 100% in certain countries throughout Africa, may allow bacteria to develop resistance in the environment that then adversely affect patients[210].

1.4.4 FQR in *S.* Typhi and Paratyphi

AMR, including fluoroquinolone resistance (FQR), is a considerable problem in *S.* Typhi and *S.* Paratyphi[19,211,212]. After fluoroquinolones were introduced, cases of fluoroquinolone-resistant *S.* Typhi and *S.* Paratyphi became apparent in the 1990s and 2000s[212–214]. At that time, there were still significant numbers of FQ-susceptible isolates, and this was the case through the mid-2000's. However, after fluoroquinolones became an established treatment for *S.* Typhi, there has been an increasing trend of FQR on top of existing MDR. The H58 (genotype 4.3.1) lineage of *S.* Typhi are often MDR, for example by carrying incH1 plasmids, which contain multiple resistance genes, and many of the isolates sequenced also have decreased fluoroquinolone susceptibility due to chromosomal and sometimes plasmid-mediated resistance[188]. In some cases, susceptibility still exists to later-generation fluoroquinolones,

such as the fourth-generation fluoroquinolone gatifloxacin[215,216]. However, given the prevalence of fluoroquinolone usage and the pressure bacteria are under to develop resistance, there is an evolution of resistance amongst typhoidal *Salmonella* against fluoroquinolones and the likelihood is that resistant strains are at least as fit fluoroquinolone susceptible ones[217]. A recent outbreak of XDR *S*. Typhi in Pakistan not only displayed resistance to the three first-line treatments but also had a decreased susceptibility to fluoroquinolones, marked by carriage of plasmid-mediated resistance gene *qnrS*, which has complicated the treatment strategy[191]. Interestingly, *S*. Typhi seems much more adept than *S*. Typhimurium at acquiring plasmid-mediated FQR. In the face of XDR *S*. Typhi, azithromycin and ceftazidime have been used successfully for FQR, although there are limited reports of resistance of azithromycin-resistant *S*. Typhi[218]. In Africa, there has been an increase in MDR *S*. Typhi, marked by a decreased ciprofloxacin susceptibility[219,220]. These outbreaks further validate the need for wider vaccination efforts in areas with endemic *S*. Typhi.

Vaccines against *S*. Typhi and *S*. Paratyphi

Recent efforts to control the spread of *S*. Typhi, especially in light of the increase in multidrug resistance, have focused on wider vaccination efforts. There are currently two broadly licensed vaccines against *S*. Typhi; an oral live attenuated Ty21a and inactivated parenteral Vi capsular polysaccharide vaccine, which have been used in at-risk populations and travellers but do not provide full immunity[221,222]. The Ty21a vaccine is administered in three doses and is largely used by travellers visiting endemic *S*. Typhi areas. The efficacy of this vaccine is 65%, and the vaccine needs to be administered every 3-5 years for continued protection[222,223]. The Vi polysaccharide vaccine, which is administered as one intramuscular injection, has a three-year efficacy of approximately 55%, and a follow-up study has shown ten-year protective antibody levels in approximately half of the immunized population[224,225]. Importantly, neither Ty21a nor the Vi polysaccharide vaccine are approved for use in children under two years of age, a high-risk population[222].

More recently, a typhoid conjugate vaccine (Typbar-TCV) manufactured by Bharat Biotech has undergone clinical trials in Malawi, Bangladesh, and Nepal, regions with high levels of endemic *S*. Typhi. This vaccine shows considerably higher efficacy, with up to 85% protection following immunization of children from nine months to 15 years of age, and it has now been pre-qualified by the WHO for wider trials and usage[226–228]. Pakistan was

one of the first countries to add this vaccine to their national immunization programme in response to the XDR typhoid outbreak, and India has introduced TCV into its childhood immunization programme in Navi Mumbai[229,230]. The IVI is currently implementing studies to continue to evaluate the vaccine and follow up cohorts to understand effects on the *S*. Typhi burden, transmission, and prevalence of other sources of bacteraemia in various sites in Africa. Countries enrolled in Typbar-TCV vaccine trials and administration are Ghana, the DRC, and independently, Bangladesh. The Effect of a novel typhoid conjugate vaccine in Africa: a multicentre study in Ghana and the DRC (THECA) will conduct clinical studies and a mass vaccination campaign, measuring vaccine efficacy[231].

There are no current licensed vaccines in general use against *S*. Paratyphi. There is a significant need for a vaccine against *S*. Paratyphi A, given the high burden in parts of South and Southeast Asia, where the incidence approaches that of *S*. Typhi. There are some vaccine candidates under development and in Phase I and II clinical trials, and more research and testing is needed to determine their efficacy and suitability for routine use[232].

1.4.5 FQR in *S*. Typhimurium

FQR in *S*. Typhimurium is not yet considered as problematic as in *S*. Typhi and *S*. Paratyphi, largely due to the fact that NTS infections generally do not require antimicrobial intervention. However, this situation is changing with the rise in ST313 and other iNTS infections. Epidemiological data suggest that *S*. Typhimurium isolates primarily acquire chromosomal mutations in *gyrA*, and there are scant examples of other chromosomal mutations in *S*. Typhimurium that confer direct FQR, unlike what is found in *S*. Typhi[233,234]. Plasmid-mediated quinolone resistance (PMQR) in *S*. Typhimurium is also significantly less than what is seen in *S*. Typhi, with only one out of 276 isolates of analysed African ST313 isolates from before 2017 displaying a PMQR gene[176]. This likely reflects that fluoroquinolone usage in iNTS is still lower than in typhoidal infections, at least in Africa. However, there are regions that have high levels of FQR, namely South and Southeast Asia where fluoroquinolone usage is high[57,235]. In Vietnam and other nearby Southeast Asian countries, *S*. Typhimurium ST34 exhibit extensive MDR, often with high levels of resistance to fluoroquinolones[236]. Many of these isolates carry multiple mechanisms of resistance to fluoroquinolones and other drugs[131,237,238].

Ciprofloxacin is one of the most commonly used broad-spectrum antimicrobials and is the first-line treatment for invasive salmonellosis in many parts of the world. Over the past several years, the prevalence of *S*. Typhimurium strains with intermediate and complete ciprofloxacin resistance has grown. The patent for ciprofloxacin expired in 2003, and the introduction of affordable generics into Africa has influenced higher usage across the continent[239]. According to CLSI, *Salmonella* are considered sensitive to fluoroquinolones excluding nalidixic acid if the minimum inhibitory concentration is ≤ 0.06 μg/ml of the FQ. Susceptibility is considered intermediate for an MIC of 0.12 - 0.5 μg/ml, and resistance is an MIC of ≥ 1 μg/ml of the FQ[200,240,241].

In African ST313 lineages, there is not yet widespread FQR; however, there has been an observable decrease in ciprofloxacin susceptibility (DCS) over the past decade. This has been notable in the DRC, where there is an ongoing outbreak of DCS ST313 infections with a high degree of MDR (Van Puyvelde *et al.*, in progress). Perhaps because of simultaneous outbreaks of *S*. Typhi and *S*. Typhimurium in the DRC, there is an increased likelihood of bacteria becoming DCS[220] (Van Puyvelde and Dyson, in progress). In particular, DCS has been observed in ST313 lineage II.1 isolates, where substructure in the phylogenetic tree reveal specific branches show a phenotypic and genotypic decrease in ciprofloxacin susceptibility. Most of these isolates do not yet have full resistance to ciprofloxacin and other FQs, but there is the potential for them to acquire further resistance[176].

1.4.6 Other AMR profiles in invasive *S*. Typhimurium

While resistance to fluoroquinolones is one of the most pressing AMR concerns in *S*. Typhimurium because of its importance as a first-line broad-spectrum antimicrobial, it is not the only class of drugs that *S*. Typhimurium clades have shown resistance to. This problem is most worrisome for invasive clades, especially ST313 and ST34 that cause more disease in immunocompromised individuals. Some isolates of ST34 carry an IncHI2 plasmid, which confers drug resistance to: fluoroquinolones, bleomycin, sulphonamides, trimethoprim, kanamycin, streptomycin, chloramphenicol, spectinomycin, florfenicol, hygromycin B, apramycin, beta-lactams, and rifampin, and they may further carry mercury resistance genes[57]. Analysis of ST313 lineage I and II isolates when first sequenced revealed that they had a substantial AMR profile carried by a Tn*21*-like transposition element[82]. Lineage I

isolates lack resistance to chloramphenicol but otherwise carry resistance genes on the pSLT virulence plasmid against cotrimoxazole, ampicillin, and sulphonamide, and trimethoprim. Lineage II isolates carry additional chloramphenicol and streptomycin resistance[80]. Recent analysis of ST313 lineage II and sublineage II.1 isolates have shown that sublineage II.1 also contains an IncHI2 plasmid, named pSTm-ST313-II.1. This plasmid carries genes conferring resistance to chloramphenicol, ampicillin, and trimethoprim/sulfamethoxazole. In XDR sublineage II.1 isolates, there are additional extended spectrum beta-lactamase (ESBL) genes that confer resistance to cephalosporins and plasmid-mediated azithromycin resistance. Sublineage II.1 isolates also carry genes that may confer silver and copper resistance[176]. Importantly, this analysis was conducted on isolates predominantly from the DRC, and it is possible that the AMR profile of ST313 lineage II and other potentially emerging sublineage isolates elsewhere in Africa differ slightly in their AMR makeup.

1.4.6.1 Threat of XDR *S*. Typhimurium ST313

An outbreak of extremely drug-resistant *S*. Typhimurium has recently been documented in the DRC, and this accounts for >10% of all *S*. Typhimurium isolates collected in the Kongo Central Province[176]. Given the trends of resistance already observed, without intervention in antimicrobial stewardship or vaccination, there is a strong likelihood of this occurring more frequently and across a wider geography. One region of significant concern is the DRC, where there are observed incidences of MDR and DCS ST313 and independently of MDR ST313 carrying plasmid-mediated azithromycin resistance, both sets of which are also ESBL positive[176] (Van Puyvelde et al., in progress). There is also an observation of three pan drug resistant ST313 isolates that carry the MDR plasmid pSTm-ST313-II.1 as well as plasmid-mediated azithromycin resistance and either chromosomal or plasmid-mediated DCS.

These two phenotypes of isolates predominate in different geographic regions of the DRC, but given the ease of horizontal transmission occurring and spontaneous DCS arising, there is a number of scenarios that could emerge. One possibility is that a spontaneous chromosomal mutation conferring DCS or even full ciprofloxacin resistance (CR) could arise in the azithromycin-resistant isolates. Or, they could acquire a plasmid-mediated resistance gene conferring CR. Alternatively, bacterial populations that are DCS could acquire a novel azithromycin resistance mutation or acquire plasmid-mediated azithromycin resistance. In

any country, MDR bacteria with DCS and azithromycin resistance pose a considerable threat. The DRC, despite increased surveillance of cases of bacteraemia, is a country ill-equipped to mitigate such an outbreak, making further surveillance and iNTS vaccine initiatives even more necessary. It will be interesting to see how the introduction of the typhoid conjugate vaccine influences the MDR profile and burden of ST313 in the DRC.

1.4.7 iNTS vaccines

Given the burden of iNTS, and specifically *S*. Typhimurium ST313, various entities have begun investigating and developing vaccines against iNTS. While iNTS remains a lesser-known problem globally, it is appreciated as an acute concern in sub-Saharan Africa, and a vaccine against iNTS could make significant differences in child mortality and reduce pressure on antimicrobial usage to treat bacteraemia[25,130]. There are currently efforts underway to develop a vaccine against iNTS that covers *S*. Typhimurium and *S*. Enteritidis, and some of these have performed positively in animal models[242,243,51]. One of the vaccine candidates is a bivalent conjugate vaccine that uses the core and O-antigen polysaccharides of *S*. Typhimurium and *S*. Enteritidis linked to the FliC (Phase I flagellin subunit) protein of each serovar. A second candidate is a live attenuated oral vaccine, which would consist of *S*. Typhimurium and *S*. Enteritidis lacking *guaBA*, *clpP* or *clpX*, and *fljB*, which encode guanine synthesis, flagellar regulation, and flagellar function, respectively. In mice, these attenuated strains have been shown to cause an attenuated response and generate some protection[244].

A third type of vaccine is led by GSK Vaccines for Global Health (GVGH) and uses the Generalized Modules for Membrane Antigens (GMMA) technology for vaccine development. The principle behind this technology is the use of Gram-negative outer membrane particles that are shed from genetically modified bacteria to deliver immune-stimulatory components, namely outer membrane antigens including O-antigen, directly to the immune system[245]. These particles include outer membrane lipids, outer membrane proteins, and soluble periplasmic components, and the bacteria can be genetically modified to shed more or less of these constituents. As these components form part of the bacterial outer membrane, they are highly stimulatory to the human immune system, thus eliciting a measurable protective responses, including a B-cell response and antibody generation against O-antigen and porins[246,247].

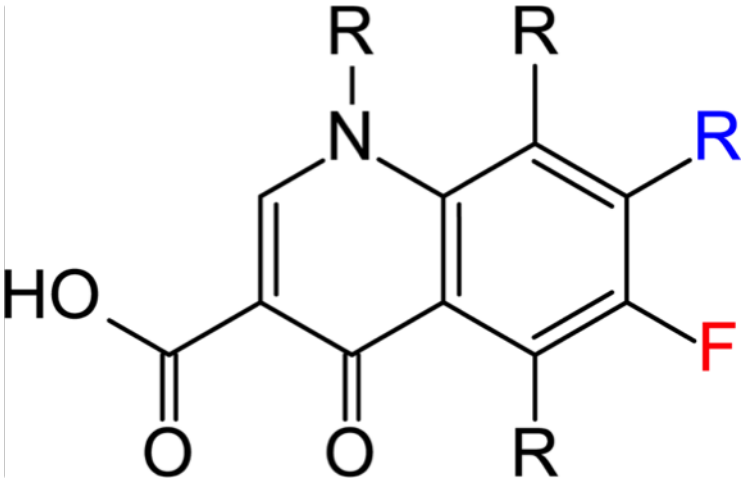

Figure 1.5 Structure of the fluoroquinolone ciprofloxacin. Ciprofloxacin is a fluorinated quinolone that binds to DNA gyrase I and Topoisomerase IV of Gram-negative bacteria.

1.5 Fluoroquinolone mechanism of action and resistance

Fluoroquinolones are a highly successful class of bactericidal antimicrobials because of their action on DNA replication machinery. Quinolones were first discovered in the 1960s, and nalidixic acid, though not strictly a quinolone based on its composition, was the first "quinolone" to be used against bacterial infections. Quinolones are comprised of a two-ringed system involving a nitrogen and a ketone. These were soon superseded by newer generations of fluorinated quinolones, the addition of a fluorine atom attached to the central quinolone ring. There are now four generations of quinolone drugs, with the 'generation' referring to the spectrum of bacteria they are active against[248]. Second generation fluoroquinolones were notable in their activity against Gram-negative bacteria, including *Pseudomonas aeruginosa*, while later generations were capable of treating Gram-positive bacterial infections as well[248,249]. The later generations have a broader spectrum, although there is no systematic measurement of assigning these drugs to a specific generation. Despite the vast number of synthesized fluoroquinolones, not all of them are licensed for human usage by the Food and Drug Administration (FDA) due to levels of toxicity and side effects[250]. As a result, some of the earlier generations of fluoroquinolones, which lack the most toxic effects, are still in use.

1.5.1 Mechanism of action of fluoroquinolones in Gram-negative bacteria

Ciprofloxacin is a second-generation fluoroquinolone bactericidal antimicrobial that acts by binding to bacterial type II topoisomerases DNA gyrase I and Topoisomerase IV in Gram-negative bacteria to disrupt DNA supercoiling. Bacteria supercoil DNA to condense and pack large quantities into the limited intracellular space, and the enzymes DNA gyrase and Topoisomerase IV function by creating temporary DNA breaks to constantly modify the level of supercoiling in response to bacterial needs and environment[251–253]. DNA gyrase I is unique in its involvement in negative supercoiling, a process that can be invoked to relax supercoils in the DNA. This comes into play as bacteria undergo DNA replication and need to release tension that builds in the double-stranded DNA as replication proceeds[254]. DNA gyrase I is an A2B2 heterotetramer comprised of four subunits: two GyrA and two GyrB complexed together. The GyrA subunit is 97 kDa, comprised of an N-terminal domain of 59 kDa and a C-terminal domain of 38 kDa, and the C-terminal domain, which directly interacts with DNA, is involved in substrate recognition and other protein interactions[255]. Topoisomerase IV is a heterotetramer comprised of two repeating units of ParC and ParE, assembled similarly to DNA gyrase. Topoisomerase IV is primarily involved in segregating the catenated DNA rings post-replication.

When bacteria are treated with fluoroquinolones, the drug enters the cell and binds to the interface between the DNA and GyrA or GyrB subunit of DNA gyrase I or ParC of Topoisomerase IV, converting transient double-stranded breaks into permanent ones[253,256,257]. As fluoroquinolones do not inhibit gyrase activity, the initial reaction of bacterial cells is a halt in growth, and cell death occurs when the permanent double-stranded DNA breaks occur.

If there is insufficient fluoroquinolone to bind all the active gyrases, the drug acts as a bacteriostatic antimicrobial. However, if there is an excess of drug compared to binding sites, enough double-stranded breaks overwhelm the SOS response and DNA repair machinery, leading to cell death. In Gram-positive bacteria, the primary target of ciprofloxacin is topoisomerase IV. Despite similarities between human and bacterial topoisomerases, there is a strong preference of fluoroquinolones for prokaryotic topoisomerases, and there appears to be minimal activity of fluoroquinolones against human topoisomerase[258–262].

1.5.2 Chromosomal resistance to ciprofloxacin in Gram-negative bacteria

A key region of the bacterial genome responsible for FQR is known as the quinolone resistance determining region (QRDR)[252,253,263,264]. Mutations in the QRDR within *gyrA*, *gyrB* or *parC* are mainly conserved, with a specific serine substitution, most commonly Ser83Tyr in GyrA, responsible for most resistance in Gram-negative bacteria by reducing the drug binding affinity[252,263]. However, some *Salmonella* isolates with more recent resistance acquisition show mutations at alternative sites within the QRDR (Van Puyvelde, personal communication)[241]. Additionally, there may be other unknown mutations within the chromosome that can contribute to FQR though not within the QRDR.

The degree to which organisms acquire *gyrA*, *gyrB*, *parC*, and *parE* mutations depends on the organism itself. *S*. Typhi isolates have been known to acquire mutations in *gyrA* and *parC* or *parE*, and in many cases a single isolate may carry multiple mutations[191,265]. They typically will not, however, acquire a *parC* or *parE* mutation on its own. This follows *gyrA*, as DNA gyrase appears to be the enzyme more susceptible to poisoning by fluoroquinolones, although there are exceptions to this[265–267]. In contrast, there are few reported cases of *parC* or *parE* mutations in *S*. Typhimurium or other NTS serovars[268]. When mutations arise in *gyrA* in *S*. Typhimurium, some of the most common clinical amino acid substitutions are D87N (aspartic acid to asparagine), S83F (serine to phenyalanine), and S83Y (serine to tyrosine), as these are residues important for quinolone binding[269–273]. There are other amino acid substitutions that may occur within the QRDR, including D87G (aspartic acid to glycine); however, these occur with less frequency in in *S*. Typhimurium. Serine substitutions are thought to be the most common because they disrupt DNA gyrase catalytic activity the least[253,269]. The level of resistance conferred is also dependent upon the specific amino acid change, and this likely explains the preponderance of specific substitutions (Van Puyvelde, personal communication)[272–274].

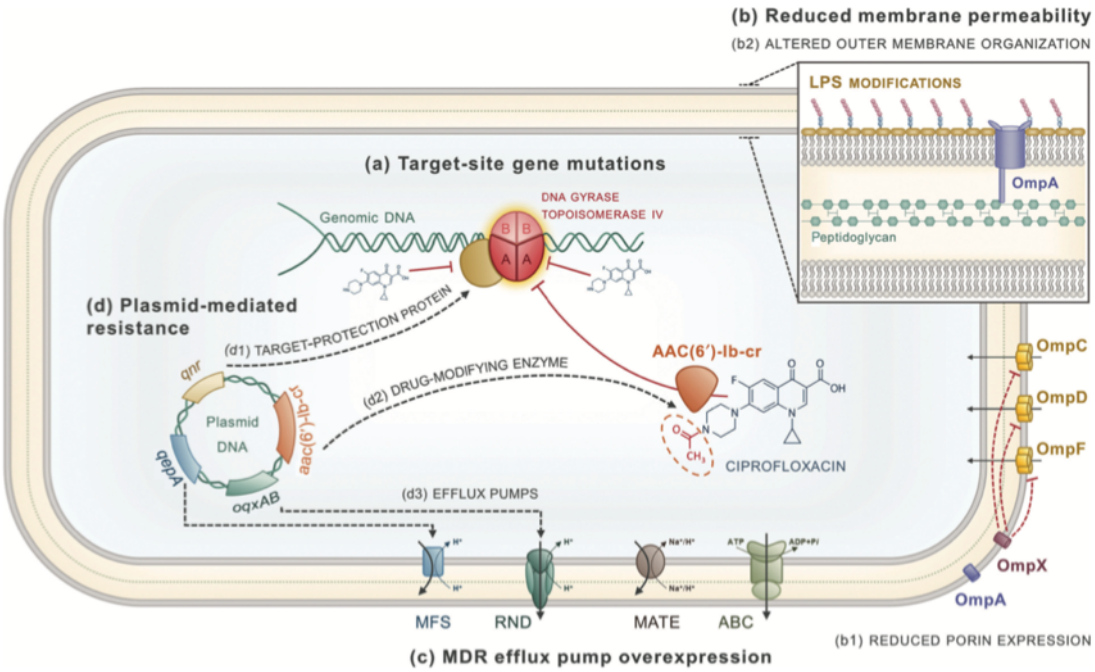

Figure 1.6 Mechanisms of fluoroquinolone resistance. Gram-negative bacteria can employ multiple mechanisms independently or in tandem to resist killing by fluoroquinolones. From Correia *et al.*, 2017[275].

1.5.3 Plasmid-mediated resistance by *qnr* to ciprofloxacin in Gram-negative organisms

In addition to chromosomal mutations in the QRDR, a secondary mechanism of FQR is due to plasmid-mediated quinolone resistance (PMQR). This is currently less common in *S.* Typhimurium, and PMQR is rarely observed independently of chromosomal mutations in the QRDR. However, families of genes known as the *qnr* genes directing the expression of the protein Qnr that can bind to topoisomerase, are responsible for decreasing ciprofloxacin susceptibility in some *Salmonella*[252,253]. The *qnr* genes fit within six different families, which are QnrA, QnrB, QnrS, QnrC, QnrD, and QnrVC, and these proteins of approximately 218 amino acids physically prevent quinolone binding to DNA gyrase and topoisomerase IV[269,276–278]. It has been observed that Qnr proteins are more inhibitory towards ciprofloxacin than nalidixic acid, and *qnrA* has been found to increase the resistance from 12.5-250-fold, while *qnrS* has increased ciprofloxacin resistance by 16-62.5-fold[279–281]. While the presence of *qnr* is understood to elevate the MIC of an organism, it is rare that *qnr* without a chromosomal mutation to DNA gyrase or topoisomerase IV confers decreased

susceptibility to fluoroquinolones[277]. However, there have been a few examples of organisms that have a *qnr* gene on the plasmid but do not have chromosomal resistance[176]. Such cases are highly unusual, and it is not clear why this does not occur more frequently.

1.5.4 Other genes involved in resistance

In addition to *qnr* genes, there are other classes of PMQR genes that are able to break down certain fluoroquinolones and yet other genes that restrict drug entry[282,275]. Gene aac(6')-Ib-cr encodes a bifunctional variant of an aminoglycoside-modifying acetyltransferase. While aac(6')-Ib-cr has primarily been implicated in resistance to aminoglycosides (including kanamycin), two known amino acid substitutions in aac(6')-Ib-cr enable the acetylation of a nitrogen in quinolones, thus reducing the quinolone efficacy and increasing tolerance to fluoroquinolones[275]. Lastly, yet another set of PMQR genes, *oqxAB* and *qepA*, code for a subset of efflux pumps that actively transport fluoroquinolones out of the cells[283]. QepA specifically acts against hydrophilic fluoroquinolones, which includes ciprofloxacin. In contrast, OqxAB acts more broadly against fluoroquinolones and also targets other antimicrobials including tetracycline, chloramphenicol, and trimethoprim[275,278,284,285]. While *oqxAB* and *qepA* are normally not solely responsible for FQR, they can play a key role in reducing drug susceptibility[253]. Of concern is the observation that some PMQR genes may be easily transferred horizontally but have been detected chromosomally, suggesting that these genes may propagate on the plasmid but insert within the chromosome to become fixed in the bacterial population[275,278].

In Gram-negative organisms, there are yet more mechanisms that occur to reduce drug uptake. Because of the LPS barrier in Gram-negative bacteria, drugs are dependent on outer membrane porin channels for cell entry, and bacteria can modulate the size, number, and conductance of porins. Expression of Omp proteins, including OmpF, OmpC, OmpD, and OmpA, controlled by expression of regulator OmpX have been implicated in decreased drug susceptibility. The expression of Omp proteins often occurs in conjunction with an upregulation of efflux pumps, such as AcrAB-TolC, to decrease quinolone uptake and increase efflux. These can be further aided by mutations in Mar, SoxRS and Rob regulons, which can regulate porin, efflux pump expression, and LPS remodelling[252,253,269,286–293].

1.5.5 Mechanisms of resistance found in African *S*. Typhimurium ST313

S. Typhimurium ST313 respond to ciprofloxacin in a similar way to how other Gram-negative bacteria respond. The primary target for ciprofloxacin is *gyrA*, and that is the gene in which the most primary mutations occur. Of the African ST313 isolates that have been analysed, with particular attention to those in sub-lineage II.1 and the emerging sub-lineage II.2 from the DRC, there is considerable variation in the ciprofloxacin MIC as determined by MIC eTest. Of 61 analysed ST313 isolates that have been studied in greater depth, the MICs range from 0.015 (susceptible) to 0.75 (intermediate). Interestingly, the isolate (4930_4) with the lowest MIC harbours a GyrA D87Y substitution, while the isolate (10530_17) with the highest MIC from within this study has an unknown mechanism of DCS (Van Puyvelde, personal communication). There have been documented GyrA substitutions at S83Y, D87N, D87Y, and G81C. Most of the lineage II.2 isolates carrying an S83Y *gyrA* mutation cluster within the MIC range of 0.094 and 0.25. However, even this smaller spectrum is still substantial enough to beg the question of what contributes to these differences. There have also been scant examples of recent isolates carrying a GyrB substitution S464Y, which is within the QRDR (Van Puyvelde, personal communication)[294]. Interestingly, the different substitutions seem to group separately in organisms dispersed on the phylogenetic tree, suggesting that different organisms are capable of withstanding different degrees of mutational stress (Van Puyvelde, personal communication)[176].

There have been at least four examples of a *qnrS* gene carried on the plasmid of an ST313 isolate. The four isolates carrying *qnrS* have all been isolated in the past four years, suggesting that this may be an emerging form of DCS in ST313. Given the ease of horizontal gene transfer, it is highly probable that *qnrS* will become more prevalent in the population over time. The pSTm-ST313-II.1 plasmid also carries *aac* gene, which may influence fluoroquinolone binding[176]. The bias in geographic sampling may mean that there are additional or different resistance mechanisms that have not yet been observed.

1.5.6 Pharmacokinetics/pharmacodynamics of ciprofloxacin

In order to appropriately administer antimicrobials, the rates of absorption, secretion, and excretion(pharmacokinetics) and the effects and mechanisms of uptake (pharmacodynamics) must be assessed. For ciprofloxacin, pharmacokinetic studies undertaken with human volunteers determined that the drug is most efficacious when administered orally twice every 24 hours to achieve a level of perfusion similar to an intravenous dose[295]. 70% of the drug is thought to be bioavailable in adults and 60% is bioavailable in children, and the maximum concentration in the serum is achieved between one and two hours after oral administration[295,296]. In adults, the mean concentration found 12 hours after a 500 mg oral dose is 0.2 μg/ml. For typhoid or infectious diarrhoea, the recommended dose is 500 mg every 12 h for 10 days or 5-7 days, respectively. Ciprofloxacin is a large solute, and after entering the bloodstream, it disperses widely, and tissue concentrations may exceed serum concentrations. The active form of ciprofloxacin is found in multiple tissue locations and secretions, including in the saliva, lymph, sputum, and bile[295]. Despite widespread use of ciprofloxacin, it is still not fully understood how and when cells take up ciprofloxacin and the efficiency of cell absorption. Because *Salmonella* are partly intracellular organisms, this directly affects killing efficiency. Beyond excretion of ciprofloxacin in urine, it has been found that ciprofloxacin is also excreted in sweat, and this has been linked to resistance in *Staphylococcus aureus* resident on the skin within one week of treatment[297].

1.6 Observation of disconnect between phenotypic and genotypic resistance

Thus, when considering phenotype, genotype and phylogeny, not all resistance to the fluoroquinolones in field isolates of *S*. Typhimurium is explained. To fully understanding how bacteria develop decreased susceptibility to stress, including antimicrobials, it is important to recognize the contribution of genetic and epigenetic components. While the primary genetic basis for resistance to ciprofloxacin has been well characterized, there are likely other factors that occur in tandem or instead of that contribute to a decreased susceptibility. This has been observed historically but has not been thoroughly considered or investigated as a substantial contributor to AMR and as a factor that may play a large role in antimicrobial treatment.

Further, there may be species and lineage-specific adaptations that may lead to resistance as some lineages appear more successful in the natural world than others.

1.6.1 Observation of phenotypic heterogeneity in ST313s

The African ST313 clades have been more thoroughly characterized in recent years, and it is known that they are a highly clonal population still undergoing evolution in the context of antimicrobial exposure. While the landscape of DCS has not yet been fully explored, there have been studies looking at genotypic and phenotypic differences between ST313 isolates under serum exposure. These studies stemmed from the observation that highly related ST313 isolates are differentially susceptible to serum killing. Ondari *et al.* exposed six different ST313 lineage II isolates to serum from healthy adults and found that there were three definable phenotypes: susceptible, intermediate, and resistant to antibody-mediated killing. This was an interesting finding because these isolates have few SNP differences, and these known SNPs could not explain the differences in serum susceptibility. Transcriptional analysis of the serum-resistant isolates grown in the presence of serum identified relatively upregulated genes associated with cell surface structures, iron use, and metabolism, while they downregulated genes associated with membrane proteins. These data suggest that there is considerable variation that occurs at the tips of the phylogenetic tree, which may not be due to readily identifiable genetic differences between the isolates[181].

Similarly, ongoing work in ST313s have indicated unique groups of colonial morphotypes when bacteria are grown under biofilm-forming conditions, another form of environmental stress. Bacteria may form biofilms when exposed to harsh environmental conditions. A biofilm is an extracellular matrix composed primarily of polysaccharides, proteins, and extracellular DNA that protect the bacteria from direct assault[298]. Biofilms have been shown to form in various conditions where a community lifestyle is advantageous, such as in adverse conditions outside of a host, on the built environment such as catheters and filters, and as part of the microbiome within hosts[299,300]. D23580 carries mutations that inactivate KatE, a catalase involved in oxidative stress, and BcsG, a cellulose biosynthesis enzyme required for biofilm formation[182]. Together these mutations are implicated in differences in biofilm formation between ST313 and ST19 (lab strain 14028s), but they do not explain variation in biofilm morphotypes within the ST313 sublineage. The genetic basis underlying the

phenotype of varying ST313 biofilm colonial morphotypes has proven elusive, and there may be multiple factors, some of them transcriptional, that explain differences in these morphologies that occur when bacteria are required to form biofilms[176,183].

A further phenotype that has been noticed in *S*. Typhimurium and is under investigation in ST313 isolates is differences in O5 antigen presence or absence, and what that correlates to[301,302]. *S*. Typhimurium is characterized by having O4 and H:I antigens, but not all of *S*. Typhimurium have O5. Interestingly, a recent analysis by Tack *et al.* has identified an increase in O5 negative invasive *S*. Typhimurium, representing 36.9% of the sampled *S*. Typhimurium in the DRC. Notably, there was an increase in O5-negative isolates each year, with much more significant numbers of O5-negative isolates from 2013-onwards[302]. It has previously been shown that variation in O-antigen length can affect macrophage uptake and complement resistance, and a length of > 4 and < 15 repeating units provides greater protection[303]. Furthermore, it appears that O-antigen length is growth phase-dependent, modulated by environmental factors, and this regulation influences serum sensitivity[304]. Additionally, past studies have shown that the presence of O5 plays a significant role in antibody binding and thus host response to infection[18]. In the context of vaccine development that is based on O-antigen specificity and binding, it is essential to understand what these differences in O5 and LPS chain length mean, how they are regulated, and how they can be detected.

1.6.1.1 High-throughput genotyping as method to distinguish between related organisms

Over the past two decades, whole genome sequencing (WGS) has become a powerful method to investigate differences between related organisms, and with the steady decrease in cost of WGS, it has become commonplace to sequence bacterial isolates on a large scale as part of surveillance and to understand population structure. The power of this technology is that it enables detection of minute differences between related organisms, such as SNPs, which may contribute to virulence or drug resistance. This has already been undertaken to great effect in many organisms, including in *S*. Typhimurium ST313. Okoro *et al.* showed significant differences between ST19, ST313 lineage I and ST313 lineage II bacteria using phylogenetic and BEAST analyses to pinpoint nodes of divergence and where in the genome these differences arose[82]. Similarly, Van Puyvelde *et al.* have been able to identify and

investigate a novel sub-lineage of ST313 bacteria that differ from lineage II using WGS of over 500 ST313 isolates. These data have revealed SNP differences and pseudogenization that indicate a greater degree of host-specificity and azithromycin resistance. However, a limitation of WGS, even when it is done for a high volume of samples, is that it cannot provide any insight of how bacteria are behaving in real-time to environmental pressures; it can only provide a prospective snapshot. Additionally, given the transcriptional and post-transcriptional modifications that occur, it is likely that some of the differences in bacterial phenotype cannot be accounted for by the genome sequences alone, and a combination of high-throughput genotyping and high-throughput phenotyping may provide greater insight.

1.6.1.2 Introduction to high-throughput phenotyping to understand AMR

In contrast, phenotyping of organisms allows scientists to examine bacterial behaviour under growth and/or duress to understand any underlying mechanisms of persistence and survival. Historically, this has been done at small scale because of technological limitations. However, with the advent of newer technologies, it is now possible to undertake large high-throughput screens to investigate multiple bacterial isolates simultaneously. There are a few different levels of phenotype that can be measured, and depending on the question, they can be used to great effect. An area of study that could significantly benefit from the use of high-throughput phenotyping is AMR measurement, because present methods of detecting, validating, and applying AMR information are often low-throughput, time-consuming, and often not clinically useful[305,306].

One method of high-throughput phenotyping is RNA sequencing (RNA-seq) of bacteria to investigate the changes in gene expression that occur upon antimicrobial exposure. Such information is extremely useful for determining the conditions under which antimicrobials might be used and concentrations that are effective, and it can be employed to understand nascent resistance mechanisms, especially when linked to genomic information[307]. A limitation of direct measurements on bacteria is that they are growing outside of the context of the host.

A second method is high-throughput metabolic profiling of bacteria. This has been carried out extensively to understand how bacteria respond to different nutrient environments and metabolic pathways associated with biosynthesis, metabolism, virulence, and AMR have

been investigated[308,309]. The power of such a method is that one can use known interactions and phenotypes to help identify unknown interactions based on similarity of patterns. Such phenotyping has been done using the Biolog assay to quantify cellular respiration of bacteria under different metabolic conditions[176,237,310]. This methodology paired with advances in analysis pipelines has opened doors to probing interactions and patterns at greater depth. An example of this is the extensive mapping of the fitness of an *E. coli* mutant library under 324 different conditions, including antimicrobial perturbation, which revealed insights into drug mechanisms and gene essentiality[311]. Another successful example from Yang *et al.* used network modelling and biochemical screening to flag pathways and mechanisms that determine antimicrobial lethality[312].

Yet another powerful method is high-throughput imaging to assess changes in morphology. With high-content microscopes that can clearly image wells of bacteria at high-resolution to distinguish individual bacteria, it is now possible to analyse single bacterial cells. The power of this technology is that it can be harnessed to assess how a given isolate responds to antimicrobial treatment at a given concentration, to study how genotypically similar isolates respond to the same treatment, and to follow antimicrobial treatment over time. While some of these questions have been addressed in the past, the combination of automated high-throughput screening and automated analysis pipelines that can distinguish morphological features allows for a new depth of investigation and understanding. In particular, these advancements enable large-scale assessment of how clinically relevant bacteria respond to standard antimicrobials in a diagnostic context. For example, a screen that can rapidly quantify growth of a clinical sample in the presence of various antimicrobials to help determine the best course of treatment would be valuable in a clinical setting. Some of this information would be obtainable via sequencing and standard laboratory diagnostics; however, observing morphological changes of the bacteria over a period of time would help predict how a given isolate and population might respond over the course of treatment and thus help clarify appropriate dosage and single or combination therapies. Additionally, these technologies can and have been used to screen for novel therapeutics and development of resistance to them as well as track development of resistance to existing antimicrobials[313,314]. Even with low-resolution microscopy, increased analytical power and methodologies have opened doors to gleaning increasing amounts of information from phenotypic data[315].

1.6.2 'Adaptive resistance' and desensitization to antimicrobials

Such phenotyping methods as described above open the door to exploring how bacteria adapt in real-time to an environmental stressor, specifically measuring how they change morphologically over time. This enables the study of adaptive resistance to antimicrobials, a concept that has not yet been as well-characterized as mutational resistance. It is understood that in a heterogeneous culture, given a large enough population of bacteria, there are likely to be some that carry mutational resistance to a given antimicrobial (or other stressor)[316]. However, what is less-well understood is how bacteria can develop resistance, or rather, 'desensitization' to a treatment in a way that is not obviously due to mutation. It has been observed that when bacteria are grown in the presence of theoretically lethal concentrations of antimicrobials, there is typically an initial death phase, which may be followed by a rebound growth phase[317,318]. However, the mechanisms contributing to this phenotype have not been studied in depth. One study proposed that a decoupling of transcription and translation in the cells leads to uncontrolled cellular division, resulting in a surge in bacterial counts after an initial death phase[317]. Overall, while this phenotype has been observed, most have not delved into its significance and the role it may play in bacterial escape of antimicrobial killing[318,319]. It is possible that the explanation for this lies in bacterial desensitization to the antimicrobial via transcriptional means, and part of this may occur from phase variation stochastically during growth or by DNA inversion[320,321]. The serum resistance in ST313 isolates observed by Ondari *et al.* may similarly be explained by this phenomenon, especially as there were no SNPs that accounted for the differences in serum sensitivity and resistance between isolates[181]. However, at the moment that remains unproven.

1.6.2.1 Mechanisms of adaptive resistance to stress in *S.* Typhimurium

Some discussion of adaptive resistance was mentioned earlier when considering the mechanisms of FQR. One of these mechanisms is upregulation of drug efflux pumps and alteration in membrane permeability that help expel drug[322]. This is an example of adaptive resistance because the bacteria have not acquired any mutations during the course of treatment, but they have changed transcription of existing machinery. One prominent mechanism that seems to enhance bacterial survival in the presence of antimicrobials is the induction of prophage. Prophages are latent viral sequences that insert into the genomes of bacteria or

exist as a plasmid and are closely associated with specific serovars. When bacterial growth is normal, these viral genomes passively reside within the bacterial genome and are consistently replicated as part of the bacterial replication process. However, bacterial cell damage initiates viral excision from the bacterial chromosome, and phage replication occurs via the lytic cycle. Viruses replicate and escape the stressed bacterium, and the cycle begins anew in a new host[323]. However, in the context of fluoroquinolone treatment, this process further amplifies the stress (SOS) response of the bacteria. Prophage induction increases transcription of prophage cell division inhibitors *kilR*, *dicB*, *dicC*, and *dicF*, amongst other stress response genes including *abc2*. This process additionally triggers the bacterial SOS response using a *lexA*-dependent pathway, which involves activation of the DNA repair gene *recA* following binding to single-stranded DNA fragments, and the DNA repair processes that follow may enhance bacterial survival under this stress[323–325]. These include upregulation of error-prone DNA polymerases UmuC and UmuD that can replicate DNA more quickly but with less accuracy as well as excision repair proteins UvrAB and the inhibition of cell division by SulA[325].

1.6.2.2 Persistence and tolerance. Differences between persistence, tolerance, and 'adaptive resistance'

Antimicrobial treatment with sub-inhibitory concentrations has been linked to persistence of bacterial populations. Persistence, though a complicated and contentious term, implies survival of bacteria in the presence of a given potentially lethal stressor. The classical understanding of persistent bacteria is that they enter a dormant or less metabolically active form while exposed to the stressor and return to normal function once the stressor has been removed[326,327]. This response is not due to genetic changes in the bacteria but transcriptional ones that allow survival in adverse conditions including antimicrobial exposure by downregulating porins and genes that are involved in drug influx, limiting replication and transcription, reducing metabolism, and the employment of toxin-antitoxin systems[327]. Microscopy of such organisms has revealed that they look morphologically distinct to non-persisters and are able to rebound as a population after the stress has been removed[328,329].

It is important to also mention the concept of drug-tolerant bacteria when discussing persistence. Tolerance to an antimicrobial implies the bacterium's ability to protect themselves during exposure to the drug. Such organisms are able to not only survive but replicate and

expand. One way in which bacteria manage is through the formation of a biofilm. Producing this extracellular matrix shields the bacterial population from contact with the antimicrobial, and these bacteria are phenotypically distinct from bacteria growing planktonically[329,330]. Importantly, these bacteria, other than expending energy and resources in creating the biofilm, do not have to directly respond to the effects of the antimicrobial. However, there are likely costs associated with forming and maintaining a biofilm that may outweigh the benefits of such a strategy.

In light of the costs and benefits each survival strategy requires, it is likely that bacteria employ a system of bet-hedging or gambling to ensure survival of at least part of the population. It has been hypothesized that bacteria use a form of game theory to spread the risk of the entire population dying[331]. This has been termed 'bet-hedging', and in the context of antimicrobial exposure, it is likely that some bacteria form a biofilm to stave off contact with the antimicrobial, some bacteria go into a persister state by downregulating or turning off any non-essential processes, and some directly encounter the antimicrobial[320,330]. One can imagine that one of these strategies may be more successful than the others in a given context. For instance, if the drug exposure is long term, the persister population may not survive because eventually it will need to metabolize nutrients to survive, and it may uptake and be killed by the drug at that point. In an alternative scenario, forming a biofilm may prove too energetically expensive or expose the population to other harm. In yet another scenario, the bacteria directly exposed to the drug may not be able to survive the dosage, or replication in the presence of the antimicrobial may lead to lethal mutations downstream. Given the likelihood of any of these challenges, the strategy of bet-hedging with various populations may help ensure bacterial survival during antimicrobial treatment and may explain phenotypic heterogeneity seen in bacterial populations.

1.6.3 Lack of understanding of additional factors involved in FQR in *S*. Typhimurium

As stated above, it is plausible that bacteria deliberately stratify their response to antimicrobials to enhance survival. However, it is still unclear what the precise and predominating epigenetic strategies are when *S*. Typhimurium are exposed to fluoroquinolones. Beyond mutational resistance, there is insufficient understanding of what phenotypic changes directly

contribute to bacteria resisting killing by fluoroquinolones. It is likely that some combination of strategies is employed, and it is important to know which to implement appropriate FQ usage. In addition, because mutations in *gyrA* are the primary and obvious changes that confer DCS, it is possible that there are concurrent SNP changes that occur in the bacteria that may influence the level of susceptibility but that are not obvious without a deep sequencing analysis. In a changing landscape of ciprofloxacin presence, it is important to understand how those additional factors may play into compounding resistance.

1.6.3.1 Transcriptional response to FQs

As described above, there are significant phenotypic differences between ST313 isolates that appear genetically clonal. One way of probing this further is by investigating transcriptional responses to fluoroquinolones, and this has been done to some extent. Transcriptional studies of bacteria exposed to ciprofloxacin have shown that pathways involved in stress response, solute and drug transport, DNA repair, and phage induction are upregulated, which can increase error-prone DNA replication and bacterial resilience in the face of antimicrobials[332–337]. Li *et al.* showed that in a *Caenorbitans elegans* model with ciprofloxacin-exposed *S*. Typhimurium, many genes involved in the stress response and DNA damage are upregulated. Furthermore, studies in *Salmonella*, *E. coli*, Staphylococcus aureus, and *Acinetobacter baumanii* have shown that stress from fluoroquinolone exposure induces a strong prophage response, and this in turn positively reinforces the bacterial stress response[325,338–340]. Furthermore, work in *A. baumanii* has shown that treatment with colistin results in an upregulation of genes such as *acrB*, *emrB*, and *mexB*, which are involved in drug efflux[341]. While colistin is not a fluoroquinolone, these data may help clarify some of the genes and processes involved in general antimicrobial evasion strategies.

1.6.3.1.1 Overview of bacterial transcriptomics to understand drug resistance

In general, the study of bacterial transcriptomics has begun to gain in popularity as a method to study AMR. This is especially in light of the concern that genomics cannot provide a complete picture of everything that occurs in the bacteria in response to stress. While transcriptomics has not yet been deployed to study and determine AMR in real-time, it can

be a powerful tool in understanding individual genes, networks, and pathways in bacteria that contribute to resistance. As previously mentioned, whole genome sequencing alone cannot capture all of the mechanisms of resistance that bacteria might employ, and transcriptomics can help bridge that gap. This has been demonstrated successfully by Boinett *et al.* in elucidating the many efflux pumps and pathways that colistin-resistant *A. baumanii* use[341]. Similarly, Siqueira *et al.* were able to demonstrate the advantages of combination therapy for MDR *P. aeruginosa* rather than dosing with either meropenem or ciprofloxacin by measuring the transcriptional response of bacteria to the antimicrobial treatments[342]. The advantage of transcriptomics is that if bacteria exposed to antimicrobials are sampled longitudinally or at discrete points during antimicrobial exposure, this can provide considerable insight into how the bacterial population is responding over time to treatment, especially if much of the response is epigenetic rather than genetic.

1.6.3.2 How bacteria respond to antimicrobial exposure longitudinally

Despite the testing that antimicrobials undergo before they are licensed for human use, there is still a dearth of information about how bacteria respond to these antimicrobials, especially over longer periods of time. When antimicrobials are licensed for humans, the majority of assessments involve safety, side effects, and efficacy. However, they do not involve extensive study of the effect on the bacteria being treated, and any studies on antimicrobial efficacy would focus on the impact of the drug within the host. Given the current climate of excess antimicrobials in the environment that are discharged through sewage, in agriculture, and in clinical environments, it becomes essential to understand how bacteria might interact with and develop resistance to antimicrobials independent of the host[343,344]. Considering pathogens like *S.* Typhimurium that can exist in the environment and are transmitted via faecally-contaminated food and water, there is an acute need to explore the interaction with ciprofloxacin outside of the human host. Often, the methodology of studying this involves treating bacteria with an antimicrobial and then measuring growth over a designated period of time. One well-established method is performing growth curves of bacteria in nutrient medium containing the antimicrobial and measuring the optical density (OD600) of the bacteria as a readout for growth[345]. An alternative method that is used extensively to study antimicrobial efficacy is to use time kill curves (TKC) follow bacterial growth dynamics over a period of time and count colony forming units (CFU) at each timepoint[346–348]. Both of these methods provide information on the growth dynamics of bacteria in the presence

of an antimicrobial and can be used to compare against growth dynamics when bacteria are not exposed. Studies of ciprofloxacin treatment of various Gram-negative bacteria have shown that at high (> 1x MIC) concentrations of ciprofloxacin, bacteria are initially killed by the antimicrobial but that there can be subsequent growth after six hours that has not been properly characterized[319,342,349,350]. Additional information can be gleaned using some of the methods discussed previously, including whole genome sequencing, RNA-seq, and microscopy. However, few (if any) of the above techniques are routinely used to describe bacterial growth longitudinally.

Specific to agents that act on DNA replication machinery, it is known that one of the early responses is cellular elongation. This occurs because when ciprofloxacin forms a complex with DNA gyrase and the DNA, this causes permanent double-stranded DNA breaks, which stall DNA replication. The stalling of replication causes a cascade of downstream reactions, specifically triggering the SOS response and DNA repair pathways. One of the effects is the halting of bacterial septation by cell division inhibitor SulA by the inhibition of cell division protein FtsZ[351–353]. As a result, filamentous, elongated bacterial cells form, which have been observed to grow up to 200 μm in length and can contain multiple copies of the chromosome along their length[352,354–357]. It has been observed that these filamentous bacteria can subsequently divide into daughter cells once growth conditions are less adverse, although the dynamics and mechanics of this occurrence are as yet unexplored[358]. There is an urgent need for further understanding of how bacteria can recover from ciprofloxacin exposure and what that means for subsequent resistance.

1.6.3.2.1 Imaging to observe killing over time

In addition to monitoring changes in growth by microbiological assays, key insights into bacterial response to antimicrobials come from morphological changes that can be captured via microscopy. Multiple studies have shown that it is possible to identify bacteria on the basis of their morphology when treated with antimicrobials[359,360]. This is particularly true of ciprofloxacin-treated bacteria due to the clear filamentation that occurs in affected cells. Multiple groups have shown the utility of single-event and time-lapse microscopy to screen for antimicrobial efficacy for specific bacteria, and this technique may also be an important way of identifying and appreciating the contribution of subpopulations to the overall phenotype. For example, Ungphakorn *et al.* were able to demonstrate similar results

between time kill curves and time-lapse microscopy of carbapenemase-producing *Klebsiella pneumoniae* and *E. coli* when treated with various colistin combinations[361]. Additionally, Barrett *et al.* were able to visually track *E. coli* response to high concentrations of ofloxacin and identify septation events in filamentous bacteria and link those morphological changes to error-prone DNA polymerase V UmuDC and activation of the SOS response[362].

1.6.3.2.2 Use of imaging, kill curves, and transcriptomics to tie together phenotype and genotype

Ultimately, to capture the greatest amount of information about and fully understand bacterial response to antimicrobial exposure, it is preferable to combine multiple approaches. With the advent of higher-throughput technologies and techniques, this multi-pronged approach is becoming increasingly feasible. Using multiple technologies, Pribis *et al.* were able to exhibit the role of a bacterial subpopulation of *E. coli* that formed in response to ciprofloxacin-induced stress by cell-sorting bacteria on their morphological characteristics, capturing growth dynamics, and identifying genes that were differentially expressed between ciprofloxacin-exposed and unexposed bacteria[331]. It is expected that more such dynamic workflows will emerge.

1.7 Summary

As fluoroquinolones, particularly ciprofloxacin, continue to be used widely, there is an increasing need to understand how this affects the fitness landscape in *S*. Typhimurium and is driving further AMR. As fluoroquinolones disrupt DNA replication, it is possible that FQR has other impacts on the bacterial cell, potentially modifying metabolism, resistance to heavy metals and other antimicrobials, invasion of host cells, and improving survival in hostile conditions[363–367]. Therefore, it is more important than ever to better understand how clinically-important bacteria are responding to ciprofloxacin.

Thanks to technological advancements, it has become easier in recent years to collect more comprehensive linked genotypic and phenotypic information. Consequently, in this project

we intend to perform a combination of RNA-seq and imaging analysis to develop our understanding of FQR. Multiple studies have shown that it is possible to identify bacteria on the basis of their morphology when treated with antimicrobials[359,360]. This is particularly true of ciprofloxacin-treated bacteria, where bacteria become filamentous and elongated when treated.

Beyond characterising treated versus untreated bacteria, linked imaging and expression analysis, performed on an Opera Phenix high-throughput microscopy platform, enables classification of multiple phenotypes in a stable 96/384-well system. In this system, we can identify live versus dead bacteria and potentially even finer characteristics such as membrane permeability, the bacterial stress response, and analysis of persister populations. The Opera Phenix is normally used to study eukaryotic cells, but it has not previously been validated as a tool to study bacteria, which we hope to do[368,369].

While some transcriptional analysis has been undertaken on the *Salmonella* response to antimicrobials, little RNA-seq analysis has been performed on MDR clinical isolates of *S.* Typhimurium to look at the transcriptional effect of drug treatment, especially comparatively across a variety of isolates and drug conditions[333]. The combination of RNA-seq and high-resolution imaging on isolates could provide considerable insight into the different effects of ciprofloxacin treatment on distinct *S.* Typhimurium isolates growing in a controlled environment. Combining this with growth assays and genomic information may further disentangle bacterial response to ciprofloxacin by distinguishing between genetic and epigenetic features. Ultimately, further insight into *S.* Typhimurium response to ciprofloxacin may help guide future antimicrobial stewardship and development of targeted vaccines, especially against iNTS isolates prevalent in sub-Saharan Africa[51].