Colonisation of the gut microbiome by Escherichia coli during international travel

UNIVERSITY OF BIRMINGHAM

Maastricht University

Colonisation of the gut microbiome by *Escherichia coli* during international travel

DISSERTATION

To obtain the degree of Doctor of Philosophy at the University of Birmingham
on the authority of the Vice-Principal, Prof. Stephen Jarvis
and the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus,
Prof.dr. Pamela Habibović
in accordance with the decision of the Board of Deans,
to be defended in public
on Wednesday, 17th of April 2024 at 13:00 hours

by Matthew Anthony Davies

Supervisors:
Prof. dr. John Penders, Maastricht University
Prof. dr. Alan McNally, University of Birmingham

Co-supervisors:
Dr. Petra F.G. Wolffs, Maastricht University
Prof. dr. Willem van Schaik, University of Birmingham

Assessment Committee:
Prof. dr. Christian J.P.A. Hoebe (chair Maastricht University)
Dr. Michael J. Cox, University of Birmingham
Prof. dr. Joan A. Geoghegan, University of Birmingham (chair University of Birmingham)
Prof. dr. Daisy M.A.E. Jonkers, Maastricht University
Prof. dr. Nina M. van Sorge, Amsterdam UMC
Dr. Donal M. Wall, University of Glasgow

Abstract

There is a high risk of acquiring extended-spectrum beta-lactamase-producing *Enterobacteriaceae* (ESBL-E), predominantly *Escherichia coli*, among international travellers visiting antimicrobial resistance hotspots. Although factors that drastically perturb the gut microbiome — e.g. antibiotic use and travellers' diarrhoea — have been shown to influence the risk of ESBL-E acquisition, it remains largely unknown whether successful colonisation of ESBL-E during travel is associated with the composition, functional capacity and resilience of the traveller's microbiome. Strains of *E. coli* often vary by geography and pathogenicity and it is unknown how their inter- and intra-species interactions associate with the acquisition of ESBL-E. The research in this thesis aims to address these problems by determining how involved the entire gut microbiome is in resisting colonisation from invading ESBL-E, or if species interactions are more important. Faecal samples from before travel and immediately upon return from Africa or Asia of 179 Dutch travellers had previously been collected and whole metagenome shotgun sequenced. Here, the gut microbiome was profiled with a metagenomics species concept approach and the microbial composition, population diversity and functional capacity were determined and how these associate with the acquisition of ESBL-E during travel was calculated. None of the measurements in pre-travel samples were predictive of risk of acquiring ESBL-E. When compared to post-travel measurements, there were also no longitudinal changes specific to acquiring ESBL-E, instead only significantly associating with the onset of travellers' diarrhoea. How individual taxa are associated was then determined, where it was discovered that the prevalence and abundance of *Citrobacter freundii* and two members of the genus *Bacteroides* increased in those remaining uncolonised by ESBL-E, suggesting that these taxa may have some capability of outcompeting ESBL-E. This was tested by growing both commensal and ESBL-producing *E. coli* in nutrient broth in the presence of travel-acquired *Citrobacter*, but *Citrobacter* could not outcompete *E. coli* in a direct manner in nutrient rich media.

As the longitudinal change to *E. coli* was not significantly associated with the acquisition of ESBL-producing *E. coli*, it was hypothesised that strain replacement was

occurring. Individual strains were assembled into metagenome assembled genomes (MAGs) and compared to reference genomes that represent the total *E. coli* strain diversity. The population of *E. coli* strains was drastically altered during travel, but this was independent from the factors driving the acquisition of ESBL-E (and perturbing the gut microbiome). The results from this thesis highlight the potential of individual species to interact with invading ESBL-E, but most importantly focusses future research towards avenues of altering a person's colonisation resistance as an intervention strategy to minimise the spread of multidrug resistant bacteria.

For Mum and Dad

Contents

List of Figures

List of Tables

Chapter One

General Introduction

1.1 *Escherichia coli*

Escherichia coli has become the most widely used model organism in microbiology after its discovery by Theodor Escherich in the 19th century. *E. coli* usually resides as a commensal member of the gastrointestinal tract of humans, cattle and birds across the world, but can also be detected in environmental samples. Various pathotypes - groups of *E. coli* that display an ability to infect a specific host tissue - exist, and these are split into two main groups: intestinal pathogenic *E. coli* (InPEC) that usually cause diarrhoea; or extra-intestinal pathogenic *E. coli* (ExPEC) that usually live in the gut microbiome as commensals [1] until the opportunity arises to colonise and infect other sites like the urinary tract or bloodstream [2]. The host range, ability to colonise various tissues and global prevalence of *E. coli* has led to a complex genomic structure allowing adaptation to multiple niches.

E. coli can be split into eight phylogenetic groups (phylogroups; A, B1, B2, C, D, E, F and G) that are genetically distinct from each other [3, 4]. Additionally, five phenotypically indistinct but genetically distinct groups of *E. coli* designated as 'Cryptic Clades' have been discovered [5], although some of these may actually be separate species of *Escherichia* [6]. There is also within-phylogroup divergence of *E. coli* into sequence types (STs), lineages that are identified by their unique allelic differences in eight housekeeping genes [7]. These are commonly distinguished by PCR, but fully *in silico* methods focusing on the genes targeted by multilocus sequence typing (MLST) PCR are now employed to classify strains using sequencing data into STs [8].

The virulence factors expressed by *E. coli* are one adaptation for infecting different types of cells and host environments, and some pathotypes infecting the intestinal tract can be distinguished by the mechanism of infection they use. The common intestinal pathogen Enteropathogenic *E. coli* (EPEC) utilises attaching/effacing machinery to infect the intestinal epithelial cell layer. After adhering, the type III secretion system (T3SS) is used to inject effector molecules that induce the epithelial cell to alter its cytoskeleton and create a projection known as a pedestal, disrupting host cell function and inducing cell death [9]. Similar to EPEC are Enterohaemorrhagic *E. coli* (EHEC), but with a major difference that they produce shiga toxins that induce apoptosis of cells with shiga toxin receptors, which in humans are predominantly the Gb_3 receptor found mainly on the cell surface of epithelial cells of the intestines and kidneys [10]. Strains of EHEC are typically found in cattle where they asymptomatically colonise due to cattle lacking these shiga toxin receptors [11]. However, colonisation in humans can lead to haemolytic uraemic syndrome (HUS) where the kidneys are damaged or inflamed and may cause kidney failure. Another major intestinal pathotype is Enterotoxigenic *E. coli* (ETEC) which is the most important cause of travellers' diarrhoea. These *E. coli* strains utilise colonisation factors, flagella and other outer membrane proteins for specialised attachment to host intestinal epithelial cells where they interrupt the cell's normal function and induce diarrhoea in the host via production of a heat labile toxin similar to Cholera toxin [12]. These examples highlight the significance of *E. coli* within intestinal infections.

Infections in non-intestinal sites are also a major burden to human health. Uropathogenic *E. coli* (UPEC) are the cause of the majority of urinary tract infections (UTIs) as they are adapted to adhering to and entering bladder epithelial cells in order to avoid clearance during urination. Successful colonisation of the bladder by UPEC can cause cystitis, but travelling upwards through the ureter can cause kidney infections [13]. It is from sites like the kidneys that *E. coli* can access and infect the bloodstream, causing urosepsis. *E. coli* is now a leading cause of sepsis, especially in neonates, infants and the elderly [14]. *E. coli* is also frequently responsible for neonatal meningitis, second only behind Group B *Streptococci* [15].

The ability of *E. coli* pathotypes to spread to multiple hosts is relevant at a global scale. The shiga toxin-producing *E. coli* O157:H7, usually associated with contaminated food, was responsible for major outbreaks of HUS across Japan [16], the United States [17] and Europe [18]. Another famous outbreak is the epidemic strain O104:H4, another shiga toxin-producing *E. coli*, that infected thousands of people across

Europe and was traced back to fenugreek sprouts grown in Germany [19]. The most globally relevant clone of ExPEC is lineage ST131 which is referred to as a pandemic strain of *E. coli*. ST131 is a globally disseminated clone of *E. coli* [20] which is genetically distinct from other ExPEC and can itself be separated into three clades (A-C). The most significant subgroup of ST131 was group C which benefited from evolving fluoroquinolone resistance and acquisition of a plasmid containing extended spectrum beta lactamase (ESBL) activity from the gene $bla_{\text{CTX-M-15}}$ [20]. These are just some of the factors that led to its dominance worldwide and becoming one of the most commonly encountered causes of AMR infections [21].

The AMR activity of ST131 group C is largely conferred by a multidrug resistance (MDR) plasmid acquired via horizontal gene transfer (HGT). HGT is the process of acquiring new DNA from external sources and can benefit the recipient with the access to new genetic traits like resistance to antimicrobials. The three most important mechanisms of HGT are conjugation, transformation and transduction. Conjugation is the most common method of sharing DNA between bacteria and works by transferring a copy of plasmid DNA to a neighbouring cell with which it is in direct contact. The plasmid being transferred usually self-encodes for the conjugation machinery that facilitates its transmission [22], but mobile genetic elements that contain their own conjugation machinery can also be found in the chromosome and are referred to as integrative and conjugative elements [23]. Antimicrobial resistance genes (ARGs) can also spread through other HGT mechanisms, like transformation where competent cells take up free DNA and integrate it into their chromosome through homologous recombination [24]. The ability to take up free DNA is not a constitutively expressed mechanism; it may be induced during DNA damage in bacteria that do not have a sufficient SOS response and this can be triggered by exposure to antibiotics [25]. Bacteriophage-mediated transmission of genetic material, transduction, is the third way in which organisms acquire foreign DNA and bacteriophages have been detected containing various ARGs [26].

1.2 Mechanisms of antimicrobial resistance

For antibiotics to be effective at halting bacterial growth or at killing the organism, they need to access their target. Bacteria can encode a number of resistance mechanisms that prevent antibiotics from physically interacting with its cellular components. Porins act

as doors for the movement of molecules and are the entry point for various antibiotics, so modulating the number or efficiency of porins can reduce the intracellular or periplasmic concentration of antibiotics and confer resistance [27]. Similarly, the intracellular concentration of antibiotics can be actively lowered by forcing them out of the cell via the use of efflux pumps [28]. Efflux pumps are linked to the plasmid-mediated spread of AMR, as a low intracellular concentration of antibiotics allows for plasmid DNA to be immediately expressed upon entry [29], of which itself may contain genes for more efflux pump machinery [30].

Instead of controlling the level of antibiotics accessing the cell components, the target molecules of antibiotics can be modified to reduce the efficacy at which the two interact. For example, mutations to the gene encoding DNA gyrase, an enzyme vital to DNA replication and the target of antibiotics like quinolones, alter the target site of the antibiotics significantly enough to prevent their bactericidal ability [31]. Targets other than genes can also be modified to protect the bacterium. For example, the antibiotic polymyxin kills Gram-negative bacteria by destabilising their lipopolysaccharide (LPS) layer and causing the intracellular contents to leak into the extracellular matrix [32], but the LPS can be modified in a way that reduces polymyxin's affinity to it [33]. This method of resistance is also mobilisable and can be spread via plasmids [34]. Although the antibiotic target can be modified, some resistance mechanisms skip the necessity of this and directly inactivate antibiotics. A widespread example of this are the ability of aminoglycoside modifying enzymes to catalyse the modification of aminoglycosides and reduce their antibiotic properties [35]. Alternatively, antibiotics can be inactivated by hydrolysis as this completely decomposes the molecule. One of the most significant cases of this is the deactivation of beta-lactams via a family of enzymes known as beta-lactamases.

1.3 Beta-lactamase classification

As the cause of some of the most globally relevant resistance traits, beta-lactamases are a diverse group of enzymes that hydrolyse the beta-lactam ring of various antibiotics. Originally, beta-lactamases were separated into classes A-D by molecular classification of their amino acid sequence, but this was improved upon with the addition of associating each enzyme to their clinical phenotype by including substrate and inhibitor profiles [36]. This resulted in three broad groups that generally follow the original

structural classification, but with functional classification of the enzymes into subgroups and better associating clinical isolates with their observed level of resistance. Group 1 beta-lactamases primarily includes cephalosporinase enzymes which most effectively hydrolyse cephalosporins; group 2 has by far the largest number of unique enzymes that confer resistance to most beta-lactam antibiotics and are referred to as extended-spectrum beta-lactamases (ESBL); and group 3 contains the metallo-beta-lactamases which are structurally the most unique enzymes as they require a zinc ion at their active site [36].

As well as having the most diversity of enzymes and conferring resistance to the highest number of antibiotics, group 2 ESBLs are also the most common in clinical infections. The enzymes most responsible for this are TEM, SHV, OXA and CTX-M; with CTX-M being especially relevant as its variants have seen a drastic expansion across the globe and are also frequently found in community settings [37].

The ARGs encoding for the CTX-M enzymes ($bla_{\text{CTX-M}}$) likely originate from the genus *Kluyvera* [38] where they were mobilised from the chromosome and into plasmids. This was so successful that CTX-M enzymes displaced most other ESBL enzymes [39]. Mobilisation of $bla_{\text{CTX-M}}$ seemingly occurred because of insertion sequences (IS) mediating the rearrangement of genetic material, as evidenced by the IS*Ecp1* being upstream of both $bla_{\text{CTX-M-14}}$ [40] and $bla_{\text{CTX-M-15}}$ [41]. IS*Ecp1* is especially efficacious at spreading ESBL genes, as proven by $bla_{\text{CTX-M-14}}$ and $bla_{\text{CTX-M-15}}$ being the most prevalent and widely disseminated $bla_{\text{CTX-M}}$ genes [42]. Its success can also be attributed to it additionally providing a promoter for the resistance gene it is transferring [43]. The plasmids that harbour the most widespread $bla_{\text{CTX-M}}$ genes are predominantly F-plasmids which are usually specific to *Enterobacteriaceae* [44].

There is more to the global dissemination of $bla_{\text{CTX-M}}$ genes than just a particularly efficient IS element. Aside from the conjugation of bacterial plasmids like those that belong to group F, phage-like plasmids have also been observed carrying $bla_{\text{CTX-M}}$ genes suggesting that bacteriophages have had a role in transferring these resistance genes [45]. The previously mentioned *E. coli* lineage ST131, and therefore the $bla_{\text{CTX-M-15}}$ gene it usually contains, became globally disseminated from several factors like its fluoroquinolone resistance, large number of virulence factors and efficient fimbriae [20]. Human-associated factors are also a major driving factor. Antibiotic usage in animal production led to an expansion of plasmids containing ESBL genes in the

animal reservoir, which increases the chance of transfer to human-host *E. coli* [46]. The physical movement of humans during international travel is a significant factor in the spread of *E. coli* harbouring $bla_{\text{CTX-M}}$ genes, especially when visiting regions of Africa and Asia where there are higher community rates of carriage of ESBL-producing *E. coli* [47]. These factors together show how ESBLs have become one of the most common types of AMR mechanisms and emphasises the importance of *E. coli* as a driver of this.

1.4 Development of a human gut microbiome

As *E. coli* and other bacteria containing ESBL genes or other ARGs can reside in the gut microbiome, it is an important feature to consider when focussing on the spread of AMR. The human gut microbiome consists of a complex array of interactions between the host and the trillions of microorganisms residing in the intestinal tract, and is arguably so important to the host health that it could be classed as an organ [48]. The development of the gut microbiome begins at birth but is highly dynamic and unstable as a neonate, eventually stabilising around the age of three years old [49]. The first contact with microorganisms occurs during childbirth and this is highlighted in vaginally delivered infants having a gut microbiome that largely resembles the vaginal microbiome [50]. The gut microbiomes of infants delivered by caesarean instead resemble an epidermal microbiome from being handled by medical staff. The mode of delivery also plays a role in the health of the infant via the gut microbiome, as vaginally born infants contain more members of Bifidobacteria which are seen as vital pioneer species in the shaping of a normal healthy gut microbiome due to their beneficial immuno-modulatory effects [51]. A reduction in Bifidobacteria is associated with necrotising enterocolitis in preterm infants, and the development of chronic inflammatory disorders of the intestinal tract and atopic dermatitis later in life which emphasises the importance of early development of the gut microbiome for future health.

An infant's diet also plays a large role in shaping the gut microbiome for the better. Human milk oligosaccharides (HMOs) are one of the largest constituents of breast milk and protect the infant by modulating the presence and abundance of bacterial species in the gut microbiome. HMOs promote the growth of beneficial species like members of Bifidobacteria, as these possess a variety of glycoside hydrolases that enable HMO metabolism and enables them to dominate the gut microbiome in early life [52]. On the other hand, some HMOs display antibiotic properties and can suppress the

growth of detrimental species like group B *Streptococcus* which can cause feverish infections in infants [53]. Other pathogenic species induce intestinal inflammation during their pathogenesis, so suppressing this can help protect against infection. HMOs are able to act as decoy receptors for enteropathogenic species, sequestering them and preventing adhesion to intestinal epithelial cells and acute diarrhoeal infections [54]. The LPS layer of gram-negative bacteria is able to induce intestinal inflammation but is modulated by HMOs to reduce its pro-inflammatory efficacy [55]. Ultimately, regulating the composition of the neonatal gut microbiome is achievable by controlling which microbiota the infant first comes into contact with, and by feeding them a diet that appropriately supports beneficial bacteria and inhibits unwanted species. The infant gut microbiota composition is the most significant predictor for the development of the autoimmune disease atopic dermatitis or other disorders like asthma and allergies [56] which emphasises the importance of giving an infant and their microbiome the best start to life.

1.5 The healthy adult gut microbiome

Although the gut microbiome starts off dynamic in early life, it reaches a state of relative stability where it is usually dominated by members of Bacteroidetes and Firmicutes with lower but equally prevalent levels of Actinobacteria and Proteobacteria [57]. The stability and diversity of the gut microbiome is generally a marker of gut microbiome health. A high taxonomic richness and diversity drives a high functional diversity and a redundancy of core functions that persist regardless of changes to species colonising the gut [49, 58]. Functional redundancy is vital in preventing niches from suddenly appearing where unwanted species could occupy, bloom and ultimately cause disease. Perturbations significant enough to decrease the bacterial diversity are frequently linked to disease [59], but the gut microbiome over time will revert to some semblance of its original stable composition after the perturbation pressure subsides [60].

As the gut microbiomes between healthy individuals can differ greatly, it is this stability of the gut microbiome that is likely the most important feature of consistent gut health. Regardless, there are still some factors that drive a difference between healthy individuals, either due to developmental reasons or behavioural reasons. A genetic driving factor was discovered when comparing the gut microbiome of twins, as monozygotic twins have a more similar microbial composition and abundance than

dizygotic twins [61]. Sexual dimorphism is a driver of the gut microbiome composition. Women tend to have significantly lower levels of Bacteroidetes than men [62] likely due to hormonal divergence during puberty as these differences are mainly observed post-puberty [63]. However, differences between the gut microbiomes of individuals is driven more by behavioural differences like hygiene, diet or the use of antibiotics. Each component of a diet will influence the microbial composition by providing nutrients to bacteria able to utilise them as a source of energy [64]. The use of probiotics - beneficial bacteria intentionally added to the diet with the aim of regulating intestinal health and reducing intestinal inflammation - are therefore a way of taking advantage of the effect that diet has on the gut microbiome [65]. Although the gut microbiome will attempt to revert to its original composition after perturbation [60], previous use of antibiotics is likely to cause a permanent alteration to the bacterial composition [66] and explain how microbiota compositions may differ between individuals. Both behaviours and genetics often differ by country, so research on the gut microbiome and its role in health is more relevant than ever with the increase in globalisation and movement of people.

The taxonomic composition of the gut microbiome is an important aspect to consider, but there is more complexity when considering the functional potential of the gut microbiome and its impact on host health. There is a balance between the intake of dietary substrates, the effects this has on the microbial composition and what microbial substrates are produced. Some of the most significant metabolites to consider in the health of the host and in beneficially regulating the microbial composition are short chain fatty acids (SCFAs), bile acid and tryptophan. SCFAs are microbially-derived fatty acids produced from the fermentation of dietary fibre, are central to the metabolic profile of the gut microbiome and have a role in the healthy function of multiple organs [67]. Within the gastrointestinal tract, the intestinal mucosa - the cells lining the intestines - utilise SCFAs for energy which allows them to maintain their protective barrier function between the intestinal lumen and the circulating blood system [68]. SCFAs also play a role in modulating the host immune response as well as suppressing tumourigenesis [68]. On the other hand, the host-derived metabolites known as bile acids, which are vital for healthy digestion due to aiding with solubilising dietary fats and cholesterol, are utilised by the intestinal microbiota. The microbiota utilising bile acids are protective, as their absence leads to a decreased function of the epithelial barrier function [69] which promotes intestinal inflammation especially in people with inflammatory bowel disorders (IBD) [70]. The third major metabolite for a healthy host and gut microbiome is the amino acid tryptophan which is required for protein biosynthesis. Humans cannot synthesise tryptophan so rely on it from their dietary

intake but also from its production by the gut microbiota. Tryptophan is a ligand for aryl hydrocarbon receptors (AhR) and this interaction elicits effects in immune homeostasis, resistance to pathogens and improved epithelial function [71]. Low levels of AhR are seen in people with IBD [71] and this increased susceptibility to enteric pathogens in mice [72]. Modern research is discovering the importance of the gut microbiome and normal cognitive development [73] and links between the two in cognitive conditions like depression, autism, Parkinson's disease, Alzheimer's disease and schizophrenia [74]. There is potential to directly or indirectly affect the brain via a gut-brain axis, a bidirectional linking of the gut microbiome and the central nervous system via routes like the neuroendocrine system (hormones), the enteric nervous system or a physical blood-brain barrier that may permit movement of certain molecules from the blood and into the brain [75]. In the context of bacterial infections and the spread of AMR, however, one of the the most interesting aspects of a healthy gut microbiome is the production of a colonisation resistance to pathogenic bacteria.

1.6 Colonisation resistance

Colonisation resistance was first discovered in the 1950s after antibiotic treated mice became more susceptible to infection from *Salmonella* Enteritidis due to a disruption to the normal gut microbiota [76]. Since then, a number of mechanisms have been discovered that contribute to the gut microbiome's ability to resist invasion from pathogens. This can be a result of a symbiosis between the host and certain bacterial species. Firstly, commensal species reducing the availability of nutrients can limit the nutritionally expensive expansion of invading species. The mouse pathogen *Citrobacter rodentium* – used as a model organism to represent EHEC and EPEC infections – was found to be unable to colonise mice with a functioning, healthy gut microbiome due to a reduced availability of amino acids needed in its colonisation [77]. This was not the case in germ-free or antibiotic perturbed mice and was reverted in those fed with a high protein diet. In another study, a diet consisting of only monosaccharides allowed *Bacteroides thetaiotaomicron* to successfully compete with *C. rodentium*, reduce its burden of infection and increase survival rate of the mice [78]. However, the addition of polysaccharides to this diet gave *C. rodentium* a nutrient it can metabolise, where *B. thetaiotaomicron* cannot, which recovered its ability to infect. It is not just metabolites used for energy that are competed for, as *E. coli* Nissle has been observed reducing the intestinal colonisation of *Salmonella* Typhimurium by competing for iron uptake [79].

Secondary bile acids are produced by the microbiota from the metabolism of host-derived bile salts and is another mechanism involved in colonisation resistance. Bile salts are secreted into the small intestine to aid with solubilisation and absorption of fat but are primarily reabsorbed here if they are not metabolised by the microbiota. *Clostridioides difficile*, a gram-positive bacteria notorious for causing severe diarrhoeal infections, requires bile salts during its infection. Secondary bile acids, however, are toxic to *C. difficile* and can inhibit its growth [80]. Antibiotic treatment negatively impacts the gut microbiome and reduce levels of secondary bile acids being produced and permits *C. difficile* infection [80], emphasising how the microbial metabolism of host-derived metabolites can be vital in protecting against invasive pathogens.

Instead of utilising host-derived metabolites, some species of bacteria are able to produce their own molecules with antibiotic properties. Ribosomally synthesised, post-translationally modified peptides (RiPPs) are a group of naturally produced peptides with a diverse range of biological activities, especially acting as antimicrobials. RiPPs can be separated into classes like lantibiotics, bacteriocins, microcins, thiazole/oxazole-modified microcins and thiopeptides [81] which can be further split into subclasses, each with a different range of effects against other species [82]. A notable class of RiPPs are microcins which are exclusive to *Enterobacteriaceae* with only a narrow spectra of action usually against other *Enterobacteriaceae* [83]. The presence of RiPPs emphasises how colonisation resistance can be set up by the resident bacteria without interacting with the host, and microcins show how this may only be a family-specific interaction likely due to similar species competing for the same nutrients and space.

The host can also play a role in creating colonisation resistance to invasive pathogens. The intestinal mucosa produces mucus that lines the inner wall of the intestinal tract which provides a barrier and physical gap between the luminal bacteria and the host epithelium [84]. As previously mentioned, SCFAs produced by the gut microbiota are utilised by the host intestinal cells. A benefit of this is a functioning epithelial cell layer that provides defense against pathogens colonising the intestinal wall [85]. This interaction between epithelial cells and the gut microbiota is two-way; some species like *Akkermansia muciniphila* are able to use mucus as an energy source without causing infection and simultaneously provide an additional barrier against invading bacteria [86].

Functional colonisation resistance is a vital aspect of the healthy gut microbiome as its loss may permit infection. Fortunately, a loss of colonisation resistance can be reconstituted by interventions that add or replace resident bacteria and this approach has considerable potential in clinical infections. Faecal microbiota transplantation (FMT) has been established as a successful treatment strategy against recurrent *C. difficile* infection [87] so now the focus is on discovering and constructing other consortia of safe bacteria that can provide a colonisation resistance to each specified pathogen.

1.7 Gut microbiome in disease

As the stable balance of the microbial consortium, its metabolism and functioning colonisation resistance are strongly associated with health, the inverse is associated with disease. The disruption of the homeostasis is referred to as dysbiosis, and this can involve various microbial factors like a drastic shift in keystone taxa; a severe reduction in species richness or diversity; an increase in pathobionts or the production of harmful metabolites. The most well known cause of dysbiosis is the use of antibiotics as they can wipe out key members of the complex network within the microbiome and this differs by each type of antibiotic used [88].

Some diseases are associated with a dysbiotic state of the gut microbiome, but it is harder to determine whether this is due to the disease or as a result of it. One of the most frequently researched disorders is IBD which is used to describe two conditions: ulcerative colitis and Crohn's disease. The gut microbiome of people with IBD greatly differs from healthy individuals, as people with IBD usually have a reduced bacterial diversity [89] which is exacerbated further due to antibiotic use [90]. The major metabolic pathways of the gut microbiome may play a role in this; those with IBD have a microbiota less able to metabolise tryptophan [91] and therefore fewer protective features induced by processing this metabolite. Obesity and type 2 diabetes are also associated with more intestinal inflammation than healthy individuals [92] and may also be due to the lower microbial richness and diversity [93]. Inflammation is clearly a condition that is associated with the gut microbiome in a diseased state and is an important aspect to consider when aiming to eradicate a disease or minimise its impact.

The niche created by inflammation is filled by microbiota that thrive in the inflamed environment, and *Enterobacteriaceae* are the most adapted to this [94, 95]. Microbial infection often drives intestinal inflammation as some species use virulence factors to induce an environment that they can survive and colonise in [97, 96]. There are various other types of virulence factors used in bacterial pathogenesis: like mechanisms to avoid the host immune response [98]; cell-surface components known as adhesins that facilitate adhesion to host cells and permit invasion [99] or effector molecules and toxins injected into the host to alter cellular processes like inducing inflammation [100]. Strains that contain a high number of virulence factors are often associated with infection and disease in humans, like with ExPEC *E. coli* [101, 102]. An inflamed gut produces more nitrate which is utilised for anaerobic respiration by *Enterobacteriaceae* [103], which when combined with the increase in luminal oxygen from leaky blood vessels [104] allows for facultative anaerobes like *Enterobacteriaceae* to thrive over obligate anaerobes. *Enterobacteriaceae* were confirmed to thrive under these conditions with blooms detected after small bowel transplantation surgery after the intestinal lumen is exposed to oxygen [105]. The increase in blood circulation during inflammation also means that iron is more accessible to bacteria, so bactericidal enzymes are released by the host to prevent this. However, *E. coli* have special iron acquisition enzymes, siderophores, that can inhibit this host defence and bypass the host immune system [106]. Finally, microcin production is induced by intestinal inflammation which allows members of *Enterobacteriaceae* to more aggressively compete with each other under these conditions [107].

1.8 Gut resistome

Regardless of whether the gut microbiome is in a healthy, stable state or in a dysbiotic state, there will be species present that carry ARGs, and the entire collection of these is termed the gut resistome [108]. Traditionally, ARGs are considered in the context of pathogenic bacteria and AMR infections, but other resident commensal bacteria are a major source of ARGs that have the potential to mobilise through HGT. As already covered, $bla_{\text{CTX-M}}$ genes originated from an environmental genus *Kluyvera* [37], but other non-pathogenic bacteria like *Shewanella* have been discovered as the origin of quinolone resistance *qnr* genes [109] and OXA-48-like beta-lactamase [110]. It is therefore important not to neglect non-pathogenic bacteria within the gut microbiome as the potential source of acquired resistance in pathogenic bacteria. The resistome is

especially significant within *Enterobacteriaceae* because during their inflammation-induced blooming [94], there is increased contact between individual bacteria and an increased rate of HGT [111]. HGT typically occurs between members of the same phylum [112], and phage-mediated transmission of ARGs contributes to movement between species [113], including clinically significant ARGs like bla_{TEM} and $bla_{\text{CTX-M}}$ [114]. The efficiency with which bacteriophages can infect and reproduce is boosted by antibiotic usage [115], therefore supporting the spread of ARGs.

1.9 Analysis of a gut microbiome

As the complexity of a system increases, so does the difficulty of analysing it. Major advancements in DNA sequencing led to the creation of a complete draft of the human genome [116] and high-throughput approaches for genome-scale research [117]. It became possible to determine the taxonomic content of a sample by sequencing marker genes – a library of genes where each gene is unique to one taxon – without the need for culturing and avoiding the bias of only detecting species through culture-based approaches [118]. A common technology used is the sequencing of 16S rRNA genes, as these are highly conserved bacterial genes that can be compared between organisms to determine evolutionary distance [119]. However, this is not accurate at distinguishing to a species level [120], so next-generation sequencing is used when determining the entire genomic content of a sample to a higher level of accuracy.

Two main technologies are used for this, short-read sequencing and long-read sequencing. Illumina is the main platform for generating short-read sequences which works by segmenting the genomes into (up to) millions of fragments - the start of a process termed 'shotgun sequencing' - that are sequenced individually. Fragmented DNA is hybridised to a flow cell and fluorescently labelled free nucleotides are washed over the flow cell, after which the fluorescent signal is measured and the DNA sequence is determined. A drawback of this approach is that in the downstream analysis it is difficult to assemble short reads in the genomic areas that are abundant in repetitive DNA [121], so other technologies that produce long-read sequences (5000 - 30000 bp) were designed to overcome this. The main platforms for long-read sequencing belong to Oxford Nanopore and Pacific Biosciences (PacBio). Nanopore sequencing functions by measuring the ionic current changes of each nucleotide as DNA is passed through a pore [122], whereas PacBio functions by measuring the real-time addition of fluorescent

nucleotides to the terminal position of DNA during DNA synthesis [123]. Although these technologies improve the downstream assembly of sequencing data and are being employed in clinical research [124], the gain in length is accompanied with a reduction in the sequencing quality. To overcome this, hybrid assemblies that utilise both short- and long-read sequences are becoming the gold standard for accurate sequencing of entire microbiomes as it improves the overall quality and completeness of the genomes produced [125, 126].

Once the genomic content of a sample has been sequenced, the difficulty lies in being able to separate reads by the organism they originate from. Reads are first assembled into contiguous sequences (contigs) which are grouped into 'bins' that belong to putative genomes. This can be achieved by taxonomically assigning a contig to a reference genome and clustering these with other contigs of the same classification, or by clustering contigs by features of their structure [127]. The latter approach avoids biases associated with relying on the quality and completeness of reference databases, and this method led to the discovery of an immense amount of organisms termed as Candidate Phyla Radiation and a significant update to the 'tree of life' [128]. A less computationally expensive approach is to quantify marker genes specific to each taxon [129], but this relies on how well characterised the reference catalogue is. Regardless of the approach, computing the taxonomic content of a microbiome allows for the species diversity, richness and composition to be calculated, as well as the determination of the metabolic and functional potential of the microbiome. This is achieved by comparing predicted amino acid sequences to protein family databases, but this is usually skewed towards housekeeping genes and highly conserved pathways that are well annotated [130].

Microbiome research is more faceted than simply sequencing the DNA content, as at a fundamental level this is simply showing what genes are present and inferring their potential. It is possible to detect the expression of genes, the proteins produced from this, and the metabolites present in the environment with transcriptomics, proteomics and metabolomics. In transcriptomics analysis, RNA produced during gene expression is used as a template to create complementary DNA which can be sequenced [131]. Proteomics involves a detailed profiling of proteins present in a sample [132] and metabolomics uses high-resolution analytics like Nuclear Magnetic Resonance (NMR) or mass spectrometry (MS) to access the structure and weight of molecules, organic compounds and proteins. These technologies individually can give some insight into an aspect of a microbiome, but a multi-omics approach would be the most comprehensive

method of analysing the gut microbiome. Combining technologies could help researchers fully understand the associations between the gut microbiota and the host, especially in the context of disease or acquisition of AMR pathogens.

1.10 The gut microbiome, ESBL-producing *E. coli* and international travel

ESBL-producing *Enterobacteriaceae* (ESBL-E) are a highly prevalent pathogen responsible for extra-intestinal infections across the world. Although rates of gut colonisation of ESBL-E are low in Europe, carriage is far more common in other regions like Africa and Asia [133]. Epidemiological studies focussing on ESBL-E identified international travel as a major contributor to the spread of AMR [47, 134, 135, 136]. Intestinal colonisation of ESBL-E can endanger a traveller's health, especially in immunocompromised people visiting hospital [137, 138], so the prevention of onwards transmission is of high importance. Acquisition of ESBL-E is significantly prevalent in travellers that develop travellers' diarrhoea; are exposed to antibiotics [47, 135, 136]; have pre-existing chronic bowel disease or consume food from street vendors [47]. Each of these represents a likely perturbation to the gut microbiome that is permitting the colonisation of ESBL-E at an increased rate and highlights an interesting target of research in preventing this. A highly diverse microbiome may have more capability with producing metabolites that positively modulate the intestinal health, like SCFAs [68, 67], and supporting beneficial bacteria that minimise intestinal inflammation that the invading species, like ESBL-E, thrive in [94, 95].

The microbiota profile and microbiome richness have previously been found to not be associated with the acquisition of MDR *Enterobacteriaceae* during international travel [139, 140], whereas its clearance one month after returning was higher in those that had a higher bacterial diversity and richness before travel [139]. The discrepancy between the acquisition and clearance shows how the microbiome likely has a role, but further research is required. This is because recent studies into how the gut microbiome is involved are limited by sample size or technologies not suited for comprehensive microbiome research (i.e. 16S rRNA gene amplicon methods), and showcases a significant gap in the field with regards to understanding the role of the gut microbiome. Uncovering the mechanisms in which the gut microbiomes of some travellers successfully

compete with invading ESBL-E could therefore be exploited to prophylactically prevent its acquisition during international travel, and minimise the impact that increased globalisation is having on the spread of AMR bacteria.

Besides the microbiome structure in its entirety, individual strains and species can interfere with pathogen colonisation. This is predominantly achieved via within-family competition, for example via the use of microcins in *Enterobacteriaceae* [83, 141] or competition for the same niche [142, 143]. There is a large amount of *E. coli* strain diversity [3, 4] which varies in prevalence across the globe [144], and many of these strains possess different microcins [145]. However, despite the high diversity in potential within-*E. coli* strain competition, there are no large-scale studies into how the commensal, pre-travel population of *E. coli* is interacting with the MDR *E. coli* that are acquired internationally.

1.11 Aims and hypothesis

Considering the above points, there is clearly a limit to our understanding of the interplay between AMR bacteria and the gut microbiome during colonisation in travellers. The aim of this thesis is to answer if there are aspects or structures of the microbiome that are protecting against invasion, or if the within-*Enterobacteriaceae* dynamic is key to countering ESBL-E. We hypothesise that **a highly diverse gut microbiome is providing significant protection against invading ESBL-E; but also that there is a microbiota or individual taxa that contribute to establishing a colonisation resistance in travellers that remain uncolonised by ESBL-E**. This was tested by studying three aspects:

1. The features of the entire microbiome that are preventing ESBL-E from colonising
2. The taxa similar to ESBL-E that are competing for the same niches
3. How these species are interacting with each other *in vitro*

Determining the microbiota composition responsible for interacting with ESBL-E could be exploited to intervene with its colonisation. The ability to manipulate the gut microbiome to create colonisation resistance would be groundbreaking in controlling the spread of AMR infections across the world, and research of this thesis contributes to the development of this.

1.12 Research summary

This thesis follows on from the study in [47] and utilises analysis of whole metagenome sequencing data from faecal samples of travellers before travel and immediately upon return. The overarching aim of each analysis is to discover ways in which colonisation resistance to ESBL-E may be formed and explain why some individuals remain uncolonised. Firstly, the microbiome-level structures were compared between those who did and did not acquire ESBL-E during travel. The species richness, diversity and composition - typical measures of a gut microbiome - were calculated to discover if the pre-travel microbiome could predict the protection against ESBL-E. Next, the longitudinal changes in gut microbiome measurements and to individual taxa were calculated and separated by ESBL-E acquisition, to illuminate how some gut microbiomes are more resilient and to detect which taxa may be contributing to this. Discovery of a consortia of bacteria inversely associated with ESBL-E acquisition would be powerful, as prophylactic interventions during international travel could help control the spread of AMR. However, the results from this approach found no microbiome-wide dynamics leading to a colonisation resistance, so the resolution of focus was increased to within-*Enterobacteriaceae* dynamics. The longitudinal change in individual taxa suggests that *Citrobacter freundii* may be protective, so travel-acquired *Citrobacter* from travellers to Laos were competed *in vitro* against commensal *E. coli* and ESBL-producing *E. coli* to test this. The viable counts of both *Citrobacter* and *E. coli* that were incubated together in nutrient-rich media were determined to discover if *Citrobacter* was able to proliferate more than *E. coli*, especially ESBL-producing *E. coli*. The aim of this was to determine if *Citrobacter* possesses mechanisms that directly minimise the growth of *E. coli* and explain why it may be protective within a microbiome. However, it was *E. coli* that outgrew *Citrobacter*. This therefore requires further research as *Citrobacter* may instead have indirect effects on *E. coli* that are only mediated through the highly complex interactions within the gut microbiome.

Interestingly, the longitudinal change to *E. coli* during travel was not associated with the acquisition of ESBL-E, which were predominantly ESBL-producing *E. coli*. We expected those acquiring ESBL-producing *E. coli* to be more associated with *E. coli* than those that remain uncolonised, but the lack of this led to the hypothesis that *E. coli* strain replacement is occurring as its overall prevalence and abundance could remain constant. To test this, only the *E. coli* DNA was assembled per sample and compared to an extensive catalogue of reference *E. coli* genomes for identification. Understanding how resident populations of *E. coli* change during travel is epidemiologically important, as travel may be a significant driver of how commensal populations develop over time and contribute to the global trends of increasing numbers of AMR infections.

Chapter Two

Enterobacteriaceae and *Bacteroidaceae* provide resistance to travel-associated intestinal colonisation by multi-drug resistant *Escherichia coli*

The following chapter was published as the manuscript:

Davies, Matthew *et al.* (Dec. 31, 2022). "*Enterobacteriaceae* and *Bacteroidaceae* provide resistance to travel-associated intestinal colonization by multi-drug resistant *Escherichi coli*". **Gut Microbes**. https://doi.org/10.1080/19490976.2022.2060676

JP, PW, AM and WvS conceived the study design and acquired funding. MD performed all bioinformatics analysis and most statistical analyses. GG provided guidance with R analyses and performed statistical analysis for the metabolic profiling. MD drafted the article with critical revision from JP, WvS and AM. All the authors read and approved the final manuscript.

Minor changes were made to the chapter for this thesis, which included additional microbiome measurements that showed comparable outcomes as the published figures; additional analysis on AMR gene abundances that were not within my remit for the consortium; incorporating supplementary results into the main text and including additional context for SCFA synthesis in the introduction.

2.1 Background

International travel has increased in parallel with globalisation, with pre-SARS-CoV-2 pandemic statistics of over 1 billion tourist arrivals occurring each year [146]. Trafficking between countries with varying degrees of AMR has resulted in the substantial spread of multidrug-resistant (MDR) bacteria across the globe [136, 150, 149, 148, 147]. In particular, travel to destinations in South(eastern) Asia and Northern Africa has been linked to a high risk of acquisition of MDR bacteria. For example, up to 84% of travellers to India and 42% of travellers to Morocco have been shown to have acquired MDR *Enterobacteriaceae* (MDR-E) in their intestines upon travel return [47]. Considering that the intestinal tract is an open system, which is confronted with a myriad of bacteria from the environment (e.g. food, water, soil, other humans or animals), the human intestinal microbiota is considered the most important reservoir through which travellers acquire and import AMR [151].

Although most travellers will lose the acquired MDR strains within months after their return from travel, the acquisition of intestinal MDR bacteria can pose a direct health threat to the traveller. Especially when hospitalised, most significantly within intensive care units that frequently use large amounts of antibiotics [137], or when undergoing specific medical procedures [138], intestinal colonisation by MDR bacteria substantially increases the risk of MDR infections and consequently mortality. *Enterobacteriaceae* are the cause of some of the most common nosocomial MDR infections [152], and with the high rates of acquisition of Extended Spectrum Beta Lactamase-producing *Enterobacteriaceae* (ESBL-E) during international travel, the admission of travellers to hospitals is posing a serious health risk. Moreover, the introduction of MDR bacteria by international travellers also poses a potential risk for public health, as this permits the transfer of MDR bacteria to household members [153]. Identifying successful intervention strategies to prevent MDR-E acquisition by travellers would therefore not only directly protect vulnerable groups of travellers that have an increased risk of infections by (translocation of) intestinal (MDR) bacteria (e.g., immunocompromised individuals or patients undergoing surgery) but also reduce the spread of AMR across the globe.

Besides travel destination, the use of antibiotics during travel [47, 135, 136], and travellers' diarrhoea [47, 134, 136, 154, 155], have been repeatedly shown to increase the risk of MDR-E among travellers. In addition, travellers with pre-existing chronic bowel

disease have a higher risk of acquiring MDR-E [47]. Each of these factors are linked to major perturbations to the gut microbiome, suggesting that the microbiome might play a pivotal role in the susceptibility to MDR-E colonisation upon travel. Indeed, the role of the intestinal microbiome in providing colonisation resistance against incoming opportunistic pathogens has long been established [76]. Endogenous microbes can directly or indirectly interfere with pathogen colonisation through a variety of mechanisms, including the competition for nutrients and the production of small antimicrobial peptides (bacteriocins) [156, 157, 158] or immunomodulatory molecules like short chain fatty acids (SCFAs) [67]. SCFAs are produced by bacteria during fermentation of dietary fibres and are used to improve host epithelial barrier function, modulate host immune response and even suppress tumourigenesis [68]. The main SCFAs found in the gut microbiome are butyrate, propionate and acetate. Butyrate has multiple beneficial effects on intestinal health, mainly as an energy source for the gut mucosa and maintenance of correct mucosal functions, and has been increasingly linked to a reduction in colorectal cancer [159]. Propionate, on the other hand, is linked to gluconeogenesis in the liver and a reduction in cholesterol [160]. Acetate may be used in butyrate synthesis but also has links with controlling appetite, lipolysis and uptake of cholesterol [161]. Together, the modulation of the host intestinal health and immune system by commensal SCFA-producers may be involved with interfering with pathogen colonisation [162]. Studies on the role of the microbiome should therefore consider SCFA synthesis as it is an important proxy for intestinal health.

To uncover the effect of travel on the gut microbiome, a previous study employed metatranscriptomics on faecal samples from 43 travellers visiting tropical regions [139]. No significant associations between the intestinal microbiota composition before or after travel and the risk of MDR-E acquisition was found. However, the microbiota profiles did significantly differ between individuals who cleared MDR-E within one month and those who were persistently colonised. In another study among 10 international travellers, profiling of longitudinally collected faecal samples by metagenomic next-generation sequencing also did not reveal alterations in microbiome diversity and composition in those acquiring MDR-E [140].

These inconclusive results highlight the need for a larger-scale longitudinal study using whole metagenome sequencing (WMGS) to examine if specific microbial taxa and their functional capacities may protect against MDR-E acquisition during travel.

Here, we performed WMGS on faecal swabs of 190 travellers before travel and immediately upon return from South Asia, Southeast Asia, North Africa and East Africa. Metagenome sequences were processed via coabundance gene clustering to create taxonomic and metabolic profiles. In order to identify taxa and traits associated with a reduced rate of MDR-E acquisition, the (dynamics in) microbiome community structure, composition and functional profile were compared between travellers who acquired MDR-E during travel and those that did not.

2.2 Methods

2.2.1 Study population and design

The present study was embedded within the COMBAT study (ClinicalTrials.gov identifier: NCT01676974), a longitudinal multicenter study among 2,001 Dutch international travellers for which the design has been described previously [163]. Briefly, all travellers provided questionnaire data and faecal swabs before travel, immediately upon return and one month after return. Faecal swabs (Fecal Swab; Copan, Brescia, Italy) were incubated in tryptic soy broth supplemented with vancomycin (50 mg/L) and after overnight culture, subcultured onto chromID ESBL agar plates (bioMerieux, Marcy l'Etoile, France). All morphologically distinct colonies were identified to the species level using a matrix-assisted laser desorption/ionisation time-of-flight mass spectrometry (Bruker Microflex LT, Bruker, London, UK). For all *Enterobacteriaceae*, antibiotic minimum inhibitory concentrations were measured with the automated susceptibility testing system Vitek 2 (bioMerieux). Phenotypical confirmation of ESBL production was performed by combination disc diffusion tests, according to current national Dutch guidelines [164]. Remaining faeces were biobanked at $-80°C$ for future analyses.

For the present study, we randomly selected paired pre- and post-travel biobanked faecal samples from 190 travellers who visited South Asia, Southeast Asia, North Africa or East Africa and who had not used antibiotics in the three months prior to travel departure. Pre-travel samples of 11 travellers were positive for ESBL-E, and both pre- and post-travel samples of these travellers were removed from future analyses in order to study shifts in the microbiome in subjects "at risk" for ESBL-E acquisition.

Table 2.1: Description of travellers included in this study (N=179) and their travel associated factors

	Median (IQR)
Age (years)	51.6 (19.65 – 81.68)
BMI	23.74 (17.18 – 35.67)
	n/N (%)
Sex	
Male	75 (41.3%)
Female	104 (58.1%)
Antibiotic Use During Travel*	
No	163 (91.1%)
Yes	10 (5.59%)
Region visited during travel	
Northern Africa	42 (23.5%)
Eastern Africa	42 (23.5%)
Southern Asia	46 (25.7%)
Southeast Asia	49 (27.4%)
Travellers' diarrhoea during travel	
No	95 (53.1%)
Yes	84 (46.9%)
Acquired ESBL-E during travel	
No	76 (42.5%)
Yes	103 (57.5%)
*Numbers do not add up to 179 due to missing values.	

The metadata for the resulting 179 subjects who were negative for ESBL-E and had not taken antibiotics within three months prior to travel are displayed in Table 2.1.

Faecal metagenomic DNA was extracted using protocol Q of the International Microbiome Standards Consortium [165], which includes bead-beating (on a FastPrep Instrument (MP Biomedicals, Santa Ana (CA), USA) with 0.1 mm zirconium-silica beads (BioSpec Products, Bartlesville (OK), USA)) followed by purification using QIAamp DNA Stool Mini kit columns (Qiagen, Hilden, Germany). A Qubit fluorometer dsDNA HS Assay (Invitrogen) was used to quantify extracted DNA, and this DNA was stored at $-20°C$.

Diluted DNA (0.5 ng/uL) was prepared for sequencing with a Nextera Library Prep kit as described previously [166], purified using the Agencourt AMPure XP system (Beckman Coulter) and quantified using the Quant-iT PicoGreen dsDNA assay (Invitrogen). 10 nM of DNA from approximately 96 samples was equimolarly pooled (three independent times), for each sequencing lane. Pools were submitted for paired-end 2 × 150 base pair sequencing (Illumina NextSeq High-Output platform) with a targeted sequencing depth of 5 million reads per sample. Raw shotgun metagenomic reads have been deposited and released to NCBI SRA under BioProject ID PRJNA688274.

2.2.2 Construction of the nonredundant gene catalog

Samples were processed using the MOCAT software [167]. Reads were trimmed and filtered using fastx [168], with a minimum quality cutoff score of 15 and a minimum length cutoff of 60 bp. High-quality reads were assembled using Soap 2.04 [169] into scaftigs of a minimum length of 500 bp.

Scaftigs underwent de novo gene prediction via MetaGeneMark [170]. Predicted genes were clustered using BLAT by single linkage [171]. Genes were clustered if there was at least a 95% sequence similarity and at least 90% of the shorter gene was covered. This created a nonredundant gene catalogue.

High-quality reads were screened against the nonredundant gene catalog with SOAPaligner from Soap 2.04 and the parameters of a length cutoff of 30 bp, a percent cutoff of 95% identity, a seed length of 30, a maximum number of mismatches of 5, a random assignment of multiple matches and the -M 4 flag indicating for it to find the best hits. Coverage of each gene per sample was calculated using the soap.coverage script (source code available at [171]), and a gene abundance table was produced.

2.2.3 Generation of coabundance gene groups and metagenomic species

The abundance table was analyzed with the canopy clustering algorithm (source code available at [171]) that groups genes by similar abundance patterns across the samples. A canopy was created by grouping genes to a randomly selected gene within a distance of > 0.9 Pearson correlation coefficient and > 0.6 Spearman's rank correlation coefficient. Multiple canopies were clustered together if their median abundances had a distance of > 0.97 Pearson correlation coefficient. Canopies were considered to be of insufficient quality when containing only 2 genes; at least 90% of their abundance was made up from only 3 samples or were present in less than 4 samples. Canopies of a high quality were identified as coabundance gene groups (CAGs) and CAGs with at least 700 genes contained enough information to potentially be classified as metagenomic species (MGS).

2.2.4 Taxonomic classification and abundance calculation of MGS

The genes of each MGS were aligned to the blast nucleotide database (blastdb_nt_v5, accessed: January 2020) using BLASTN (v2.6.0; [172]) at a percentage identity of >45% on at least 100 bp and a maximum of 50 hits for each gene were kept. Bacterial taxa that received fewer than 25 hits were filtered out as noise. An MGS was assigned to a species if at least 70% of the total hits matched at a sequence similarity of >95% on at least 100 bp. Parameters were reduced to 60% of hits at >85% similarity to be assigned to a genus.

An MGS was quantified by mapping reads at >95% on at least 100 bp to the top 100 most correlated genes; counts were normalised by the total nucleotide length of these 100 genes and by sequencing depth of the samples. Individual CAG counts were only calculated for samples where at least three genes of that CAG were present.

2.2.5 Functional profiling

All genes in each MGS were converted into amino acid sequences with Transeq [173]. Amino acid sequences were orthology assigned to the eggNOG v5.0 database [174] (accessed: January 2020) by eggNOG-mapper [175]. KEGG annotations were selected from the results, and pathways were grouped into KEGG modules for downstream analysis. KEGG abundance per sample was calculated as the sum of the MGS abundances in which the KEGG module was detected.

Additionally, the initial nonredundant gene catalogue was annotated to KEGG Orthology (KO) and KOs belonging to butyrate, succinate and proprionate production were selected. Total microbiome abundances were calculated from the gene catalogues' equalised gene abundance table, and by summing every instance, a KO was annotated to a gene in each traveller.

2.2.6 Microbiome richness, diversity and community structure measurements

Statistical analyses were carried out in R on the entire study population as well as stratified according to travel destination. Alpha diversity measurements of species richness, Chao1 index and effective Shannon index were calculated using Vegan [176]. The dissimilarity in the microbial community structure (beta diversity) between samples was calculated by means of the Bray–Curtis dissimilarity index. Associations between the microbial community structure and metadata were determined using Envfit from the package Vegan.

To examine changes in microbial richness and diversity over time, the delta in alpha-diversity (intra-individual change in microbial richness/diversity between pre- and post-travel samples) was calculated. To examine the (in)stability of the microbial community structure, the within-subject Bray–Curtis distance between pre- and post-travel samples was calculated.

Univariate analysis consisted of the t-test or Mann Whitney U-test, depending on normality as determined by the Shapiro-Wilk normality test, which was applied to

examine associations between ESBL-E acquisition and microbial richness, diversity or SCFA abundance at baseline, as well as longitudinal changes in richness (observed species, Chao1), diversity (effective Shannon), SCFA abundance and community structure (Bray Curtis dissimilarity) measurements. Multivariate analysis consisted of fitting the data to a linear regression model, with the diversity measurement as the outcome and ESBL-E acquisition as the main covariate of interest, while adjusting for the following potential confounding factors: age (years), sex (male/female), BMI, travellers' diarrhoea (yes/no) and food consumption at food stalls (never/sometimes/daily).

2.2.7 Taxonomic and functional statistical analysis

To reduce the number of comparisons, taxa or KEGG modules present in <10% of samples or in which the relative abundance in the top 10% of samples was below 0.1% were removed from further taxonomic and functional analysis.

Zero-inflated two-step beta regression (ZIBR) analysis [177] was carried out on all remaining taxa; except for taxa absent in less than 10 samples as this is a prerequisite of the zero-inflated aspect of the model. The model was fit using the post-travel CAG as the dependent variable and its pretravel abundance, ESBL-E acquisition, age, sex, BMI and travellers' diarrhoea as covariates for both the logistic and beta components of the model. The time point variable was set to 1 due to the study containing only two time points, and the Gaussian quadrature points was set to 30. Analysis was carried out on all samples collectively and repeated for each travel destination separately.

A negative binomial generalised linear model [178] was used for the remaining taxa detected in the majority of samples (i.e. absent in fewer than 10 samples). Post-travel CAG abundance was predicted against ESBL-E acquisition while using pre-travel CAG abundance, age, sex, BMI and travellers' diarrhoea as additional covariates. Negative binomial generalised linear models were also used to study changes in KEGG module abundance in association with ESBL-E acquisition.

2.2.8 AMR gene abundance calculation from the entire metagenome

The abundance of AMR protein families was calculated using ShortBRED [179]. As setup, the ResFinder DNA sequence database [180] was downloaded (accessed: October 2022) and converted into amino acid sequence with Transeq [173]. The UniRef90 database was also downloaded (accessed: October 2022) from UniProt [181] and converted into a BLAST-ready database with the BLAST package makeblastdb (version 2.7.1+) [172]. The script *shortbred_ identify.py* was used to collapse ResFinder proteins of interest into protein families and screen them against the UniProt90 database, then create markers that distinguish each protein family. The script *shortbred_ quantify.py* was then used to screen each metagenomic sample and quantify the markers. The output abundance count is normalised to reads per kilobase million (RPKM) but counts that were created from <10 hits were removed from subsequent analysis.

A generalised liner model was used for statistical analysis comparing counts across time point or between travellers with and without ESBL-E. Abundance was the predicted variable with age, sex, BMI and travellers' diarrhoea as additional covariates.

2.3 Results

2.3.1 Baseline microbiome is not significantly associated with risk of acquiring ESBL-E

Faecal culturing revealed that 103 (57.5%) out of the 179 travellers who were negative for ESBL-E prior to travel acquired ESBL-E during their trip. From the post-travel faecal samples (T1) of these 103 travellers, a total of 148 morphologically distinct ESBL-E strains were isolated. The vast majority of strains were identified as *E. coli* (136/148, 91.9%), 9 strains (6.1%) were characterised as *Klebsiella pneumoniae* and the remaining 3 strains were characterised as *Proteus mirabilis*, *Citrobacter freundii* and *Klebsiella ornithinolytica*.

Faecal swabs were whole metagenome sequenced and processed into coabundance genomes (CAGs; Section 2.2.3). Upon quality filtering and preprocessing of WMGS data of the pre- and post-travel samples, 37,381 CAGs were generated of which 144 CAGs contained more than 700 genes and were classified as MGS.

The ability of the baseline microbiome to predict a traveller's risk of acquiring ESBL-E was determined by the richness, diversity, community structure and taxonomic composition of the MGS observed in pre-travel samples of travellers who did or did not subsequently acquire ESBL-E during travel. The observed number of metagenomic species was significantly higher (coefficient estimate: 2.85, 95% CI [0.11, 5.59]) in baseline faecal samples of travellers who acquired ESBL-E than those who did not, as determined using linear regression analysis (Figure 2.1; Appendix Table A). When stratifying the analyses according to travel destination, these differences were no longer observed (Appendix Table A).

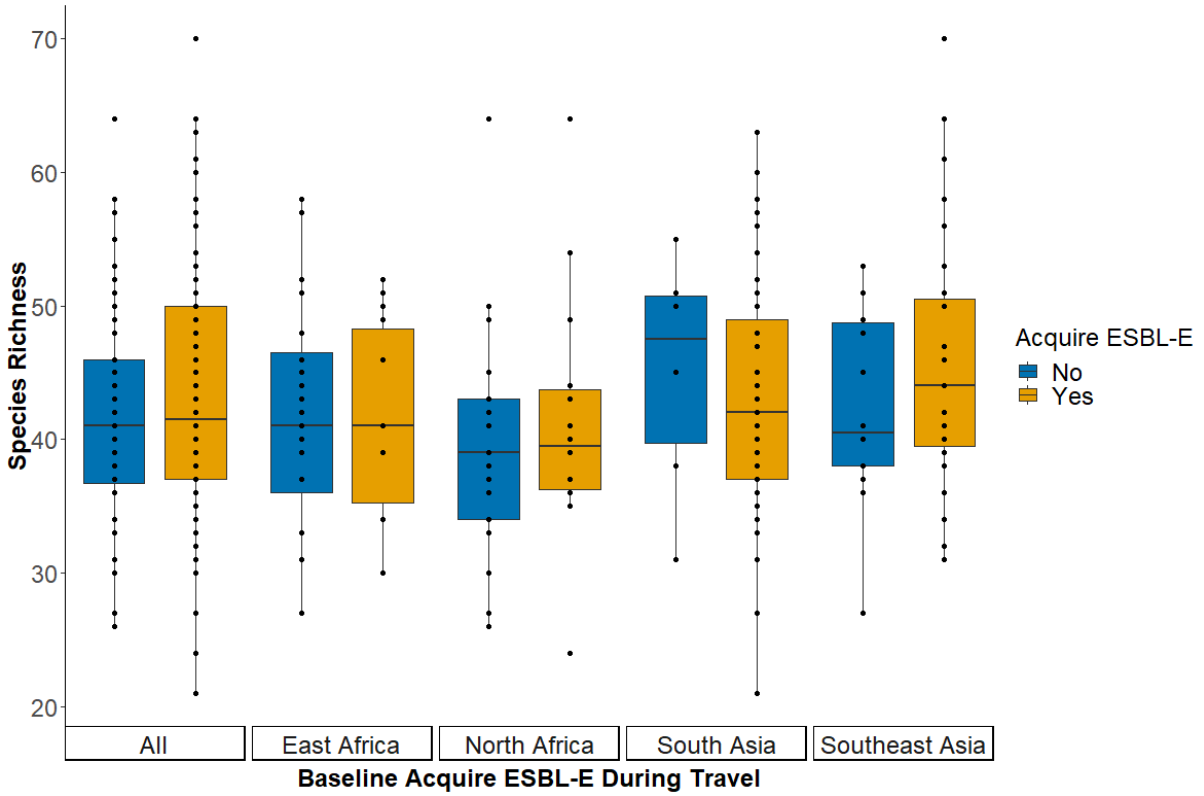

Fig. 2.1: Baseline observed species richness in association to ESBL-E acquisition.
Observed species richness in faecal samples collected at pre-travel baseline. Only metagenomic species (coabundance gene groups with >700 genes) were included in the diversity calculations. Differences in baseline alpha diversity metrics between individuals that did or did not acquire ESBL-E were tested using linear regression analyses. To adjust for potential confounding factors, alpha diversity metrics were treated as dependent variables and ESBL-E acquisition, sex, age, BMI and travellers' diarrhoea as independent variables, to produce coefficient estimates, standard errors and p-values (Appendix Table A).

The estimated number of species was also significantly higher (coefficient estimate: 2.82, 95% CI [0.09, 5.56]) in baseline faecal samples of travellers who acquired ESBL-E than those who did not, and this significance was also lost when stratifying to travel destination (Figure 2.2; Appendix Table A).

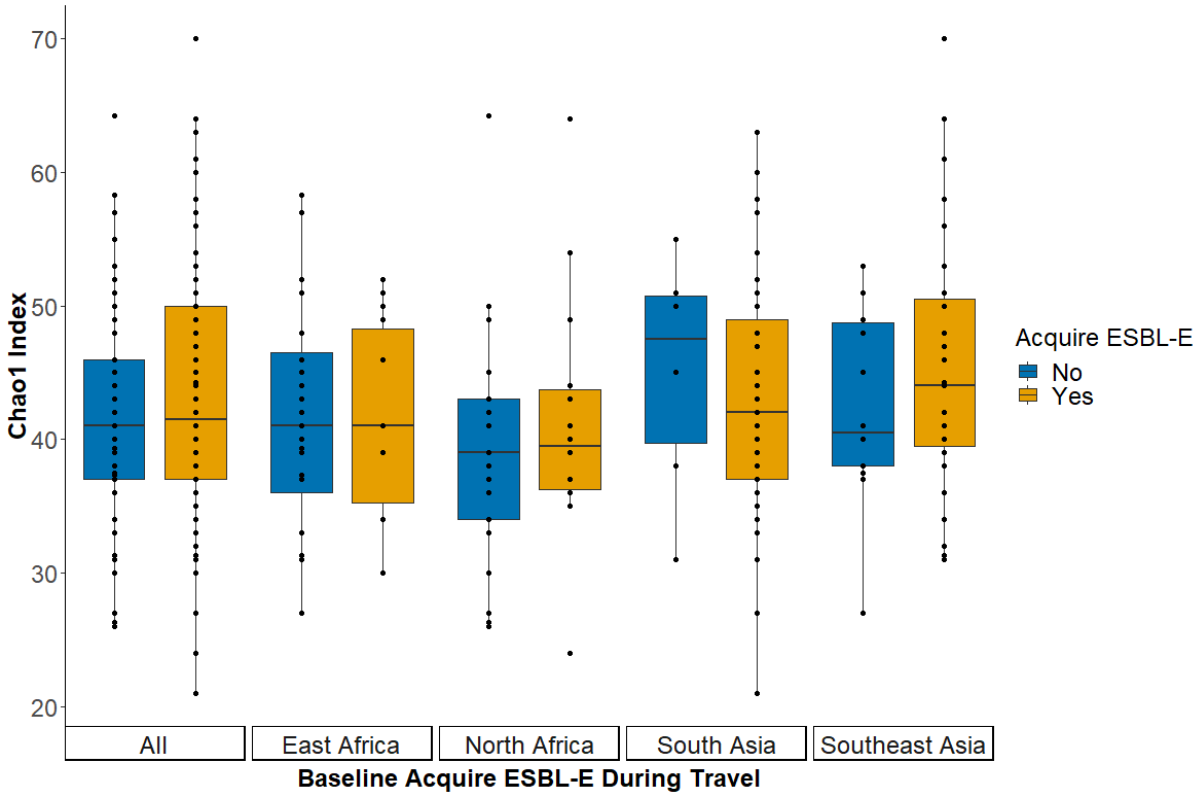

Fig. 2.2: Baseline estimated species richness and in association to ESBL-E acquisition.
Estimated species richness (Chao1) in faecal samples collected at pre-travel baseline. Only metagenomic species (coabundance gene groups with >700 genes) were included in the diversity calculations. Differences in baseline alpha diversity metrics between individuals that did or did not acquire ESBL-E were tested using linear regression analyses. To adjust for potential confounding factors, alpha diversity metrics were treated as dependent variables and ESBL-E acquisition, sex, age, BMI and travellers' diarrhoea as independent variables, to produce coefficient estimates, standard errors and p-values (Appendix Table A).

In contrast, the effective number of species as calculated from Shannon diversity was neither statistically significantly associated with ESBL-E acquisition in the overall study population ($p = 0.235$) nor when stratified to travel destination (Figure 2.3; Appendix Table A).

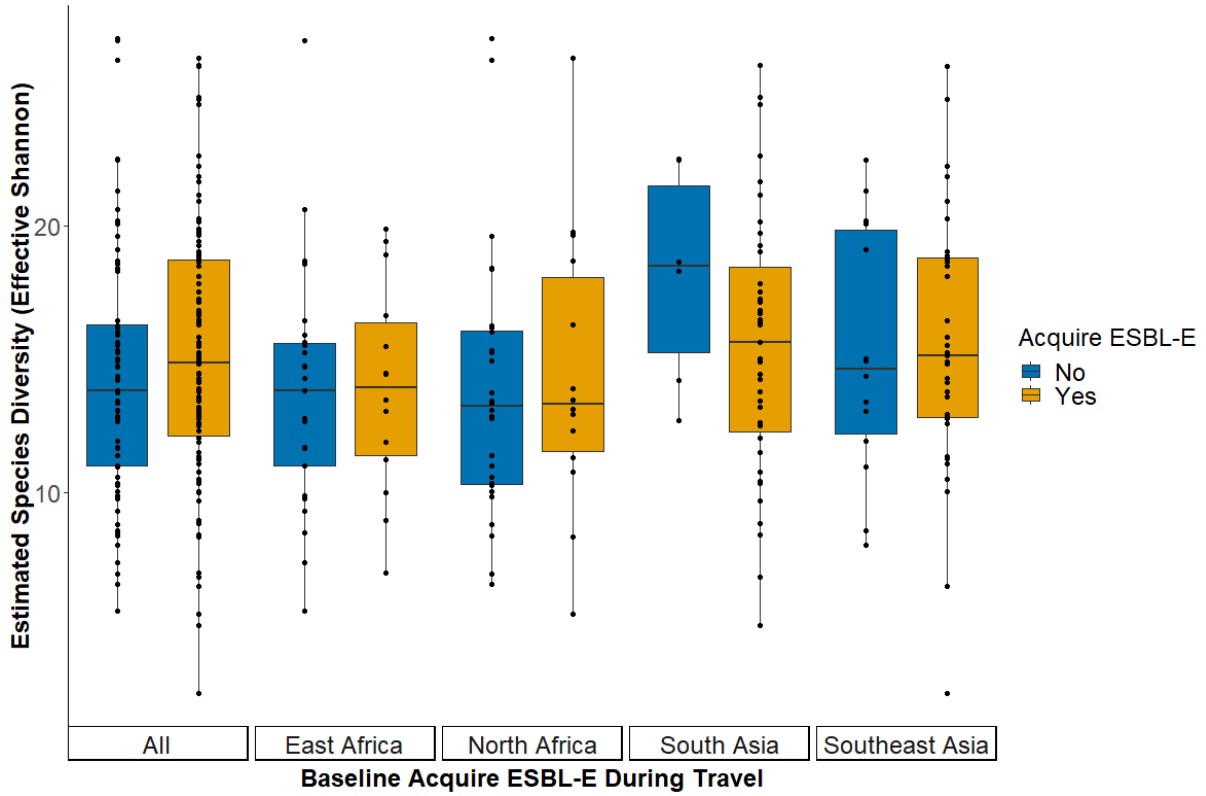

Fig. 2.3: Baseline microbial diversity in association to ESBL-E acquisition. Effective number of species as calculated from Shannon Index in faecal samples collected at pre-travel baseline. Only metagenomic species (coabundance gene groups with >700 genes) were included in the diversity calculations. Differences in baseline alpha diversity metrics between individuals that did or did not acquire ESBL-E were tested using linear regression analyses. To adjust for potential confounding factors, alpha diversity metrics were treated as dependent variables and ESBL-E acquisition, sex, age, BMI and travellers' diarrhoea as independent variables, to produce coefficient estimates, standard errors and p-values (Appendix Table A).

We subsequently examined whether the overall microbial community structure prior to travel, as examined by means of the Bray–Curtis dissimilarity, was associated with ESBL-E acquisition risk. Ordination of pre-travel samples in Principal Coordinates Analyses (PCoA) did not, however, show any clustering according to ESBL-E acquisition. Fitting the ESBL-E acquisition status to the ordination, using envfit, confirmed the lack of association between the baseline microbial community structure and the risk of subsequently acquiring ESBL-E during travel (Figure 2.4; $p = 0.3$; Appendix Table A). The significance of age, BMI, sex and ESBL-E acquisition in the microbial community structure was determined using the envfit function in vegan. The significance value was determined based on 999 permutations. All P values derived from envfit were adjusted for multiple comparisons using FDR adjustment (Benjamini–Hochberg procedure) and were not significantly associated with ESBL-E

acquisition (Appendix Table A).

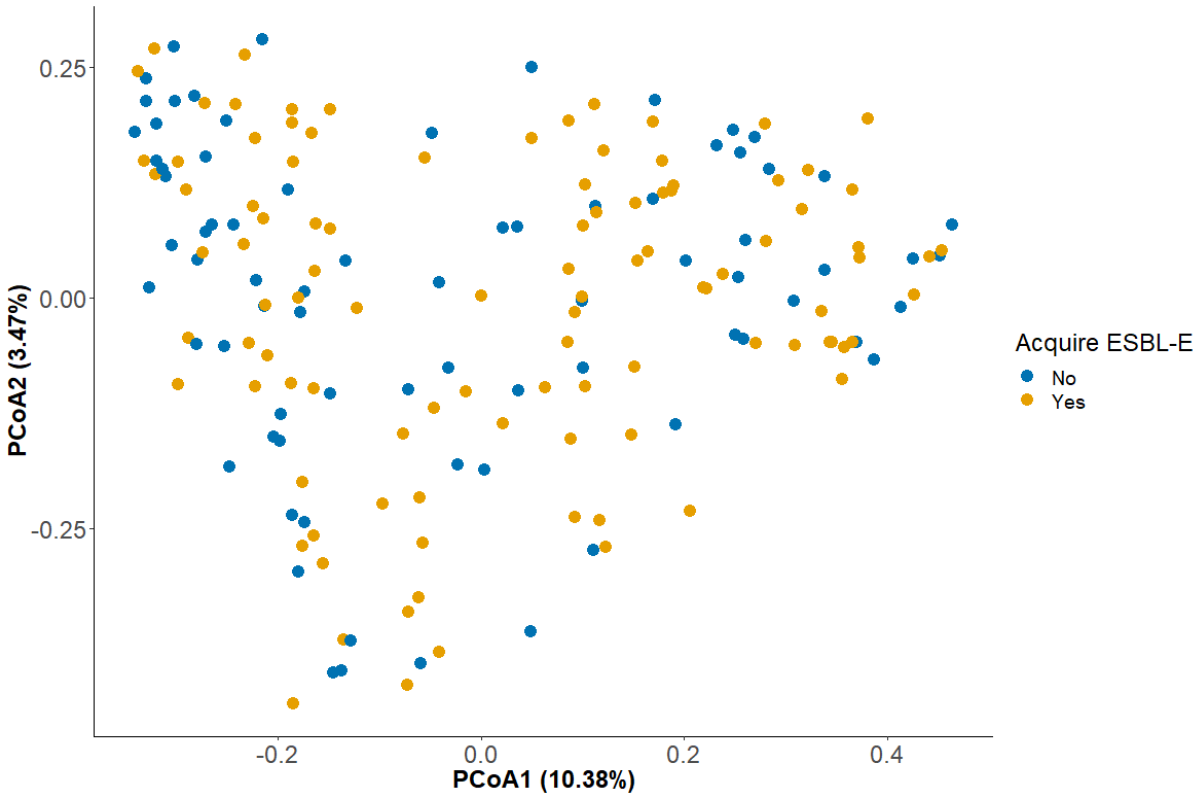

Fig. 2.4: Baseline microbial community structure (Bray-Curtis) in association to subsequent ESBL-E acquisition during travel.
Ordination was performed using PCoA based on Bray-Curtis dissimilarity. Only metagenomic species (co-abundance gene groups with >700 genes) were included in the diversity calculations. Samples are coloured according to subsequent ESBL-E acquisition during travel.

2.3.2 Microbiome structure and diversity are altered during travel, but mainly associated with the development of travellers' diarrhoea and not the acquisition of ESBL-E

To examine the association between microbiome stability and the susceptibility to ESBL-E acquisition, longitudinal analyses were performed. The observed microbial richness was not associated with ESBL-E acquisition (Figure 2.5). However, a statistically significant (p = <0.001) decrease in the observed microbial richness (coefficient estimate: −4.47, 95% CI [−6.82, −2.14]) was associated with the onset of traveller's diarrhoea (Appendix Table A). When stratified for travel destination, the lack of association to ESBL-E acquisition was still observed, and the negative association

with diarrhoea was seen for travellers visiting East Africa (coefficient estimate: −7.1, 95% CI [−12.62, −1.58], p = 0.0133) and South Asia (coefficient estimate: −6.52, 95% CI [−12.05, −1], p = 0.0219) (Appendix Table A).

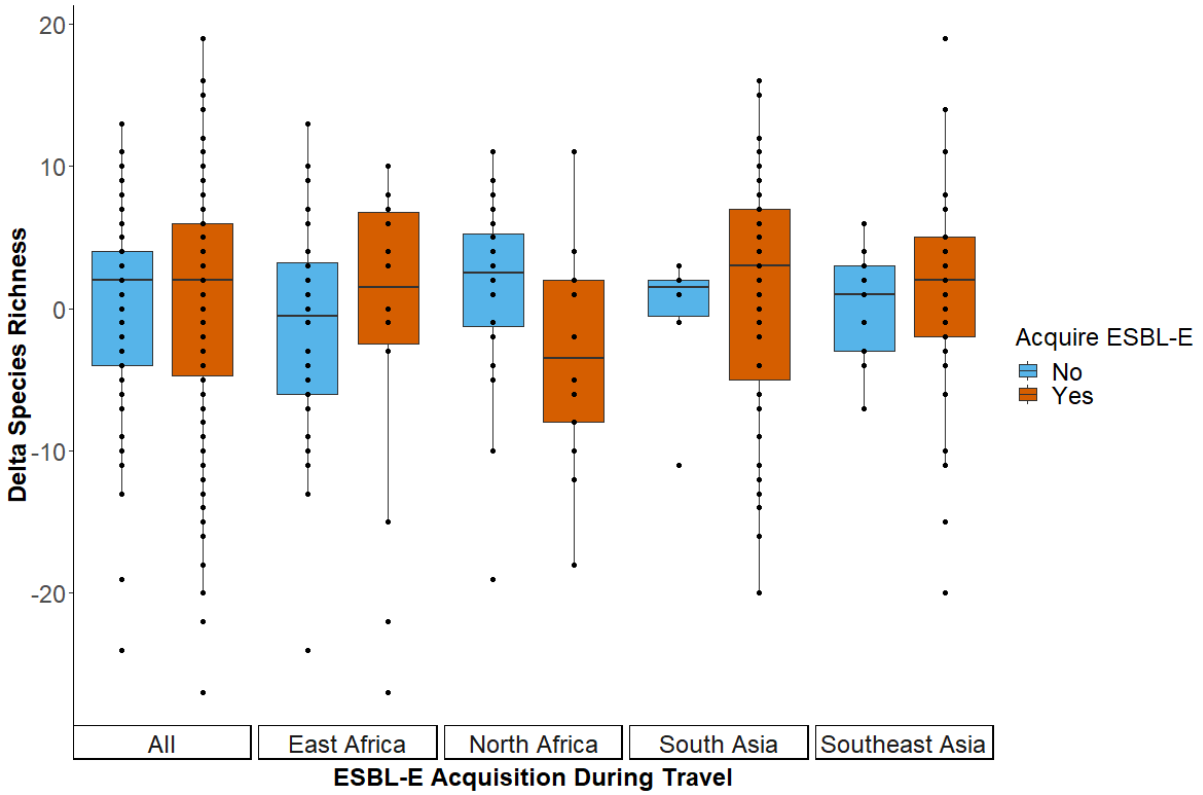

Fig. 2.5: Change in observed species richness in association to ESBL-E acquisition.
The change in observed species richness between pre- and post-travel faecal samples was calculated for each individual. Only metagenomic species (coabundance gene groups with >700 genes) were included in the diversity calculations. Differences in stability of species richness between individuals who did or did not acquire ESBL-E were tested using linear regression analysis. To adjust for potential confounding factors, analyses were performed with delta in alpha diversity metrics as dependent variables and ESBL-E acquisition, sex, age, BMI, travellers' diarrhoea and consumption of food from street food stalls as independent variables (Appendix Table A).

The estimated microbial richness was also not associated with ESBL-E acquisition (Figure 2.6) but had a significant (p = 0.001) decrease (coefficient estimate: −4.5, 95% CI [−6.9, −2.1]) associated with the onset of traveller's diarrhoea (Appendix Table A). When stratified for travel destination, the lack of association to ESBL-E acquisition was still observed, and the negative association with diarrhoea was seen for travellers visiting East Africa (coefficient estimate: −7.60, 95% CI [−13.36, −1.83], p =

0.0113) and South Asia (coefficient estimate: −6.31, 95% CI [−11.9, −0.73], p = 0.0277) (Appendix Table A).

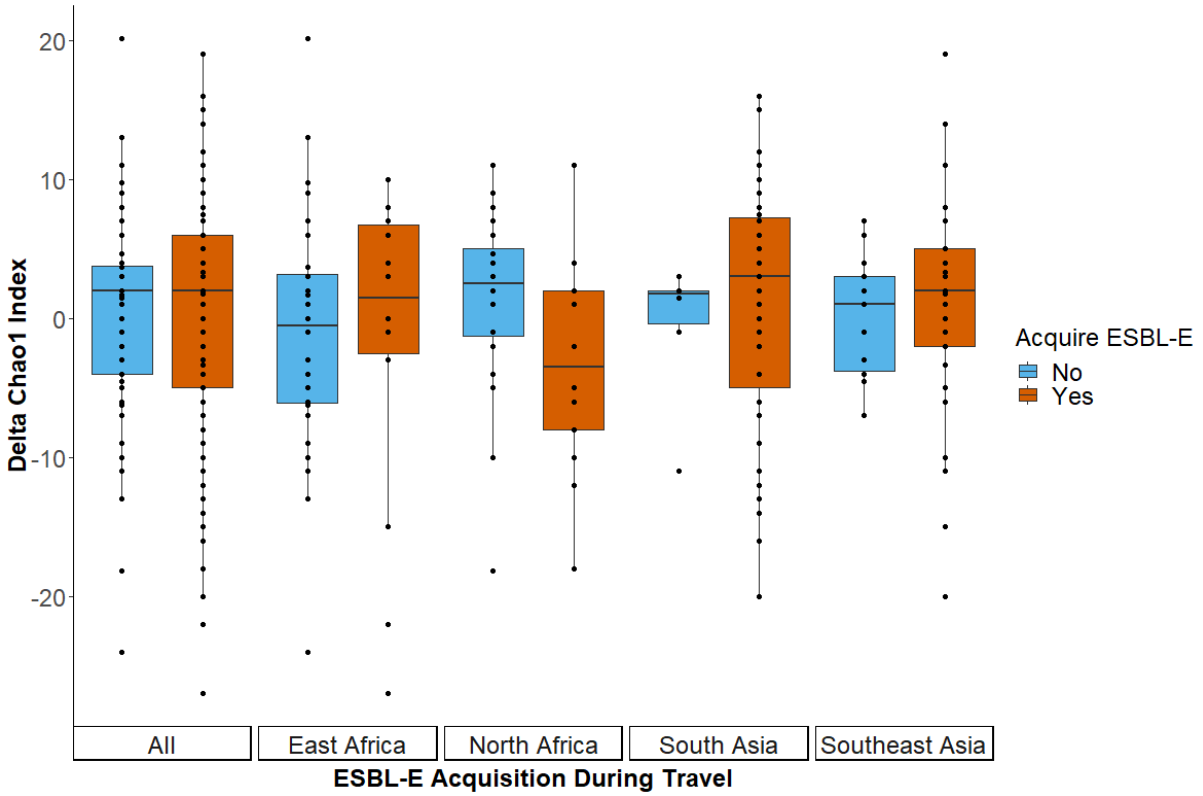

Fig. 2.6: Change in estimated species richness in association to ESBL-E acquisition.
The change in estimated species richness between pre- and post-travel faecal samples was calculated for each individual. Only metagenomic species (coabundance gene groups with >700 genes) were included in the diversity calculations. Differences in stability of species richness between individuals who did or did not acquire ESBL-E were tested using linear regression analysis. To adjust for potential confounding factors, analyses were performed with delta in alpha diversity metrics as dependent variables and ESBL-E acquisition, sex, age, BMI, travellers' diarrhoea and consumption of food from street food stalls as independent variables (Appendix Table A).

The effective number of species as calculated from Shannon Index was not associated with ESBL-E acquisition (Figure 2.7) but had a significant (p = <0.001) decrease (coefficient estimate: −5.5, 95% CI [−5.05, −1.9]) associated with the onset of traveller's diarrhoea (Appendix Table A). When stratified for travel destination, the lack of association to ESBL-E acquisition was still observed, and the negative association with diarrhoea was observed for travellers visiting East Africa (coefficient estimate: −5.46, 95% CI [−8.11, −2.81], p = <0.001) and South Asia (coefficient estimate: −4.05,

95% CI [−7.56, −0.54], p = 0.0249) (Appendix Table A).

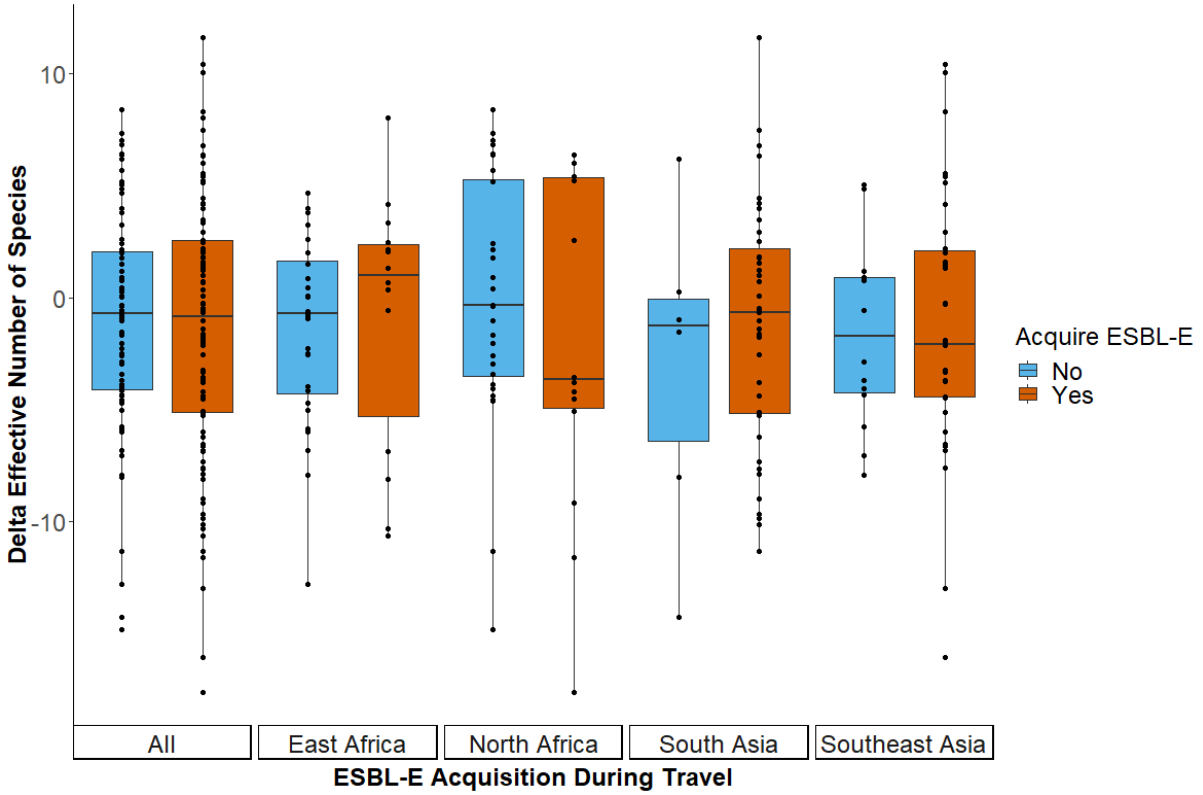

Fig. 2.7: Change in effective number of species in association to ESBL-E acquisition.
The change in effective number of species between pre- and post-travel faecal samples was calculated for each individual. Only metagenomic species (coabundance gene groups with >700 genes) were included in the diversity calculations. Differences in stability of species richness between individuals who did or did not acquire ESBL-E were tested using linear regression analysis. To adjust for potential confounding factors, analyses were performed with delta in alpha diversity metrics as dependent variables and ESBL-E acquisition, sex, age, BMI, travellers' diarrhoea and consumption of food from street food stalls as independent variables (Appendix Table A).

As for the stability in microbial richness and diversity, the stability in the microbial community structure was also examined among travellers. The within-subject Bray–Curtis distance between pre- and post-travel samples was not significantly different between travellers who did or did not acquire ESBL-E (Figure 2.8; p = 0.504). Upon stratification to travel destination, the lack of significance remained (Appendix Table A).

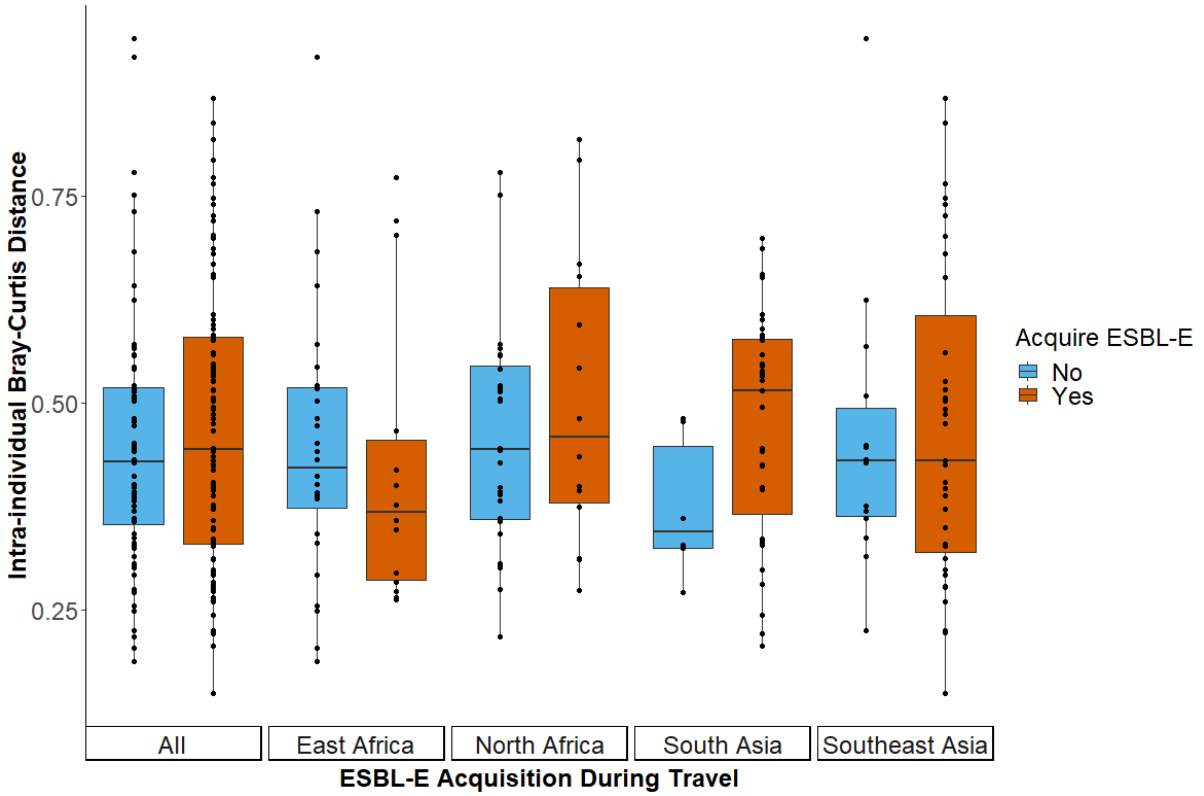

Fig. 2.8: Stability in the microbial community structure in association with ESBL-E acquisition.
Intraindividual Bray–Curtis dissimilarity between pre- and post-travel samples. Differences in stability in the microbial community structure between individuals who did or did not acquire ESBLs were tested using linear regression analyses. Analyses were performed with intraindividual Bray–Curtis dissimilarity as the dependent variable and ESBL-E acquisition, sex, age, BMI, travellers' diarrhoea and consumption of food from street food stalls as independent variables (Appendix Table A).

However, the microbial community structure is less stable among travellers who developed travellers' diarrhoea as the within-subject Bray–Curtis distance was significantly larger among travellers that experienced diarrhoea compared to those that did not (coefficient estimate: 0.074, 95% CI [0.03, 0.12]; p = 0.002; Appendix Table A). There was also a significantly larger within-subject Bray–Curtis distance between travellers who ate at local street food stalls daily (Coefficient estimate: 0.176, 95% CI [0.07, 0.28]; p = 0.001; Appendix Table A). Upon stratification according to travel destination, this signal remained for travellers eating daily from street food stalls in Southeast Asia (coefficient estimate: 0.27, 95% CI [0.12, 0.42], p = <0.001; Appendix Table A)

2.3.3 Changes in the taxonomic and functional composition of the microbiome

The shifts in prevalence and relative abundance of individual microbial taxa in association with ESBL-E acquisition were next examined with ZIBR [177]. Eight MGS were significantly associated with ESBL-E acquisition during travel (Figure 2.9). Three belonged to the family *Enterobacteriaceae* (MGS-003 and MGS-102 were classified as *Citrobacter freundii* and MGS-016 as *Klebsiella pneumoniae*), and three belonged to the genus *Bacteroides* (these MGSs did not meet the threshold for species level classification, i.e., <70% of hits to the same species). The genes within the remaining two MGS assign to various phyla, and interpretation for the latter two MGS should therefore be done with caution as these might be potential artifacts.

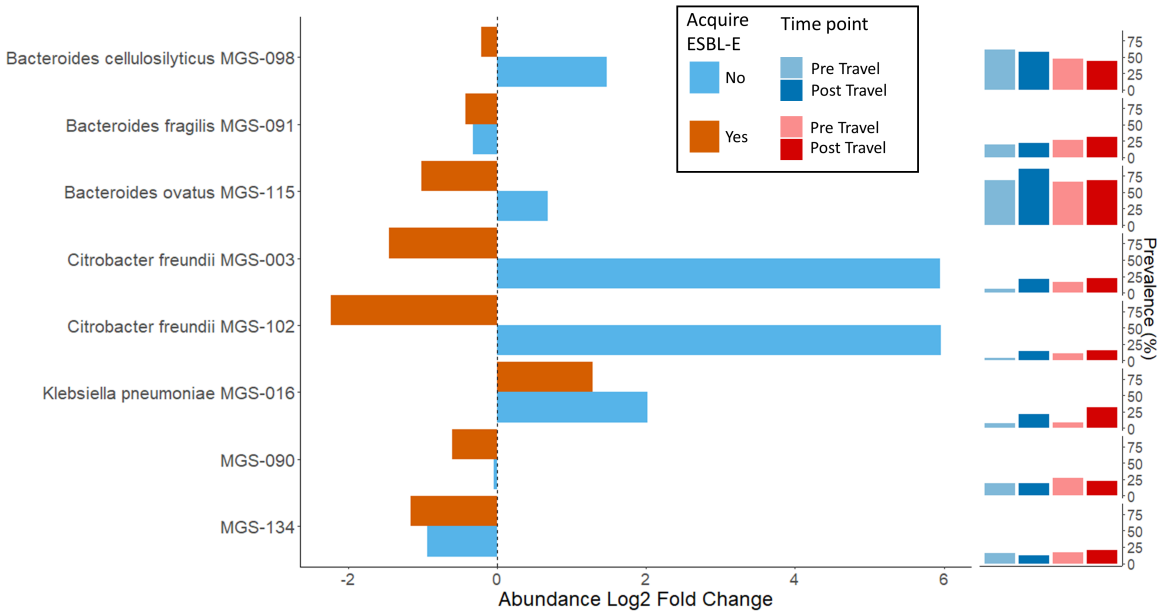

Fig. 2.9: Taxonomic composition shift in association with ESBL-E acquisition. Plot depicting the log2-fold change of abundance, total presence count (prevalence), median relative abundance in total population and p-value of each taxa that significantly changes during travel. Only taxa significantly associated with ESBL-E acquisition were selected for visualisation. Differences in pre-travel to post-travel composition were tested using zero inflated beta regression with a random effects (ZIBR) statistical model. Analysis was performed with the post-travel abundances as dependent variables and pre-travel abundances, ESBL-E acquisition, sex, age, BMI and travellers' diarrhoea as the covariates for both the logistic and beta components.

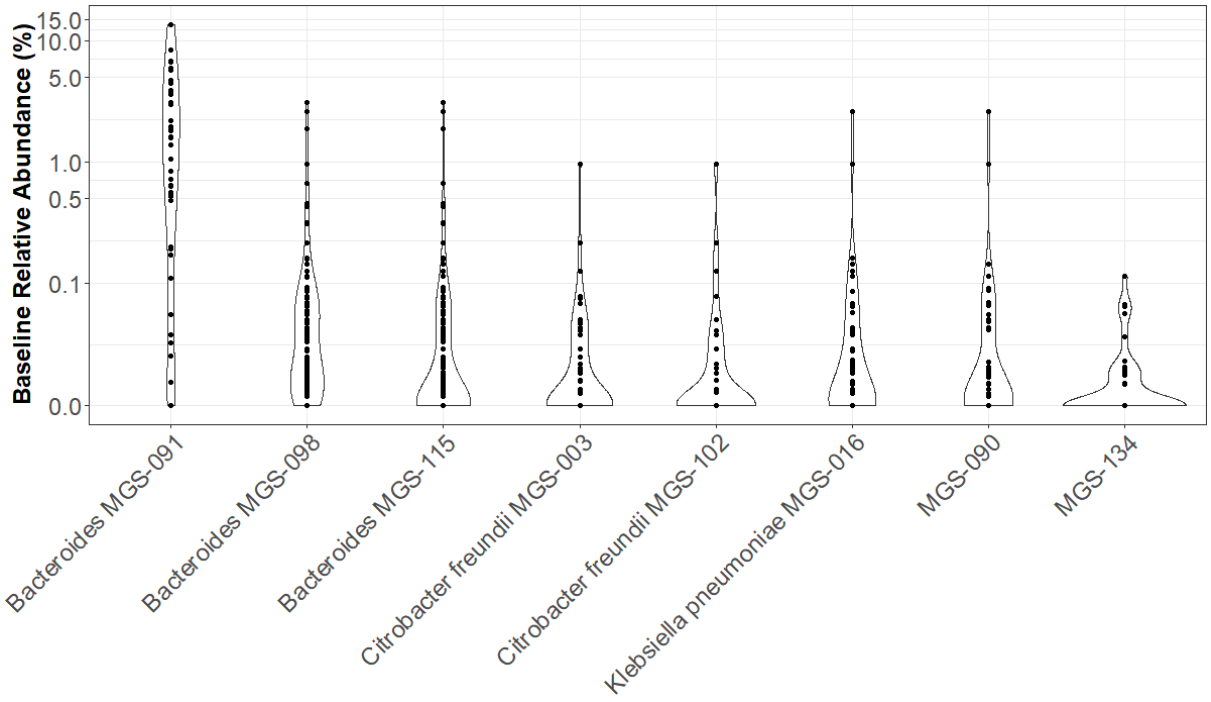

Fig. 2.10: Pre-travel relative abundances of MGS significantly associated with ESBL-E acquisition.
Violin plots depicting the pre-travel relative abundances of each taxa that is significantly associated with the acquisition of ESBL-E during travel.

MGS-091, the most abundant *Bacteroides* MGS (see Figure 2.10 for baseline relative abundances), had both a 0.42-fold decrease and a 0.33-fold decrease in travellers who did and did not acquire ESBL-E, respectively. *Bacteroides* MGS-098 had a 1.47-fold increase in abundance in travellers who did not acquire ESBL-E, but a 0.21-fold decrease in travellers who acquired ESBL-E. Additionally, MGS-115, the most prevalent *Bacteroides*, had a 0.69-fold increase in travellers remaining ESBL-E-negative and a 1.01-fold decrease in travellers who acquire ESBL-E.

The most pronounced shifts were, however, observed for two MGS identified as *Citrobacter freundii*. Both *C. freundii* MGS showed similar longitudinal patterns; increasing 5.96 and 5.95 fold, respectively, in travellers who did not acquire ESBL-E while decreasing 2.23 and 1.45 fold in travellers who did. The abundance of MGS-016, identified as *Klebsiella pneumoniae*, has both a 1.28- and 2.02-fold increase in those subjects who did and did not acquire ESBL-E, respectively. However, the increase in its prevalence from 8.9% to 31.7% of subjects is larger in travellers who acquire ESBL-E during travel, compared to 7% to 20.8% of subjects in those who remain negative for ESBL-E.

MGS-090 had a 0.04-fold decrease in travellers who remained negative for ESBL-E, but a 0.6-fold decrease in those who acquired ESBL-E. MGS-134 had a 0.94- and 1.16-fold decrease in travellers who remained negative for ESBL-E and acquired ESBL-E, respectively.

The metabolic profile of the samples was determined by functionally annotating each gene of each MGS and clustering by the KEGG module. Samples were hierarchically clustered by distance to each other, based on the fold changes of each module, but did not cluster by ESBL-E acquisition (Figure 2.11).

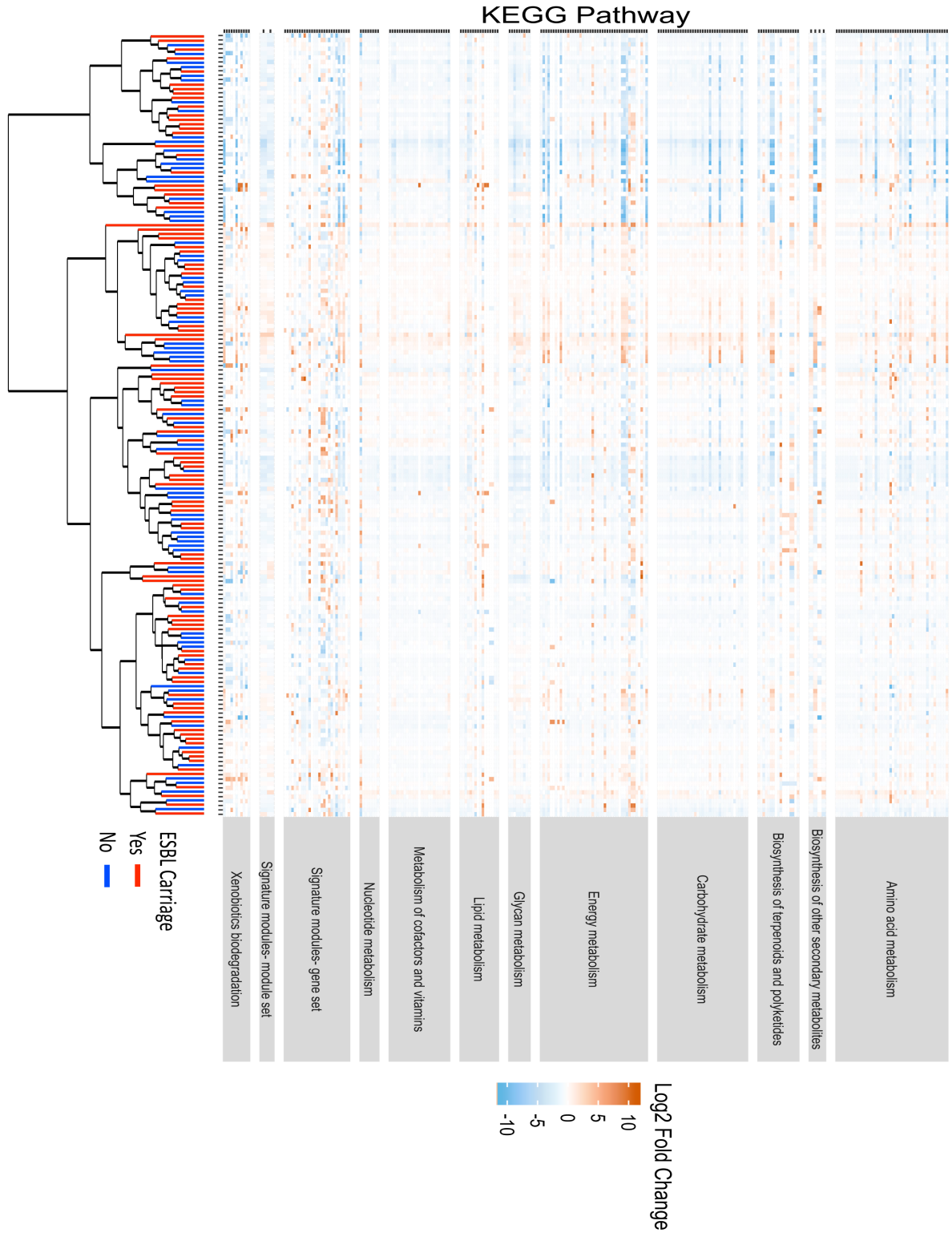

Fig. 2.11: Metabolic profile shift during travel.
Heatmap depicting the log2-fold changes of each KEGG module per sample during travel. An increase in KEGG module is coloured orange and decrease in blue. KEGG modules are clustered into 12 KEGG pathways whose height is weighted by the number of constituent modules detected. Samples are clustered by dissimilarity distance and are visualised in the dendrogram, with branch ends coloured as blue for ESBL-negative and red for ESBL-positive during travel.

As the production of SCFAs is a proxy for good intestinal health, we were interested to know if more SCFA synthesis protected the gut against ESBL-E acquisition. We examined the abundances of KEGG Orthology genes involved in butyrate, succinate and propionate production. Neither the baseline abundances of these genes nor the abundances in post-travel samples or shifts in abundance between pre- and post-travel were associated with ESBL-E acquisition (Table 2.2).

Table 2.2: SCFA abundances statistical association to ESBL-E acquisition, separated into pre-travel, post-travel and change in abundances

		Pre-travel Abundance	**Post-travel Abundance**	**Delta Abundance Change**
KEGG Orthology	**Gene**	**FDR p-values**	**FDR p-values**	**FDR p-values**
	Butyrate production			
K00248	*bcd*	0.789134	0.772783	0.9276193
K00634	*ptb*	0.789134	0.772783	0.9276193
K00929	*buk*	0.789134	0.8133193	0.9276193
K01034	*but*	0.966684	0.7933541	0.7454598
	Succinate production			
K00024	*mdh*	0.966684	0.9197282	0.9276193
K01676	*fumA, fumB*	0.966684	0.772783	0.9276193
K01677	*fumA*	0.789134	0.772783	0.9276193
K01678	*fumB*	0.789134	0.772783	0.9276193
K01679	*fumC*	0.966684	0.772783	0.9276193
K00239	*sdhA, frdA*	0.8164514	0.772783	0.9276193
K00240	*sdhB, frdB*	0.966684	0.7933541	0.9276193
K00241	*sdhC, frdC*	0.966684	0.9197282	0.9276193
	Propionate production			
K01902	*sucD*	0.966684	0.772783	0.9276193
K01903	*sucC*	0.966684	0.772783	0.9276193
K01847	MUT	0.966684	0.939518	0.9276193
K01848	*mcmA1*	0.966684	0.772783	0.9965073
K01849	*mcmA2*	0.966684	0.9197282	0.9276193
K05606	MCEE, epi	0.9623947	0.772783	0.9276193
K11264	*scpB, mmcD*	0.966684	0.772783	0.9965073
K01026	*pct*	0.966684	0.772783	0.9276193
K00932	*tdcD*	0.966684	0.772783	0.9276193

2.3.4 International travel increases the number of AMR genes, especially in those who acquire ESBL-E

The AMR genes from the ResFinder database were collapsed into protein families with ShortBRED, and their abundances were calculated in each metagenomic sample. Low prevalence (detected in <10% of travellers) AMR families were removed, and the remaining were visualised in Figure 2.12, with Timepoint, ESBL-E carriage, Diarrhoea and Subregion added below. There is a cluster of samples with high abundance of AMR towards the left of the heatmap, and this represents a large portion of the post-travel samples, especially those who acquired ESBL-E and developed travellers' diarrhoea. This pattern, however, is not specific to subregion travelled to.

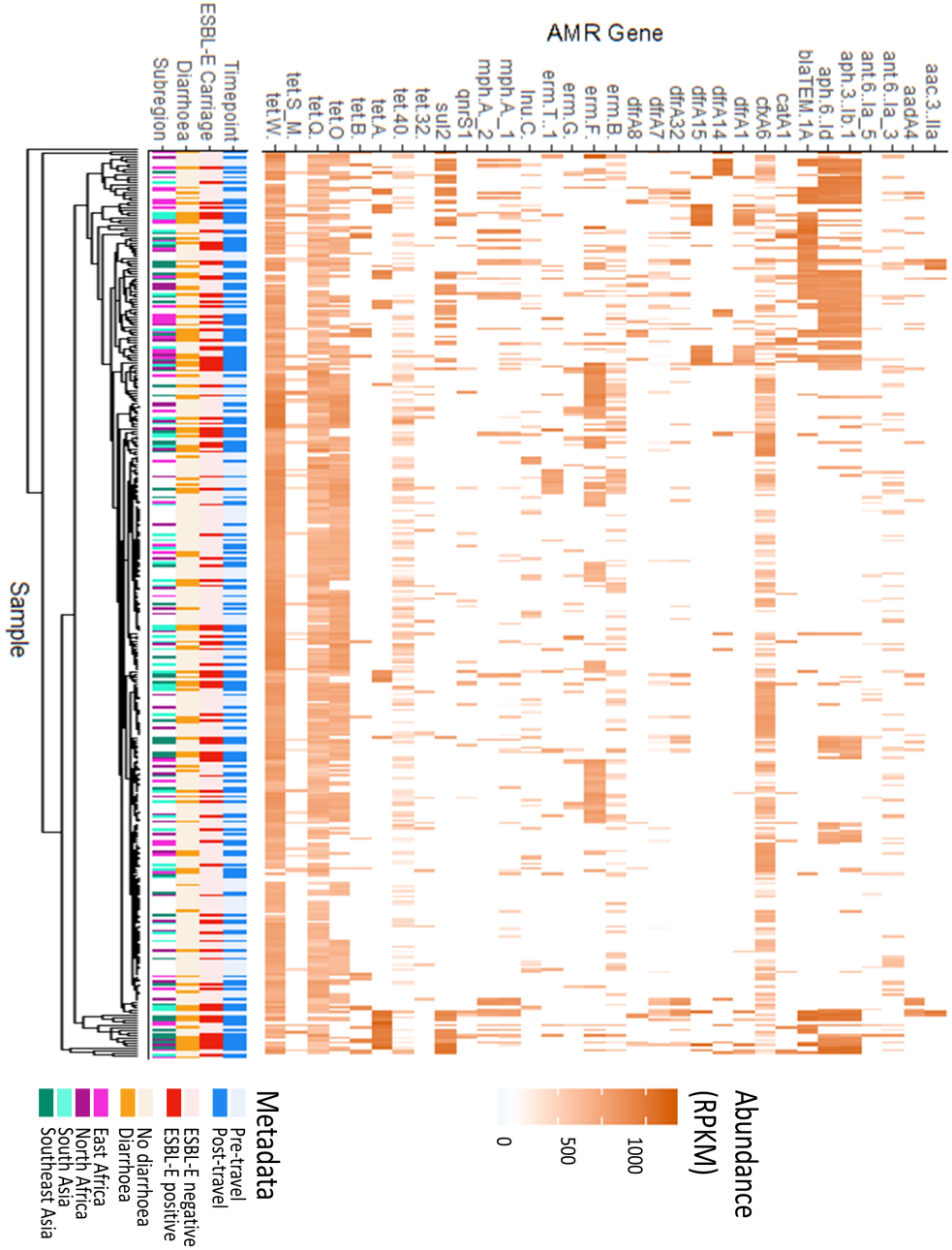

Fig. 2.12: AMR protein abundances during travel
Heatmap depicting the RPKM-normalised abundances of AMR protein families, as determined with ShortBRED from the ResFinder database. A darker shade represents a higher abundance. The metadata for Timepoint (blue), ESBL-E carriage (red), Diarrhoea (orange) and Subregion (Purple + Green) are added below. Samples are clustered by dissimilarity distance and visualised as a dendrogram.

To examine this pattern further, the AMR gene richness was calculated and compared across timepoint (now including low prevalent AMR genes as these are relevant to each individual). There is a significant increase (coefficient estimate: 2.465, 95% CI [1.52, 3.41]; p = <0.001) in AMR richness during travel. When adjusting for potential confounding factors, generalised linear modelling reveals that the onset of travellers' diarrhoea is a significant driver of this (coefficient estimate: 1.385, 95% CI [0.33, 2.44]; p = 0.0107).

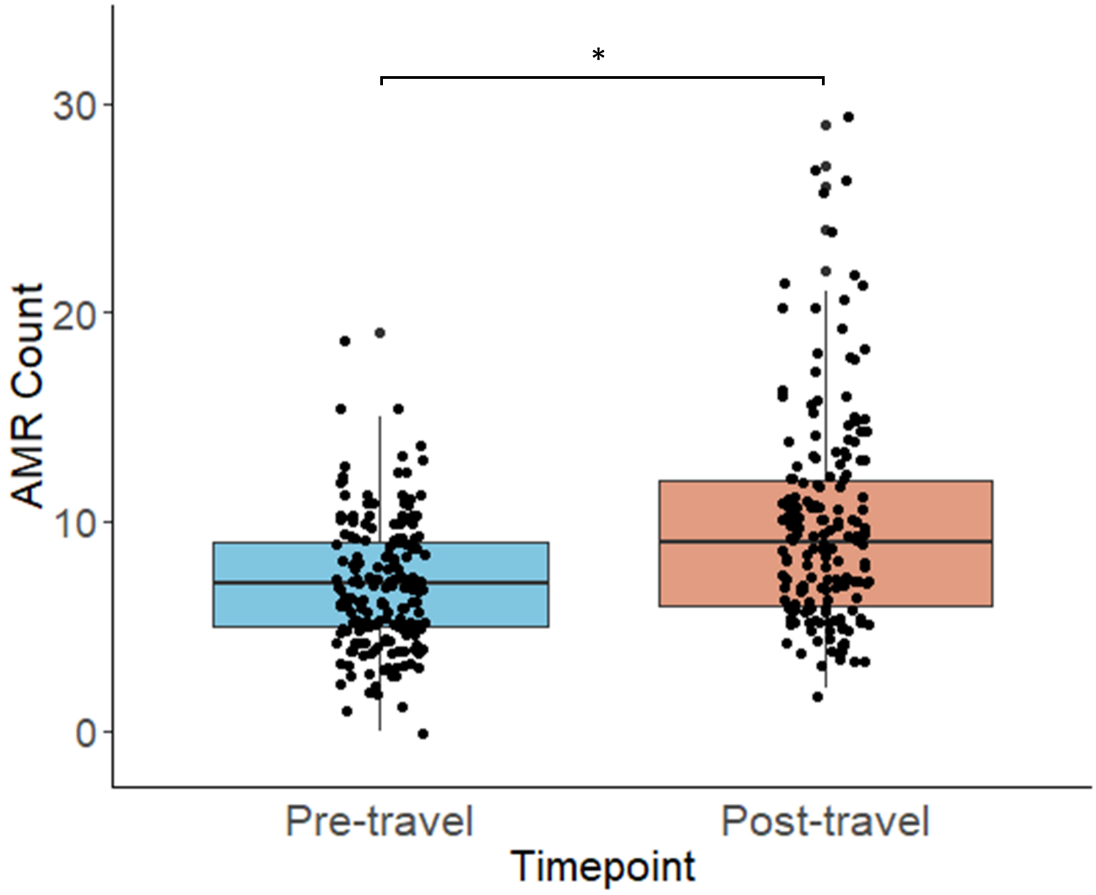

Fig. 2.13: AMR gene richness in association to timepoint
AMR gene richness in faecal samples collected both before and after international travel. Difference in AMR gene richness between pre- and post-travel samples was tested using generalised linear regression analysis, and a significant difference is visualised with the p-value. To adjust for potential confounding factors, AMR gene richness was treated as the dependent variable and age, sex, BMI and diarrhoea before or during travel as independent variables.

To determine the extent at which acquiring ESBL-E during travel affects the acquisition of other AMR genes, the AMR gene richness was also compared between travellers who did and did not acquire ESBL-E. This time, however, only post-travel samples were included in the analyses. Travellers colonised with ESBL-E have a significantly higher amount of AMR genes (coefficient estimate: 2.053, 95% CI [0.55, 3.56]; p = 0.008) compared to travellers who remain uncolonised. When adjusting for potential confounding factors, generalised linear modelling reveals that the onset of travellers' diarrhoea is a significant driver of this (coefficient estimate: 1.717, 95% CI [0.25, 3.18]; p = 0.023).

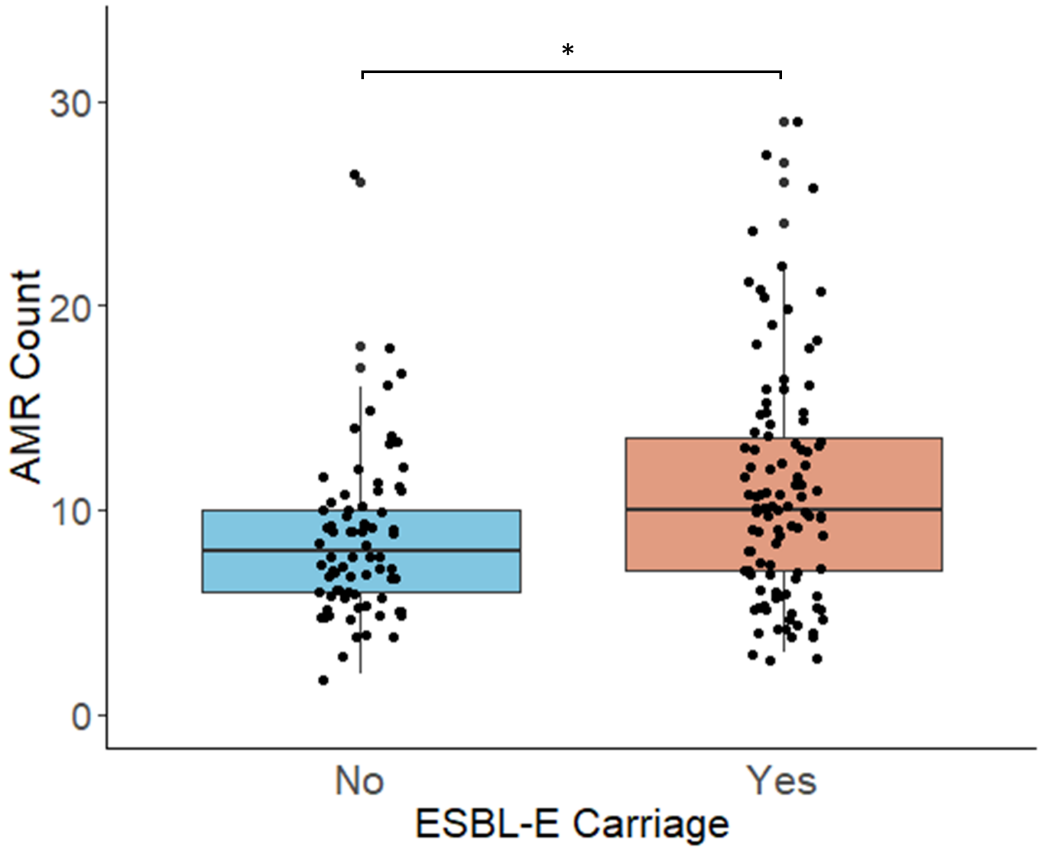

Fig. 2.14: AMR gene richness in association to ESBL-E carriage
AMR gene richness in faecal samples collected after international travel. The AMR gene richness of post-travel samples was calculated and the difference between travellers who did and did not acquire ESBL-E during travel was tested using generalised linear regression analysis, and a significant difference is visualised with the p-value. To adjust for potential confounding factors, AMR gene richness was treated as the dependent variable and age, sex, BMI and travellers' diarrhoea as independent variables.

2.4 Discussion

The gut microbiome's role in providing colonisation resistance against invading bacteria is of particular relevance in people travelling internationally to countries in which MDR bacteria are endemic. This study aimed to discover which aspects of the microbiome structure – or consortia of bacteria – contribute to colonisation resistance against MDR bacteria. Neither the microbial diversity and community structure in pre-travel samples nor the longitudinal change of these metrics during travel were predictive of the susceptibility to acquiring ESBL-E. Alternatively, prevalence and abundance changes in individual microbial taxa belonging to the family *Enterobacteriaceae* and genus *Bacteroides* were significantly associated with ESBL-E colonisation risk.

Studies focusing on the association between colonisation with MDR bacteria and the gut microbiota primarily utilise 16S rRNA gene amplicon methods, which have limitations in their lack of resolution to a species level analysis, and/or cross-sectional studies, which do not capture the longitudinal microbial changes. ESBL-E acquisition in travellers has been researched with meta-transcriptomics [139], where reads were mapped to 16S and 23S rRNA databases for identification, and small-scale WMGS on a data set of up to 10 travellers [140]. There is therefore a limited scope and depth of current research into the role of the microbiome in this field, so in this study, the WMGS of 179 travellers provides a significant contribution to the closure of a major gap in the literature.

In our study, the diversity and community composition did not differ between travellers before travel, or within individuals during travel, in relation to ESBL-E acquisition. Instead, after the development of travellers' diarrhoea, the diversity is significantly reduced and community composition is drastically altered, consistent with the literature [150, 139, 140]. Travellers' diarrhoea is known to be associated with an increase in ESBL-E acquisition [47], but ESBL-E acquisition without travellers' diarrhoea seemingly has little effect on the microbiome richness and diversity. This is unlike previous studies that associated increased pathogenic bacteria with a reduced microbial richness and altered microbiome in hospital patients [182]. However, samples collected from hospitalised patients are often confounded by antibiotic use, which reduces the diversity and richness measurements, as well as significantly perturbing even stable microbiomes. The results here are therefore suggesting that ESBL-E can colonise a traveller's microbiome without perturbation and without causing symptoms, but in

those who develop travellers' diarrhoea, it is difficult to determine if ESBL-E colonisation precedes the onset of diarrhoea.

In this study, MGS both identified as *Citrobacter freundii* had a statistically significant increase in abundance in travellers who did not acquire ESBL-E. Commensal *Enterobacteriaceae*, of which *Citrobacter* is a member, are continuously competing with other commensal or invading members of the family. This is achieved through various mechanisms, such as the secretion of antimicrobial proteins [141], the metabolism of nutrients [142] or competition for oxygen [183]. These mechanisms may change the host's susceptibility to colonising species [184], as it must have the necessary pathways to outcompete. Further studies on the interactions between *C. freundii* and ESBL-E in the human gut are required to uncover the mechanisms by which ESBL-E colonisation might be inhibited by *C. freundii*.

A statistically significant increase in abundance of two MGS of *Bacteroides* was also associated with travellers who remained negative for ESBL-E, whereas a decrease of these MGS as well as an additional *Bacteroides* MGS was observed in travellers who acquired ESBL-E. The healthy gut microbiome is usually dominated by the strict anaerobe phyla of *Bacteroidetes* (of which *Bacteroides* is a member) and *Firmicutes*, whereas the facultative anaerobes of *Enterobacteriaceae* populate at a much lower proportion [185]. Interestingly, in a recent cross-sectional study on the association between the faecal microbiota and ESBL-E colonisation among 200 healthy volunteers living in rural Thailand, the most profound difference was detected in the phylum *Bacteroidetes*. In particular, the abundance of *B. uniformis* was significantly lower in ESBL-E carriers as compared to noncarriers [186].

Various *Bacteroides* species exhibit immunomodulatory effects that are beneficial to the host and can increase the colonisation resistance to invading bacteria. A major mechanism of this is the inhibition of host inflammation via the secretion of sphingolipids [187], a signalling molecule, or outer-member vesicles containing polysaccharide A [188], a molecule that induces the production of T cells to suppress inflammation [189]. Low-grade intestinal inflammation perturbs the gut microbiome and can enhance the colonisation of *E. coli* [190], so reducing this is vital to maintain colonisation resistance to invading bacteria. *Bacteroides* species may have further immunomodulatory effects via the improvement of host macrophage's phagocytic function [191] or via an increased xylan degradation [192], which has links to an

enhanced immune system and improved health [193, 194]. Moreover, *Bacteroides* species can produce a variety of short-chain fatty acids and organic acids, including acetic acid, propionic acid, isovaleric acid and succinic acid. Short-chain fatty acids have been shown to provide colonisation resistance against antibiotic-resistant *Enterobacteriaceae*, among others, by triggering intracellular acidification [195]. To reveal a potential protective role of metabolites, including sphingolipids, xylan breakdown products and short-chain fatty acids, produced by *Bacteroides* species or other members of the microbiome, future studies on the faecal metabolome are warranted.

We found a significant increase in the diversity of AMR genes during travel, especially in association with ESBL-E acquisition. Prior research using a different AMR gene reference database but on the same data cohort also found an increase in the diversity of AMR genes during travel [196], and hypothesised that travel is a significant perturbation to the gut resistome. They did, however, also find that the beta-diversity decreased during travel. When comparing the patterns in AMR diversity with the taxonomic data of our study, both changes seem to occur often without a perturbation to the gut microbiome microbial composition. Together, this shows how the microbial composition is not drastically altered during travel, but the few distinct species that are acquired have a similar, high number of AMR genes.

A major strength of this study is the use of whole metagenome shotgun sequencing as a representation of the total microbiome. The longitudinal nature of the data set additionally strengthens the study as it allows for temporal associations on the effects of travel on the microbiome, with a large sample size to correct for confounding factors. However, metagenomics limits the resolution at which to confidently identify individual strains. Instead, the taxa-specific analysis should be interpreted as where to focus potential future research. This study has a limitation with the sampling method employed, as the transport at room temperature and use of Cary–Blair medium on faecal swabs can increase the microbial richness and diversity [197] and can lead to increase in abundance of taxa, notably *Enterobacteriaceae*, that can grow at room temperature in the presence of oxygen [198]. However, as the samples all experienced the same conditions, the effects are not likely to differ between samples; instead, the quantities should not be extrapolated to other sampling methods.

The observations of this study highlight the likely importance of discrete species in the dynamics of ESBL-E acquisition, especially within the family *Enterobacteriaeceae*.

The identification of species associated with ESBL-E acquisition necessitates for higher resolution research. The current analyses cannot easily detect how the commensal *Enterobacteriaceae* population is interacting with the invading ESBL-producing *E. coli*, so it would be of significant interest to perform experimental studies to explore the competition between various species and strains within this genus and family. The extent at which each member alters the microbiome's colonisation resistance could be discovered. Alternatively, the knowledge on factors facilitating the elimination of previously acquired ESBL-E is still limited, so research on samples taken on additional time points after travel is worthwhile being examined.

This study emphasises how the overall microbiota diversity, community structure and functional metabolism play little role in the acquisition of ESBL-producing *Enterobacteriaceae*, instead highlighting the importance of *Bacteroides* and other members of *Enterobacteriaceae*. The discovery of taxa that may contribute to colonisation resistance will aid in the development of intervention strategies to minimise the acquisition and spread of MDR bacteria.

Chapter Three

Determining the ability of travel-acquired *Citrobacter* to compete *in vitro* with extended-spectrum beta-lactamase producing *Escherichia coli*

3.1 Background

Citrobacter is a genus of the family *Enterobacteriaceae* and is found in the human gut. *Citrobacter* is considered an opportunistic pathogen as it can be responsible for a number of bloodstream, urinary tract or pulmonary infections, albeit less commonly than other members of *Enterobacteriaceae* like *E. coli* or *Klebsiella*. All species of *Citrobacter* have been isolated from clinical infections, except *C. rodentium* which is a mouse-specific pathogen that is used as a model organism in the study of human intestinal disease [199]. *C. rodentium* colonises the gut by attaching to the intestinal epithelium in a manner similar to both enterohaemmorrhagic and enteropathogenic *E. coli* [200]. Interestingly in kanamycin-treated mice, *C. rodentium* was displaced from the colonic musoca and caecum into the caecal lumen where it stabilised, except in mice with *C. amalonaticus* already blooming in the caecal lumen as these outcompeted in a contact-dependent manner [143]. This highlights the potential that species of *Citrobacter* could outcompete invading strains of *E. coli*.

Colonisation resistance to pathogens is a vital characteristic of the gut microbiome, and previous research in Chapter 2 suggests that *C. freundii* may provide a colonisation resistance to invading ESBL-E in travellers. This warranted further research of this organism and its ability to compete with *E. coli*. Here, we grew travel-acquired *Citrobacter* with both antimicrobial-sensitive and ESBL-producing *E. coli* to determine if *Citrobacter* has the ability to outcompete *E. coli* in a contact-dependent manner and explain how it may be providing a colonisation resistance within travellers.

3.2 Methods

3.2.1 Isolate selection

Commensal *E. coli* were isolated from healthy volunteers from the West Midlands, United Kingdom, courtesy of Professor Peter Hawkey. Volunteer inclusion required no history of infection or antimicrobial treatment in the prior 12 months. *Citrobacter* and ESBL-producing *E. coli* were selected from [155] during international travel to Laos.

Isolate sequences were trimmed with Trimmomatic [201] and *de novo* assemblies were created with SPAdes [202], both with default settings. The multi-locus sequence type (MLST) of *E. coli* was determined with MLST (https://github.com/tseemann/mlst) with PubMLST database [203]. ESBL-producing *E. coli* were selected from isolates with a chromosomally inserted CTX-M gene. Abricate (https://github.com/tseemann/abricate) was used to determine which contigs contained a CTX-M gene and contigs larger than 10kb and containing a CTX-M gene were classified by mlplasmids [204] into chromosomal or plasmid of origin. Four *E. coli* isolates, representing various phylogroups and STs, were chosen from both the commensal and the ESBL-producing *E. coli* groups for subsequent competition analysis.

Table 3.1: Description of *E. coli* used in this study (N=8)

Taxon	Phylogroup	ST		Phenotype on UTI *ChromoSelect* Agar	Phenotype on MacConkey Agar
E. coli	B2	95	Commensal	Purple	Red
E. coli	B2	73	Commensal	Purple	Red
E. coli	A	10	Commensal	Purple	Red
E. coli	D	69	Commensal	Purple	Red
E. coli	B1	515	$bla_{\text{CTX-M-14}}$	Purple	Red
E. coli	D	69	$bla_{\text{CTX-M-14}}$	Purple	Red
E. coli	A	48	$bla_{\text{CTX-M-14}}$	Purple	Clear
E. coli	D	38	$bla_{\text{CTX-M-15}}$	Purple	Red

28 isolates that initially identified as *Citrobacter* in the Laos traveller study were selected. To taxonomically identify to species level, the core genomes were compared to publicly available reference *Citrobacter* genomes. Up to 20 reference sequences, if available, for each named *Citrobacter* species were downloaded from NCBI RefSec ([205]; Accessed: 21/01/22). This included 20 isolates from *Citrobacter amalonaticus, Citrobacter braakii, Citrobacter cronae, Citrobacter freundii, Citrobacter koseri, Citrobacter portucalensis, Citrobacter werkmanii*; 17 *Citrobacter farmeri*; 16 *Citrobacter youngae* and less than 10 from each of *Citrobacter arsenatis, Citrobacter europaeus, Citrobacter gillenii, Citrobacter murliniae, Citrobacter pasteurii, Citrobacter rodentium, Citrobacter sedlakii, Citrobacter telavivensis* and *Citrobacter tructae*. Additionally, to determine if the sample was contaminated with other species, the assemblies were analysed with Kraken2 [206]. Isolates with more than 45% identity to *Citrobacter* and less than 10% identity to a single other genus were considered for competition analysis.

To detect *E. coli* within the travel-acquired *Citrobacter* sequences, the distance between the genomes of all reference and traveller *Citrobacter*, as well as 40 reference *E. coli* genomes, was calculated and visualised as a tree with mashtree [207]. To cluster genomes by their core genome, *Citrobacter* assemblies were annotated with Prokka [208] and the core genomes were calculated and aligned with panaroo [209], with settings "-clean-mode strict" and "-a core". A phylogenetic inference by maximum likelihood was performed on the core gene alignment by IQ-TREE [210] using the 'GTR +G +I' substitution model.

All isolates of *Citrobacter* and *E. coli* that were candidates for competition analysis were grown on UTI *ChromoSelect* Agar (Sigma-Aldrich) and MacConkey Agar (Sigma-Adlrich) to see if *Citrobacter* isolates and *E. coli* isolates were distinguishable between each other. If there was no phenotypic difference between the genera then they were removed from downstream analysis. Four *Citrobacter* isolates that were phylogenetically diverse, consisted of a pure culture and were distinguishable (based on their colony morphology) from *E. coli* were chosen for subsequent competition analysis. One ESBL-producing *E. coli* could not be distinguished from two of the *Citrobacter* isolates with the chromogenic agar used in this study. A total of 12 isolates were selected for two sets of competitions; *Citrobacter* vs commensal *E. coli* and *Citrobacter* vs ESBL-producing *E. coli*.

Table 3.2: Description of *Citrobacter* used in this study (N=4)

Taxon	**Closest Reference**	**Phenotype on UTI *ChromoSelect* Agar**	**Phenotype on MacConkey Agar**
Citrobacter 1	*C. amalonaticus*	Turquoise	Clear
Citrobacter 2	*C. amalonaticus*	Turquoise	Clear
Citrobacter 3	*C. amalonaticus*	Blue/Purple	Clear
Citrobacter 4	*C. amalonaticus*	Blue/Purple	Clear

3.2.2 Competition Assay

Isolates were stored in 50% glycerol at −80°C until required. Isolates were cultured for 15 hours in a shaking 37°C incubator and the culture was diluted 1:10 six times in phosphate-buffered saline (PBS). 1 μl of the 10^{-4} dilution was added to 99 μl Nutrient Broth in a flat-bottom 96-well plate, as this is approximately 1000 bacterial cells, and this was incubated in a Tecan (Tecan) plate reader for 18 hours at 37°C with OD 600 measurements at 30 minute intervals. This was repeated in duplicate. Length of incubation for the competition experiments was decided as eight hours due to isolates being approximately in their late log phase of growth.

For the competition experiments, overnight cultures were diluted 1:10 six times. In triplicate, 20 μl of the 10^{-6} dilution was plated onto Luria-Bertani agar and incubated at 37°C overnight. Equal amounts of 100 μl of the 10^{-5} dilutions of the two isolates being competed were added to 4.8 ml Nutrient Broth and incubated for eight hours at

37°C with aeration. The competed culture was then diluted 1:10 six times in PBS and 100 μl was plated out, in triplicate, onto either UTI *ChromoSelect* agar or MacConkey Agar, depending on which was needed to distinguish between the two species. The plates were then incubated overnight at 37°C.

The colonies from both the input culture plates and from the competed culture plates were counted and the colony forming units (CFU) were calculated using the Miles and Misra method [211]. A competition index (CI) is calculated as (*E. coli* output CFU/*Citrobacter* output CFU)/(*E. coli* input CFU/*Citrobacter* input CFU). The entire experiment was repeated in triplicate for each combination of isolates.

3.3 Results

3.3.1 Choice of *Citrobacter* isolates from the Laos traveller dataset to compete with *E. coli*

The design of the competition analysis was to compete different combinations of four isolates each of *Citrobacter*, commensal *E. coli* and ESBL-producing *E. coli*. The four isolates from each group were selected to consider a variety of phylogroups, sequence types and phenotypes (Table 3.1 and Table 3.2.1).

28 isolates from the Laos traveller dataset that had previously been considered as *Citrobacter* or 'Other' were taxonomically identified with Kraken2. 13 isolates had more than 45% identity to *Citrobacter* at a genus level, and less than 10% to a single other genus and were considered for downstream analysis. The assemblies of the travel-acquired *Citrobacter* isolates were compared to reference *Citrobacter* and *E. coli* genomes and the resulting tree is represented in Figure 3.1. No travel-acquired isolates clustered with the reference *E. coli* or separately from the reference *Citrobacter* and therefore no further isolates were removed as targets of competition analysis.

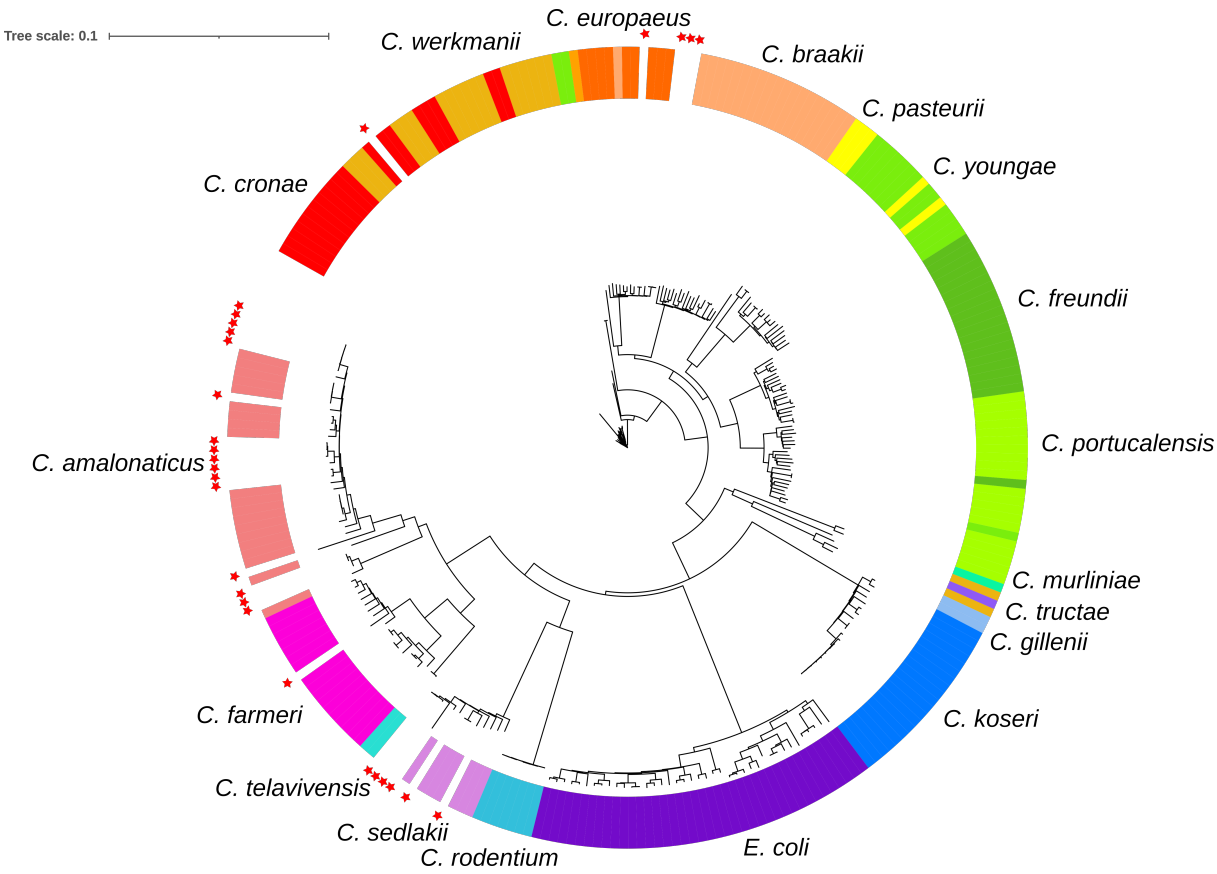

Fig. 3.1: *Citrobacter* and *E. coli* genomes as clustered by min-hash distances. Genomes from 208 reference *Citrobacter* from 17 species, 40 reference *E. coli* and 28 travel-acquired *Citrobacter* compared with Mashtree. Species of reference isolates are represented as different colours of the coloured bar, and travel-acquired *Citrobacter* are represented by red stars. Tree was visualised with iTOL.

The core genomes of both travel-acquired and reference *Citrobacter* were calculated and aligned (2483 genes present in >99% of sequences and an additional 232 genes present in >95% of sequences), and a phylogenetic inference was made (Figure 3.2). Considering the taxonomic identity from Kraken, the phenotype on chromogenic agar and the phylogenetic relatedness to reference *Citrobacter*, the four isolates chosen for competition analysis clustered closely to *C. amalonaticus*, the most commonly travel-acquired species of *Citrobacter* in this dataset after *C. sedlakii* and *C. europaeus*.

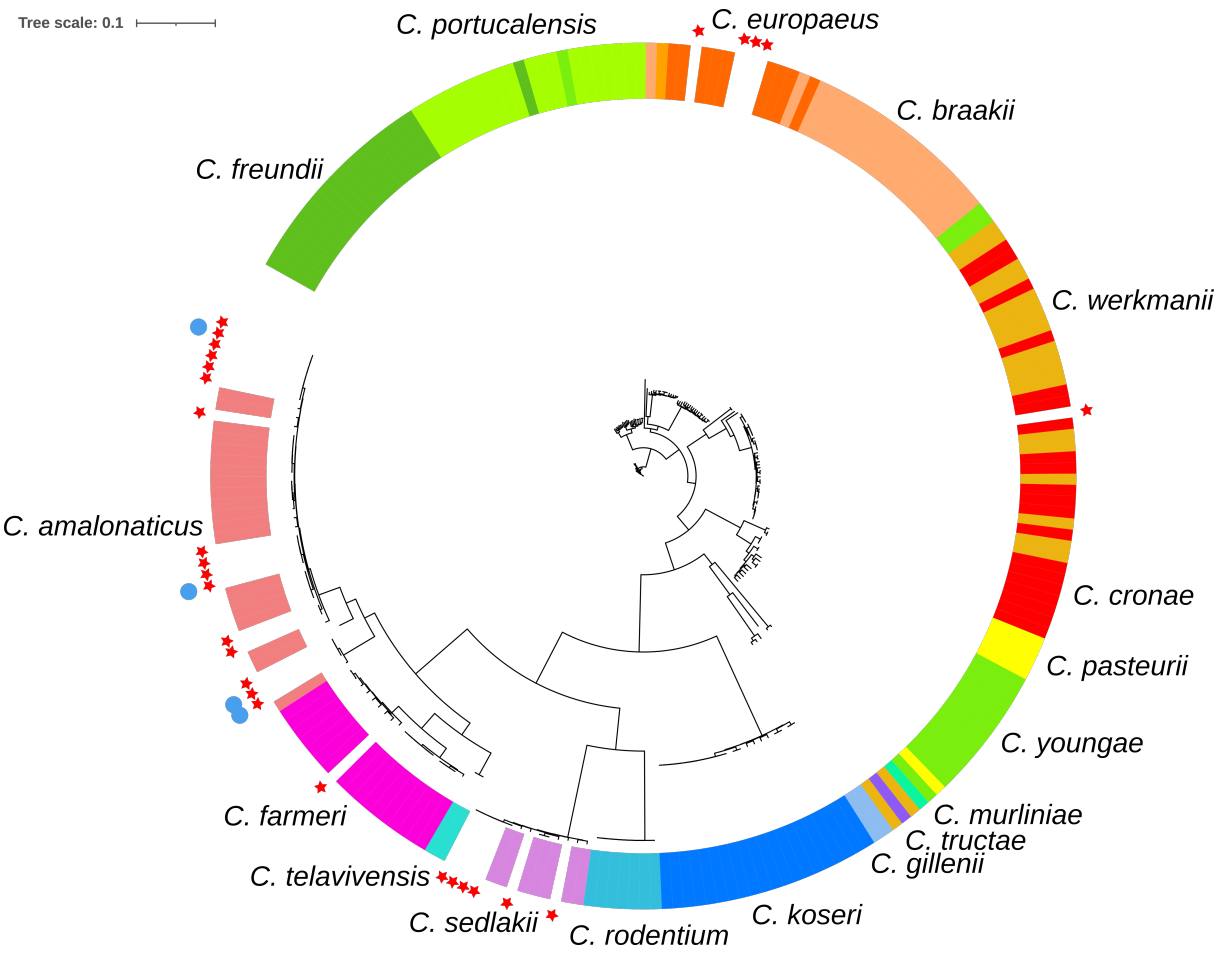

Fig. 3.2: Travel-acquired and reference *Citrobacter* phylogeny.
The core genes of 208 reference *Citrobacter* and 28 travel-acquired *Citrobacter* were aligned and a phylogenetic inference was performed with IQ-TREE. Species of *Citrobacter* are represented as different colours in the coloured bar, travel-acquired *Citrobacter* as red stars and the isolates chosen for competition analysis as blue circles.

3.4 Growth Kinetics

Citrobacter and *E. coli* isolates were incubated for 18 hours in Nutrient Broth and the OD was measured every 30 minutes and represented in Figure 3.3. Late logarithmic phase of growth is at approximately 8 hours after inoculation and was used as length of time for the competition assay.

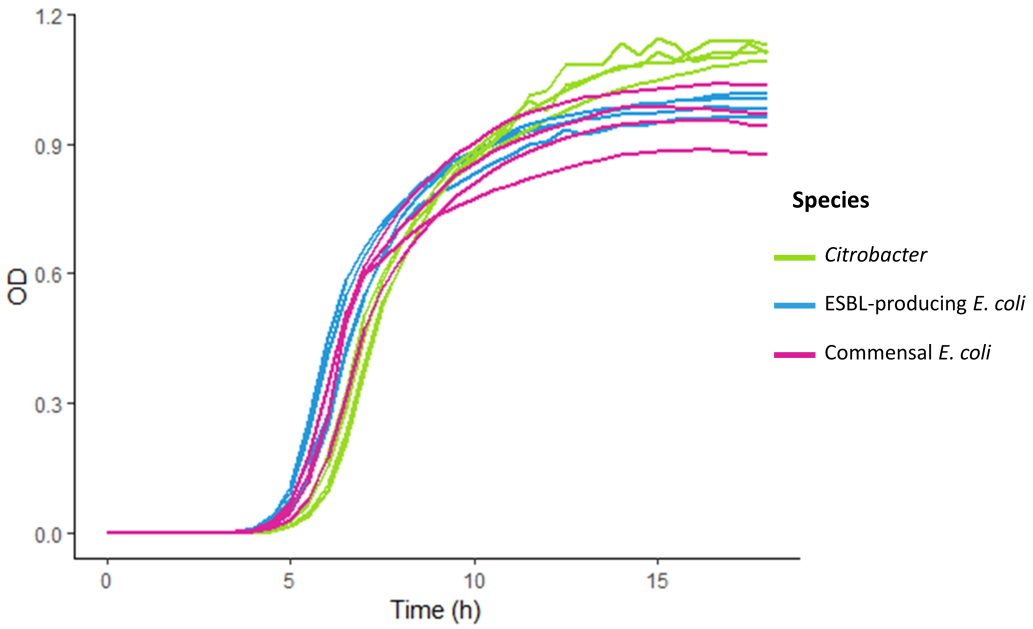

Fig. 3.3: Growth curves of *Citrobacter*, commensal *E. coli* and ESBL-producing *E. coli* at 37°C in Nutrient Broth
Bacterial growth curve for four isolates of each of *Citrobacter* (green), commensal *E. coli* (magenta) and ESBL-producing *E. coli* (blue) grown in Nutrient Broth at 37°C. OD measurements were taken every 30 minutes for 18 hours.

3.5 Commensal *E. coli* outcompete travel-acquired *Citrobacter*

All combinations of *Citrobacter* and commensal *E. coli* were competed in Nutrient Broth for 8 hours, plated onto chromogenic agar to differentiate between the two species and the resulting CFUs were counted. A competition index (CI) was determined as (*E. coli* output/*Citrobacter* output)/(*E. coli* input/*Citrobacter* input) so a CI of 1 represents equal competition. As all CIs are above 1, all commensal *E. coli* are outcompeting each isolate of *Citrobacter* (Figure 3.4). This is higher for *Citrobacter* isolates 3 and 4 compared to isolates 1 and 2.

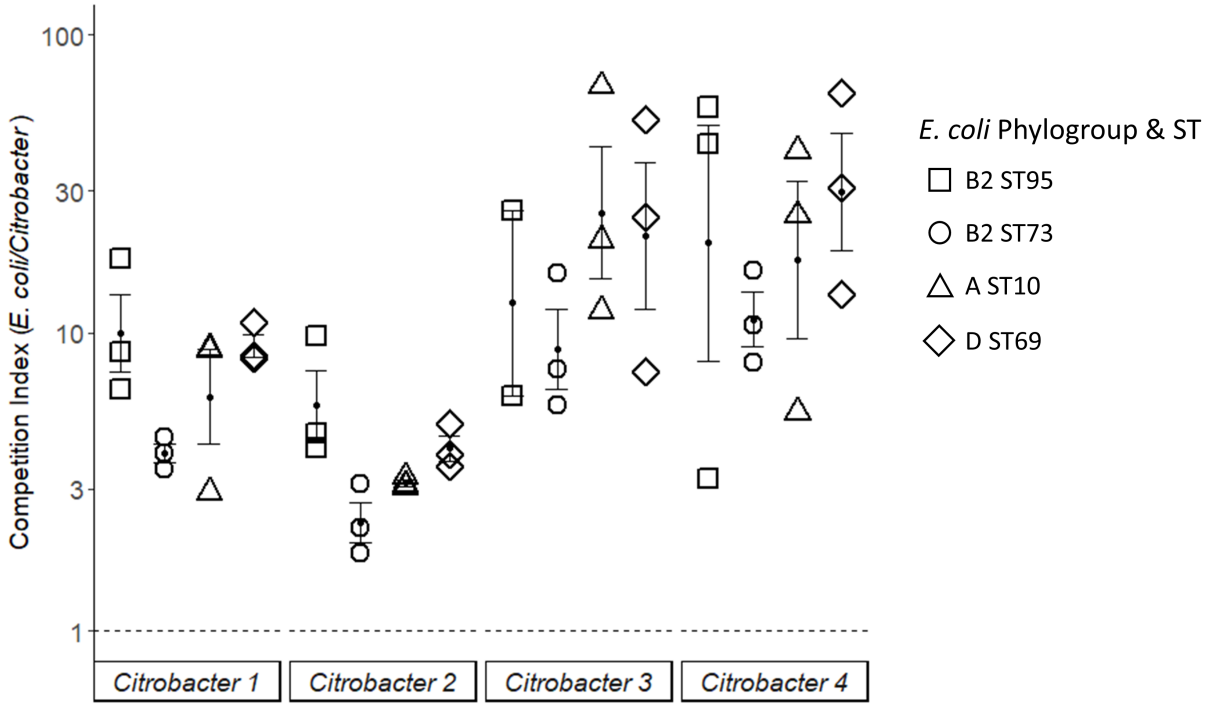

Fig. 3.4: Competition between commensal *E. coli* and travel-acquired *Citrobacter*.
An equal number of *Citrobacter* and *E. coli* cells were added to Nutrient Broth and incubated for eight hours. Competed culture was diluted and plated onto chromogenic agar and the number of isolates of each species was counted. Competiton Index (CI) is *E. coli*/*Citrobacter* and a CI larger than 1 represents competition in favour of *E. coli*; individual CIs are illustrated with shapes and the mean value and standard error are overlaid. Different shapes represent different *E. coli* strains (phylogroup, sequence type); square= B2 95, circle= B2 73, triangle= A 10, diamond= D 69.

3.6 Potential for ESBL-production does not decrease ability of *E. coli* to outcompete travel-acquired *Citrobacter*

All combinations of *Citrobacter* and ESBL-producing *E. coli* were also competed, except for one phylogroup A sequence type 48 isolate of *E. coli* as it was indistinguishable from *Citrobacter* on the chromogenic agar used. Except for one repeat of *E. coli* ST48 with *Citrobacter* isolate 2, all ESBL-producing *E. coli* also outcompete each isolate of *Citrobacter*, and is higher for isolates 3 and 4 compared to 1 and 2 (Figure 3.5).

3.7 ESBL-producing *E. coli* also outcompete travel-acquired *Citrobacter*

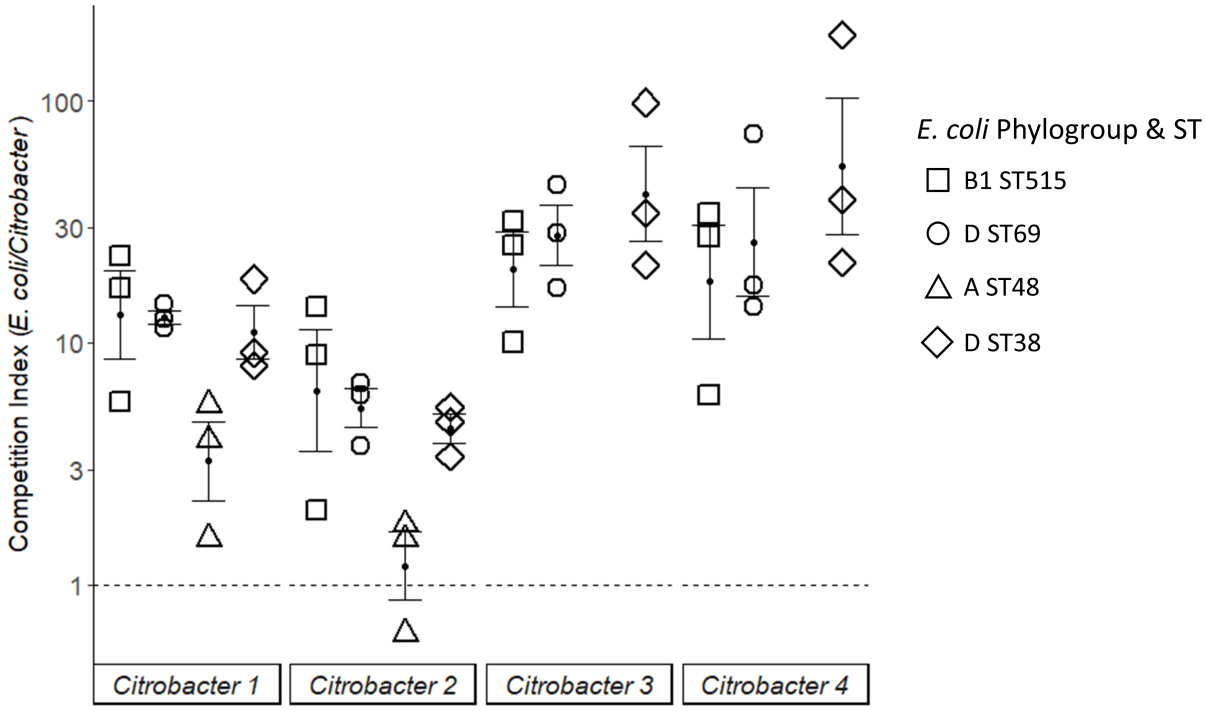

Fig. 3.5: Competition between ESBL-producing *E. coli* and travel-acquired *Citrobacter*.
An equal number of *Citrobacter* and *E. coli* cells were added to Nutrient Broth and incubated for eight hours. Competed culture was diluted and plated onto chromogenic agar and the number of isolates of each species was counted. Competiton Index (CI) is *E. coli*/*Citrobacter* and a CI larger than 1 represents competition in favour of *E. coli*; individual CIs are illustrated with shapes. Different shapes represent different *E. coli* strains (phylogroup, sequence type); square= B1 ST515, circle= D 69, triangle= A 48, diamond= D 38. Triangle is missing on *Citrobacter 3* and *Citrobacter 4* due to being indistinguishable on the chromogenic agar used. Error bars represent standard error.

3.8 Discussion

Previous results suggested that within-*Enterobacteriaceae* interactions could be creating a colonisation resistance against ESBL-E, so this study aimed to test if *Citrobacter* can outcompete *E. coli in vitro* and if this is an *E. coli* phylogroup dependent manner. The strains of *E. coli* used in this experiment outcompete travel-acquired *Citrobacter amalonaticus*. However, the level at which it is outcompeted varies by both *Citrobacter* isolate and *E. coli* strain.

Travel-acquired *Enterobacteriaceae* is often the focus of traveller studies, yet *Citrobacter* is a genus that is still under-researched in this field, usually as it is sampled from clinical infections. This study provides a unique perspective on travel-acquired *Citrobacter* which, as Figure 2.9 in Chapter 2 shows, can be found in nearly 25% of international travellers returning home. *C. amalonaticus*, a facultative, chemoheterotrophic species known for its ability to produce hydrogen [212], was the most frequently acquired species of *Citrobacter* when travelling to Laos.

The aim of the isolate selection in this study was to, where possible, ensure strain variety and diversity among the competitions. As the previous microbiome study in Chapter 2 used data from travellers, travel-acquired *Citrobacter* and travel-acquired ESBL-producing *E. coli* were selected so that meaningful comparisons can be made. Additionally, these were selected from a study of travel to Laos, as Southeast Asia has high rates of ESBL-E acquisition [47]. Although Figure 2.9 suggests that *C. freundii* is the species of interest, there was no *C. freundii* acquired in the Laos dataset. The analysis for Figure 2.9 taxonomically identified an MGS by keeping the top 50 hits for each gene, yet more *C. freundii* is present in the Blast nucleotide database than any other species of *Citrobacter*, creating a bias towards it. This bias is especially strong in the 2483 core genes shared across all species of *Citrobacter* as they are more frequently annotated as *C. freundii*. Regardless, many species within the family *Enterobacteriaceae* are known to compete with each other and potentially provide colonisation resistance [141], so researching any travel-acquired members of *Enterobacteriaceae* is relevant. Even though this study focuses solely on *C. amalonaticus*, it has been observed as providing a colonisation resistance to *C. rodentium in vivo* [143], which shows it has the capability to outcompete closely related organisms and is a meaningful organism to study.

A difference in competition index is observed between *Citrobacter* isolates. Isolates 1 and 2 were outcompeted to a lower level than isolates 3 and 4 and, as determined from plating on UTI ChromoSelect agar, isolates 1 and 2 have β-glucosidase activity whereas isolates 3 and 4 have both β-glucosidase and β-galactosidase activity. Although the four isolates clustered with reference *C. amalonaticus* genomes, both isolates 3 and 4 have some genetic diversity from the rest of the genomes in the species and do not cluster as closely. Together, subtle differences in phenotype and genotype could be causing a drastic change to their ability to compete with *E. coli*. Notably, the experiment in this study used media that does not represent the complexity of the gut microbiome or contain various metabolites available in the gut. For example, some genera including *Citrobacter*, *Klebsiella* or *Lactobacillus* are able to ferment the highly prevalent dietary metabolite glycerol [213, 214]. The fermentation of glycerol by these genera can produce the antimicrobial agent reuterin [215] which has been observed decreasing the abundance of *E. coli* in the gut [216]. *Citrobacter* and *E. coli* could therefore be competed against each other in a variety of media to discover if the presence of growth substrates, in both aerobic and anaerobic conditions, which could modulate the ability of *Citrobacter* to compete with *E. coli*.

The gut microbiome is orders of magnitudes more complex than just two species interacting with each other, and an *in vitro* experiment in a nutrient rich environment is missing a multitude of host and microbiota factors. Within the gut, there are spatial niches that the microbiota occupy. For example, *C. rodentium*, a model organism for murine enteric infections, was displaced from mice colonic mucosa and caecum into the caecal lumen after treatment with kanamycin, but only the mice with *C. amalonaticus* blooming in the lumen were not recolonised with *C. rodentium* after treatment of kanamycin ended [143]. This example highlights the potential of *C. amalonaticus* to compete with ESBL-producing *E. coli* by occupying spatial niches in the gut, but the *in vitro* nature of this experiment discovered that *E. coli* is more adapted at replicating quicker in nutrient rich media. The interaction between *Citrobacter* and *E. coli* could instead be studied in mouse models to discover if competition occurred in distinct regions of the digestive tract, and how this changes over time. Additionally, metabolomics and transcriptomics of the gut microbiome could then also be performed to understand further what metabolic changes occur during addition of competing bacteria.

As *Citrobacter* colonisation was shown to be associated with a lower prevalence of ESBL-E acquisition (Figure 2.9), we had predicted that ESBL-producing *E. coli* would

be less able to compete with *Citrobacter* than commensal *E. coli*. However, the ESBL-producing *E. coli* chosen for this experiment have chromosomally inserted CTX-M genes in order to avoid confounding the experiment with maintaining carriage of a plasmid containing CTX-M genes. In some cases there is a fitness cost to hosting a plasmid [217], and this cost may be lost in isolates with chromosomally inserted AMR genes and explain why there is no decrease in competition index for ESBL-producing *E. coli*.

This study shows that *E. coli* are able to outcompete *C. amalonaticus* in a nutrient rich media, and this occurs regardless of the carriage of AMR genes for ESBL-production. However, there are many interactions within the gut microbiome to consider that may explain how within-*Enterobacteriaceae* competition is occurring, and these were not captured with this *in vitro* competition assay.

Chapter Four

Phylogenetic analysis of the change in *Escherichia coli* colonising the human gut during international travel

4.1 Introduction

Escherichia coli (*E. coli*) is a commensal member of the human intestinal flora [218] but also a major contributor to intra- and extra-intestinal infections [219]. Usually, most people harbour only one or two different strains of *E. coli* in the gut, though it is possible to be colonised by various strains [220]. Although commensal *E. coli* can exist in the gut without causing harm to the host, there can be a shift from commensalism to pathogenesis, whether opportunistically or with specialised virulence mechanisms. Commensal *E. coli* can also act as a reservoir of antimicrobial resistance (AMR) genes [221] and are able to transfer these AMR genes and plasmids to other pathogenic species [222]. Tracking of *E. coli* populations that exist at an individual, community and global scale is vital in monitoring for disease and AMR.

The genetic background of *E. coli* creates differences in diversity of potential host, pathogenecitiy and virulence. For example, in a commensal population phylogroup B2 was found to have higher virulence potential than all other phylogroups, based on 20 virulence factors involved in production of adhesins, protectins, toxins and iron acquisition [223]. Phylogroup B2 is also frequently the lineage containing the extraintestinal pathogenic *E. coli* (ExPEC) causing bloodstream and urinary tract infections in humans [101, 102]. On the other hand, enterotoxigenic *E. coli* (ETEC), the

causative agent of diarrhoeal diseases that cause 18,700-42,000 deaths per year [224], are predominantly from phylogroups A and B1 [225]. The within-phylogroup level of virulence can also vary [223]. As well as virulence and pathogenicity, levels of AMR can vary between phylogroup, like extended-spectrum beta-lactamase (ESBL)-producing *E. coli* most frequently belonging to phylogroups A and D in a healthy German community [226], or phylogroup B2 in Dutch hospitalised patients [227].

To further emphasise the complexity and fluidity of *E. coli* populations, lineage prevalence can vary temporally and by country. In a healthy Parisian population, phylogroup A drastically decreased in prevalence and was replaced with phylogroup B2 between 1980 and 2010 [223, 228]. Large phylogroup differences are generally seen between tropical and temperate environments [144], likely due to a number of factors such as geographic variations, socioeconomic status and diet [218, 229]. As total numbers of pre-SARS-CoV-2 pandemic tourists reached an estimated 1 billion [146], geographic distinction of *E. coli* and its varying degree of AMR is of particular importance in understanding the risk that international travel poses on the spread of AMR.

International travel is one of the largest contributors to the spread of AMR bacteria globally. Up to 84% of travellers to India and 42% of travellers to Morocco acquired ESBL-producing *Enterobacteriaceae* (ESBL-E), predominantly *E. coli*, whereas pre-travel levels sit at around 6.2% [47]. As seen in Chapter 2, the diversity and community structure of the gut microbiome is not associated with the carriage of ESBL-E or its acquisition during travel. However, the change in abundance and prevalence of discrete species, members of *Enterobacteriaceae* but not *E. coli* itself, may be involved.

It is of note that abundance and prevalence changes of *E. coli* were not associated with acquiring ESBL-producing *E. coli* in Chapter 2, suggesting that commensal *E. coli* are replaced with invading *E. coli*. We hypothesise that displacement of *E. coli* occurs in a strain-specific manner as STs with a higher amount of virulence factors will be better able to colonise the host, and that ESBL-producing *E. coli* belong to these virulent STs which drives the previously described high levels of ESBL-E acquisition. Here, we focused on the *E. coli* present in faecal swabs of 179 travellers before travel and immediately upon return from South Asia, Southeast Asia, North Africa and East Africa. Metagenome sequences were binned into metagenome assembled genomes

(MAGs), and *E. coli* MAGs were selected per sample and assigned to phylogroups and STs. Differences across time point, travel destination, onset of travellers' diarrhoea or acquisition of ESBL-E were observed in order to associate factors involved in a change to the *E. coli* detected in faecal samples of travellers. We determined if the level of virulence potential or abundance of AMR genes drives this change in *E. coli*.

4.2 Methods

4.2.1 Production of metagenome-assembled genomes (MAGs)

Raw sequencing reads were selected from Section 2.2.1 and trimmed with Trimmomatic [201] and subsequently assembled with metaSPAdes [230]. Reads were mapped back to their assemblies in order to calculate contig depth and coverage. Contigs that were too small or shallow were removed with the script jgi_summarize_bam_contig_depths [231].

Contigs were binned via Metabat 2 (v2.12.1) [231], with a minimum contig size of 2000 bp and parameters minCV 1.0, minCVSum 1.0, maxP 95%, minS 60 and maxEdges 200. Parameters were decided by checking for contamination and completeness of bins with dRep [232], which utilises checkM [233]. A default minimum completeness was used and a maximum contamination of 10% was allowed. Of the 4430 input bins, 1446 (32.6%) passed checkM filtering and these dereplicated into 788 MAGs.

4.2.2 MAG taxonomic identification

MAG DNA sequences were converted into amino acid sequence via prodigal [234]. These were aligned to DIAMOND database [235] (Accessed: September 2021) for identification with parameter –max-target-seqs 10. 50 dereplicated MAGs were identified as *Enterobacteriaceae* and selected. The bins that contributed to each dereplicated *Enterobacteriaceae* MAG were selected from the 1446 non-dereplicated bins that passed quality measures. A total of 229 bins now represent the *Enterobacteriaceae* strain variation across all samples.

4.2.3 *E. coli* identification

In order to identify MAG *E. coli* sequence type, the core genome of MAGs was compared to the core genome of various reference sequence types.

Reference *E. coli* were downloaded from Enterobase *E. coli* database [236] with the Python script EnterobaseGenomeAssemblyDownload [237]. The downloaded sequences consisted of 20 genomes from 21 STs that represent the diversity of *E. coli* lineages found across its phylogeny; ST3, ST10, ST11, ST12, ST14, ST17, ST21, ST28, ST38, ST69, ST73, ST95, ST117, ST127, ST131, ST141, ST144, ST167, ST372, ST410 and ST648.

MAGs fasta files were annotated into gff format with Prokka [208]. Core genes of all MAGs and reference *E. coli* were calculated and aligned with Panaroo [209] with the flags '–clean-mode strict' and '-a core'. A phylogenetic inference by maximum likelihood was performed on the core gene alignment by IQ-TREE [210] using the 'GTR +G +I' substitution model. 14 MAGs that grouped distantly from *E. coli* references were determined as a different species and removed from analysis. Metadata was added to the tree: time point (pre-travel/post-travel), post-travel carriage of ESBL-E (Yes/No); UN subregion travelled to (East Africa, North Africa, South Asia, Southeast Asia), diarrhoea within 3 months before travel (pre-travel samples) or developed during travel (post-travel samples, Yes/No), consumption of food from street vendors (No/Occasionally/Daily).

ClermontTyping [238] was used with default parameters to assign each MAG and reference *E. coli* a phylogroup. A further two MAGs identified as non-*Escherichia* were removed from analysis. Multi-locus sequence type (MLST) of MAGs was identified with MLST (https://github.com/tseemann/mlst) with PubMLST database [203]. Antimicrobial resistance genes were identified with Abricate (https://github.com/tseemann/abricate) using the resfinder database [180]. Virulence genes were identified with ABRicate using the virulence factor database (VFDB; [239]), and statistical differences between metadata with only two groups were tested for with a Mann-Whitney U test [240]. Multiple comparisons of means via Tukey contrasts from R

package multcomp [241] was performed on metadata of more than two groups and multiple comparisons was corrected with Holm-Bonferroni method [242]. Longitudinal change to phylogroup was visualised with R package circlize [243].

The two large clusters of MAGs, positioned near ST10 and phylogroup B1, were compared to additional reference STs in order to distinguish their identity at a higher resolution.

For each ST within Enterobase that was annotated to ST complex 10 and had at least 20 available sequences were selected. 20 genomes from each (40 for ST10) were downloaded, which consisted of: ST4, ST10, ST34, ST43, ST44, ST48, ST167, ST215, ST218, ST227, ST378, ST617, ST656, ST685, ST716, ST744, ST746, ST752, ST757, ST761, ST772, ST993, ST1060, ST1091, ST1141, ST1201, ST1284, ST1286, ST1303, ST1312, ST1415, ST1434, ST1491, ST1585, ST1638, ST1721, ST2223, ST2353, ST2491, ST2496, ST2690, ST3107, ST3476, ST3489, ST3877 and ST4981.

The 92 MAGs clustered within phylogroup A were analysed with the reference genomes as before.

For each ST within Enterobase that had been annotated to phylogroup B1 and had at least 20 available sequences were selected. 20 genomes from each ST were downloaded, which consisted of: ST3, ST16, ST17, ST20, ST21, ST23, ST25, ST26, ST29, ST40, ST56, ST88, ST99, ST101, ST109, ST120, ST129, ST145, ST148, ST152, ST155, ST187, ST192, ST200, ST205, ST223, ST243, ST252, ST270, ST311, ST332, ST388, ST392, ST441, ST442, ST539, ST677, ST678, ST811, ST829, ST1086, ST1125, ST1147, ST1148, ST1172, ST1246, ST1252, ST1273, ST1304, ST1326, ST1502, ST1611, ST1623, ST1730, ST1792, ST1992, ST2079, ST2165, ST2328, ST2354, ST2836, ST4213, ST4748, ST5418, ST5475, ST5559, ST5891, ST6069, ST6096, ST7375 and ST8405. The 36 MAGs clustered within phylogroup B1 were analysed with the reference genomes as before.

4.3 Results

179 travellers provided faecal swabs before travel and immediately upon return. These were whole metagenome sequenced, the reads processed into metagenome assembled genomes (MAGs) and all MAGs per sample that identified as *E. coli* were selected. 4430 MAGs were generated from 358 samples and after quality filtering, 213 were identified as *E. coli*.

4.3.1 Commensal *E. coli* differ from *E. coli* acquired during travel

The commensal, pre-travel population of *E. coli* was dominated by phylogroups B2 (33/90), A (24/90) and D (14/90). Whereas post-travel *E. coli* belonged mainly to phylogroups A (60/123) and B1 (29/123) (Figure 4.1).

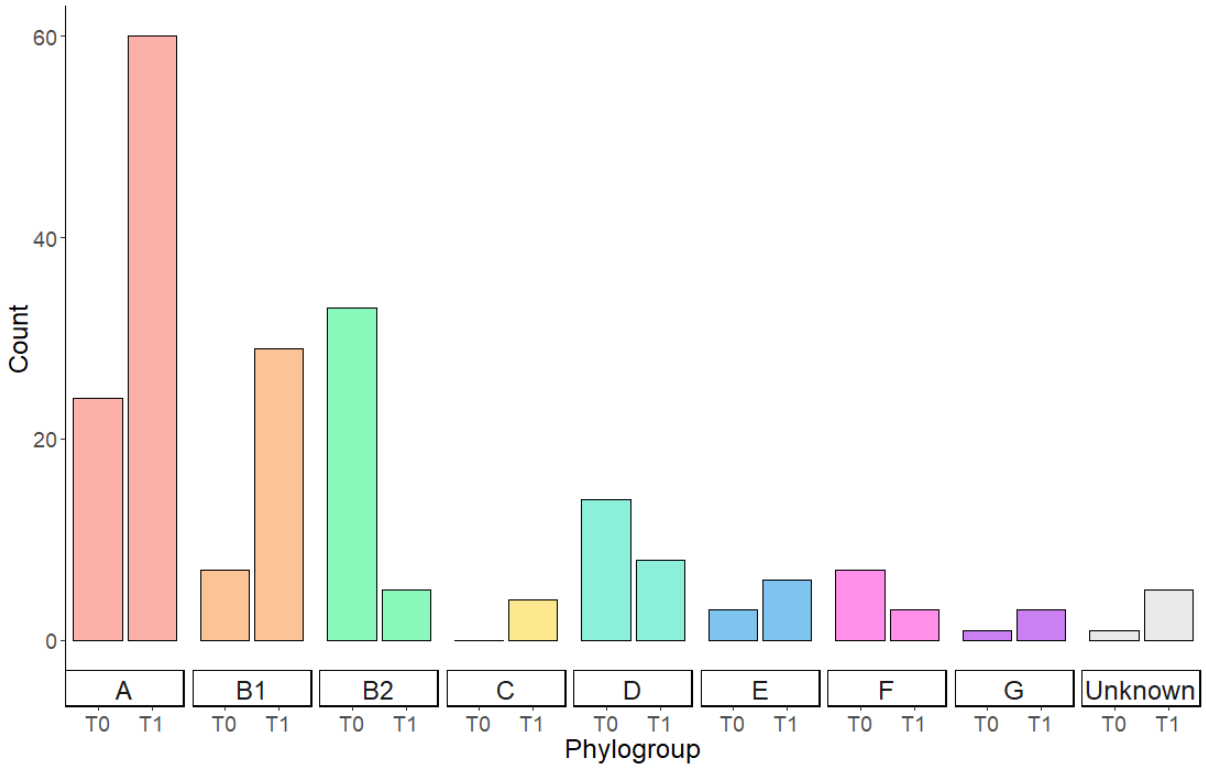

Fig. 4.1: *E. coli* MAG counts by phylogroup and time point.
E. coli MAGs from pre-travel (T0) and post-travel (T1) samples, separated into phylogroups as annotated by ClermontTyping. MAGs that did not annotate to a phylogroup are labelled as 'Unknown'.

4.3.2 There is a low prevalence of AMR in MAGs

The binning methods used in the production of the MAGs could not accurately cluster plasmid DNA with chromosomal DNA, so few AMR genes were detected. Most pre-travel MAGs contained *mdfA*, a multidrug transporter gene found ubiquitously in *E. coli*, but very few other AMR genes and, as expected, no $bla_{\text{CTX-M}}$ genes (Figure 4.2).

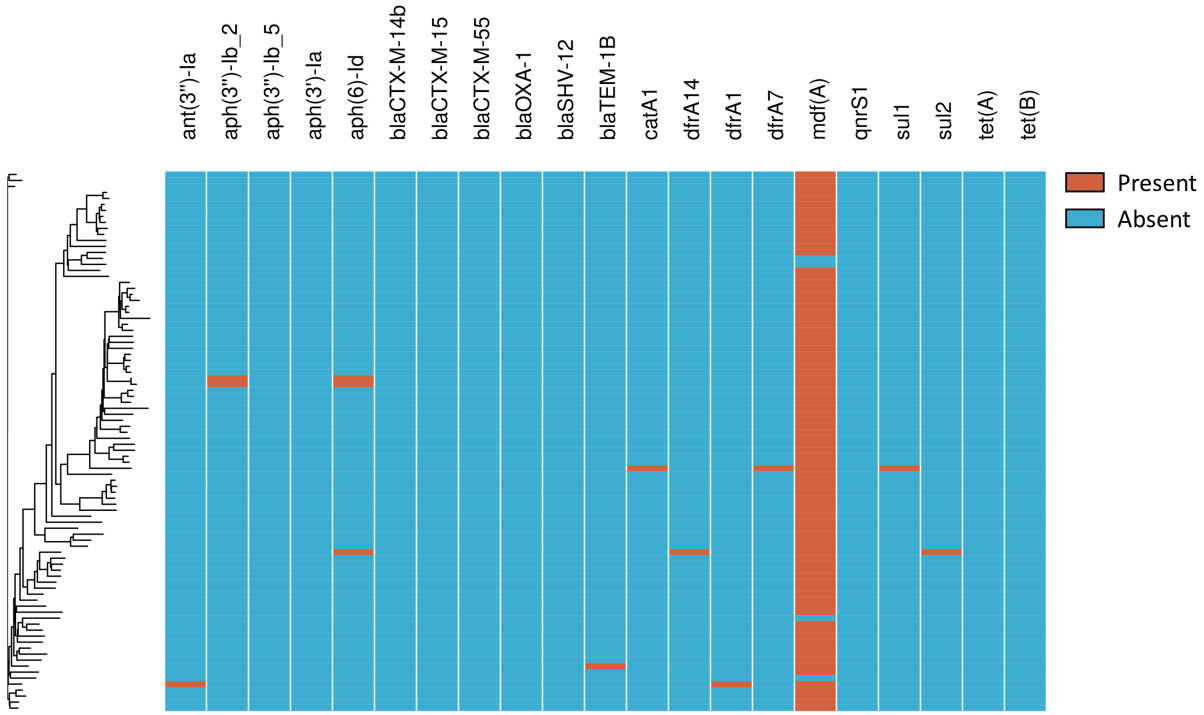

Fig. 4.2: Presence of AMR genes in pre-travel MAGs.
MAGs from pre-travel samples were screened for AMR genes from the resfinder database with ABRicate. Orange represents presence, blue represents absence. Samples have been clustered with mashtree and are represented as a dendrogram.

Post-travel MAGs also had few detected AMR genes, albeit more than pre-travel MAGs, and a ubiquitous *mdfA* gene (Figure 4.3). However, five MAGs contained $bla_{\text{CTX-M}}$ genes that are likely to be chromosomally inserted due to being successfully binned into the putative *E. coli* genomes.

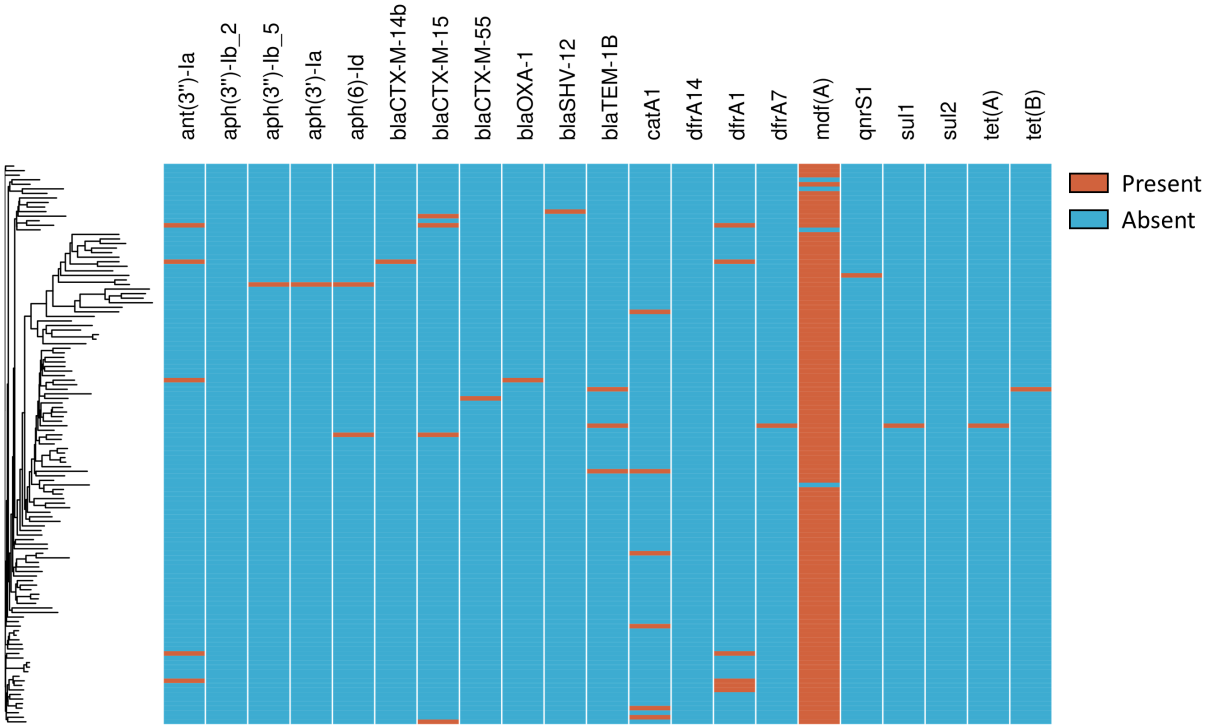

Fig. 4.3: Presence of AMR genes in post-travel MAGs.
MAGs from post-travel samples were screened for AMR genes from the resfinder database with ABRicate. Orange represents presence, blue represents absence. Samples have been clustered with mashtree and are represented as a dendrogram.

4.3.3 Phylogenetic relatedness of reference and MAG *E. coli*

The core genome between 213 MAGs and 420 reference genomes consisted of 1677 genes shared by >99% of genomes and an additional 1402 genes shared by 95% - 99% of genomes. The core genes were aligned to calculate the similarity distances between all samples and a corresponding phylogenetic tree was produced, with time point metadata added (Figure 4.4). Pre-travel MAGs are heavily skewed towards phylogroup B2 whereas post-travel MAGs are predominantly present in phylogroups A and B1.

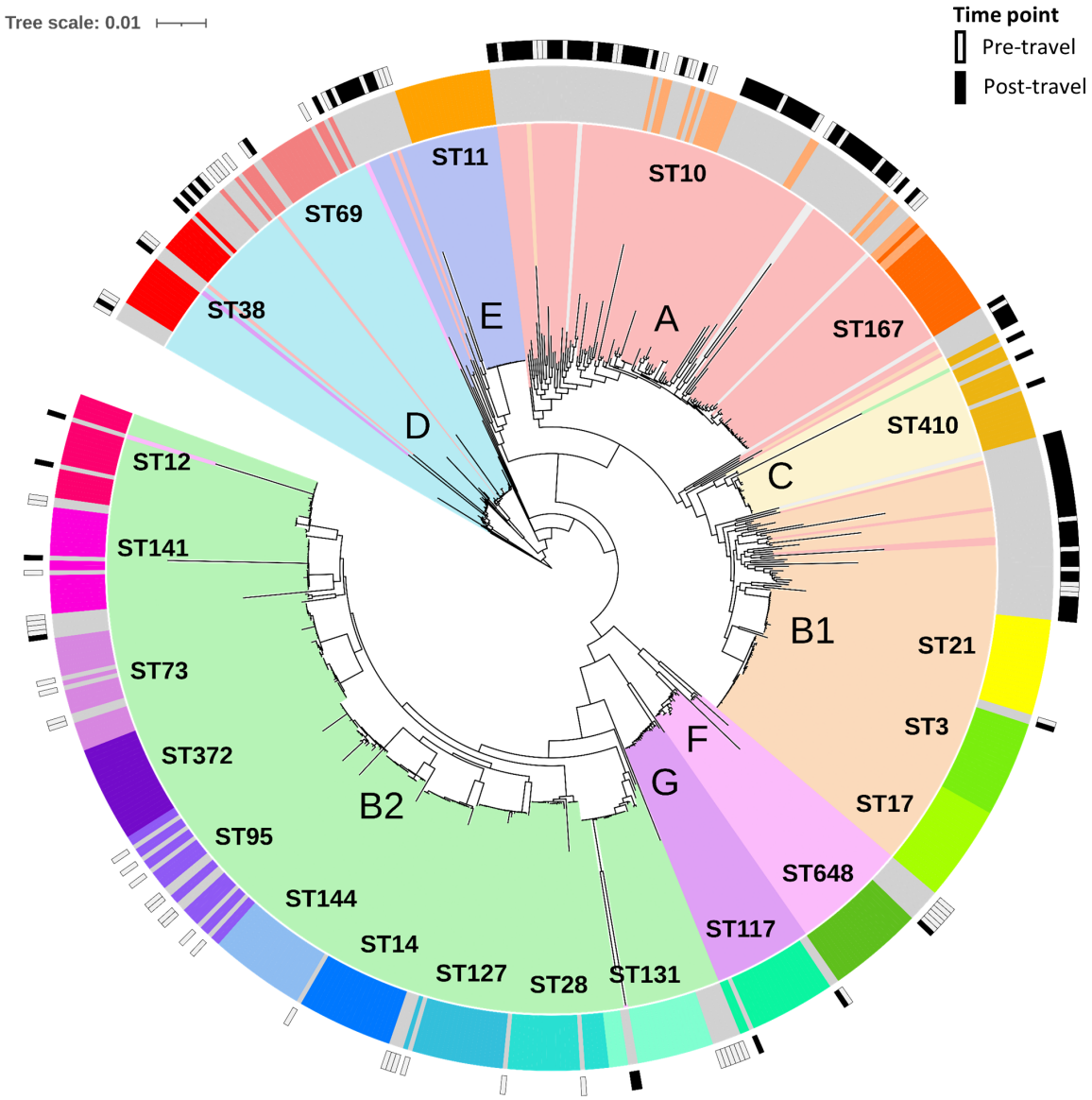

Fig. 4.4: *E. coli* MAG and reference *E. coli* phylogeny of travellers.
The core genes of 213 *E. coli* MAGs and 20 genomes each from 21 reference sequence types were aligned and phylogenetic inference was performed by IQ-TREE. Inner shading surrounding branches represents *E. coli* separation by phylogroup, as determined by ClermontTyping; grey is unknown. Innermost ring represents MAGs (grey) and 21 different reference sequence types (coloured). Empty black boxes represent pre-travel MAGs; filled black boxes post-travel MAGs. Tree was visualised with iTOL.

In order to associate additional factors to phylogroup, other than time point, an identical tree was created. Added to this was phenotypically and PCR confirmed carriage of ESBL-E in post-travel sample and the subregion travelled to. Factors that perturb the microbiome and were significant associated with an increase in ESBL-E acquisition [47] were also added: diarrhoea and consumption of food from street

vendors. There are comparable amounts of each metadata within each phylogroup and across ESBL-E acquisition, suggesting that there are no specific *E. coli* displacement associated with the countries travelled to (when grouped into subcontinents) or with the factors that affect the microbiome.

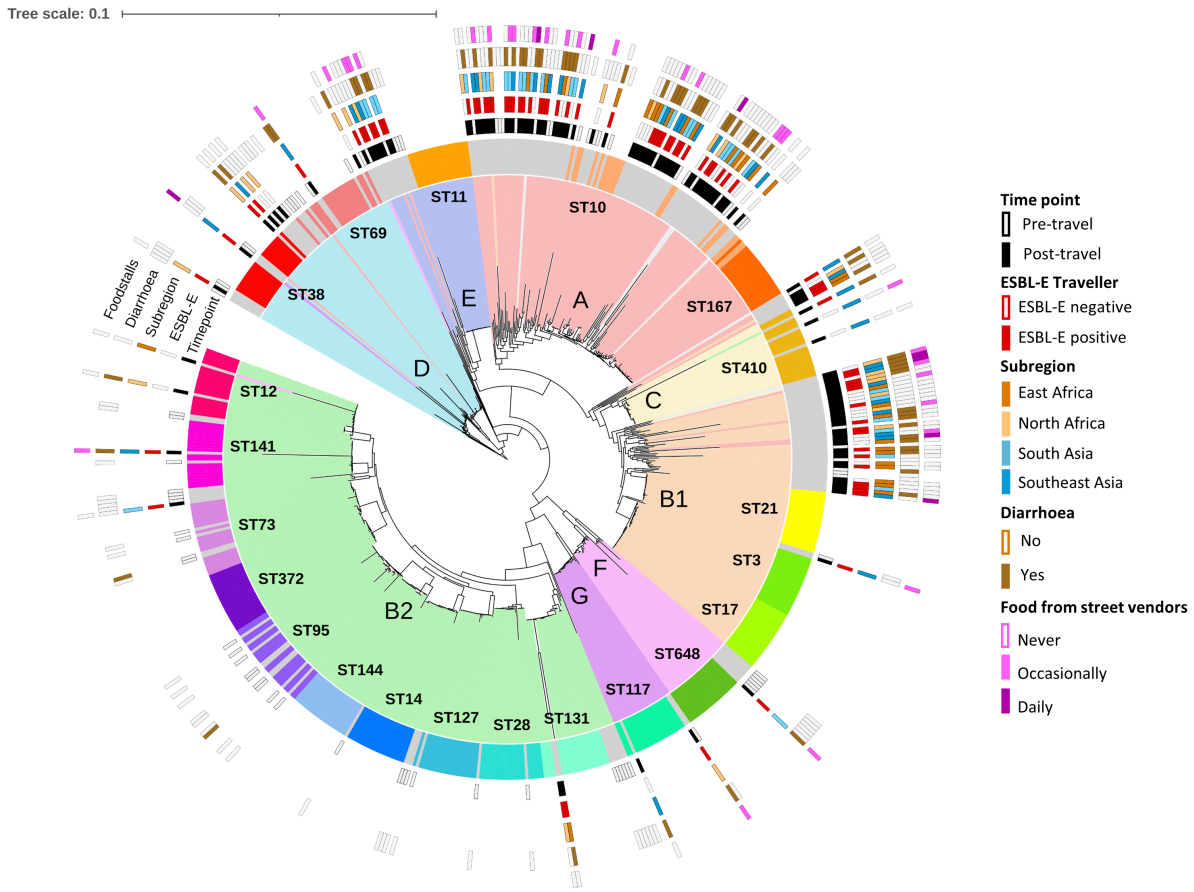

Fig. 4.5: *E. coli* MAG and reference *E. coli* phylogeny of travellers with various metadata.
The core genes of 213 *E. coli* MAGs and 20 genomes each from 21 reference sequence types were aligned and phylogenetic inference was performed by IQ-TREE. Inner shading surrounding branches represents *E. coli* separation by phylogroup, as determined by ClermontTyping; grey is unknown. Innermost ring represents MAGs (grey) and 21 different reference sequence types (coloured). Empty black boxes represent pre-travel MAGs; filled black boxes post-travel MAGs. Empty red boxes, only assigned to post-travel MAGs, highlight travellers who remained negative for ESBL-E during travel, whereas filled red boxes highlight travellers who acquired ESBL-E. Region travelled to is Northern Africa (light orange), Eastern Africa (dark orange), South Asia (light blue) and Southeast Asia (dark blue). Diarrhoea within 3 months before travel (pre-travel MAGs) or the onset of diarrhoea during travel (post-travel MAGs) is represented in brown, no diarrhoea is represented as empty box. Daily food consumption from street vendors is represented as purple, occasional represented as pink, and not at all as an empty box. Tree was visualised with iTOL.

The core genome between 92 MAGs that were identified as phylogroup A and clustered near ST10, as well as 940 reference genomes belonging to members of ST complex 10, consisted of 2453 genes shared by >99% of genomes and an additional 803 genes shared by 95% - 99% of genomes. The core genes were aligned in this subset of MAGs and additional reference *E. coli* in order to further distinguish the large cluster of MAGs around ST10 in Figure 4.4. A phylogenetic tree was produced (Figure 4.6) with IQ-TREE.

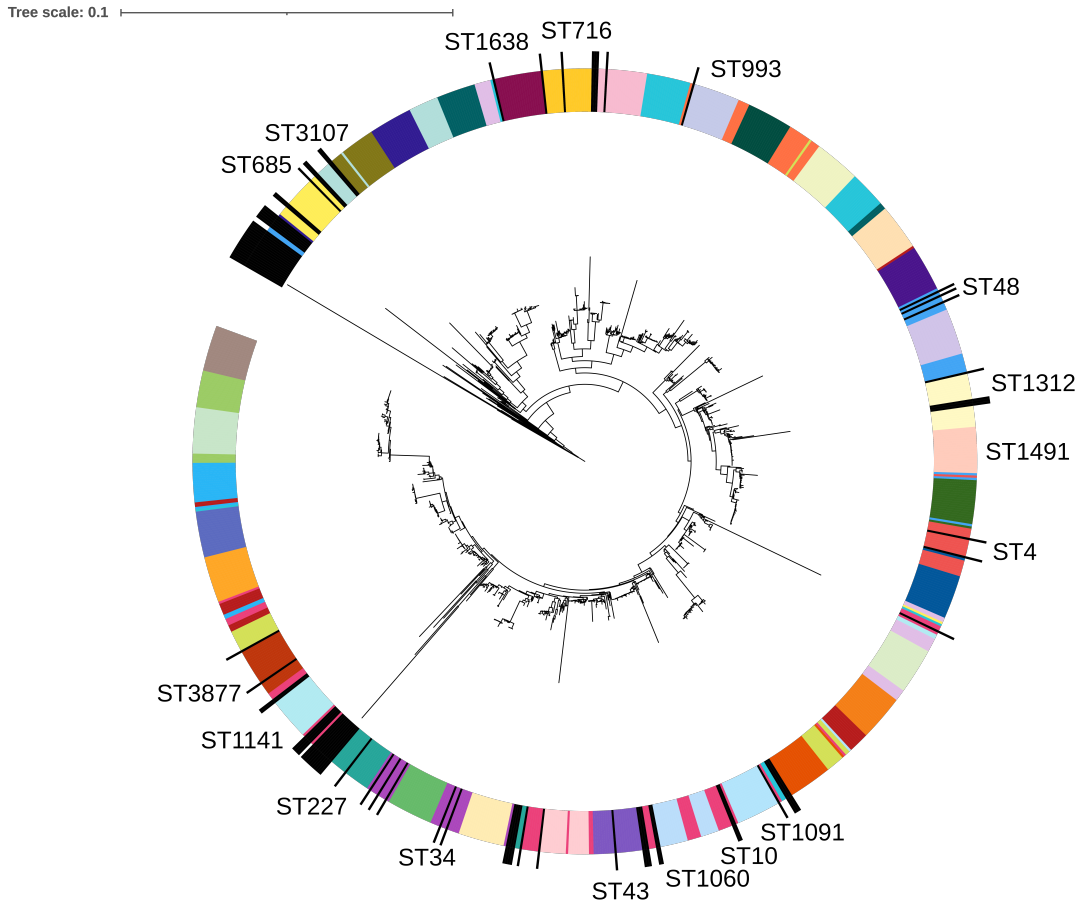

Fig. 4.6: Reference ST complex 10 *E. coli* and phylogroup A MAGs phylogeny of travellers.
The core genes of 92 MAGs identified as phylogroup A and clustering around ST10 were aligned with 940 reference sequences - 40 sequences of ST10 and 20 sequences from 45 other STs belonging to ST complex 10 - and phylogenetic inference was inferred by IQ-TREE. Coloured ring represents different STs and MAGs (black). Closest related STs to MAGs are labelled with text. Tree was visualised with iTOL and rooted at midpoint.

The core genome between 36 MAGs that were identified as phylogroup B1 and clustered with other members of that group, and 1420 reference *E. coli* from 71 STs within phylogroup B1 were compared. 2739 genes were shared between >99% of genomes and an additional 461 genes were shared between 95% and 99% of genomes. The core genes were aligned for all phylogroup B1 MAGs to further distinguish the closely clustered group of these MAGs in Figure 4.4, and a phylogenetic tree was created (Figure 4.7).

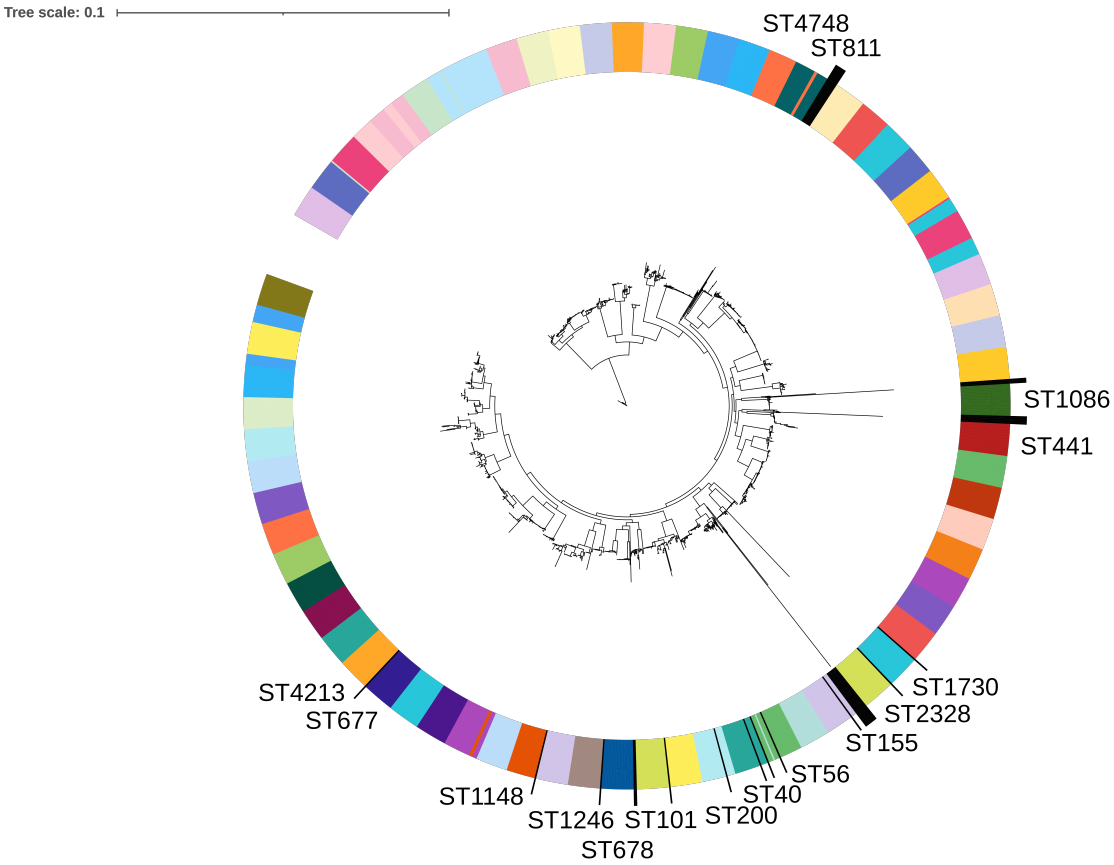

Fig. 4.7: Reference phylogroup B1 *E. coli* and phylogroup B1 MAGs phylogeny of travellers.
The core genes of 36 MAGs identified as phylogroup B1 and clustering with other members of phylogroup B1 were aligned with 1420 reference sequences - 20 sequences from 71 other STs belonging to phylogroup B1 - and phylogenetic inference was inferred by IQ-TREE. Coloured ring represents different STs and MAGs (black). Closest related STs to MAGs are labelled with text. Tree was visualised with iTOL.

For each MAG, the combined results from each set of core gene alignments is represented in Appendix Table A.2. The phylogroups and STs as defined by ClermonTyper and MLST (PubMLST database), respectively, are listed. The phylogroup and ST of the closest branch of reference *E. coli* is also listed, as determined from the (more specific) ST complex 10 and phylogroup B1 trees first then followed by Figure 4.4. Using this method to assign phylogroups had a mismatch in 18/213 (8.5 %) of cases, but was able to assign an additional 49 STs in the 90 MAGs that were not assigned an ST with MLST. The alleles of the seven MLST genes in each MAG are listed in Appendix Table A.3.

4.3.4 Virulence potential is lower in post-travel MAGs

MAGs were screened for virulence genes using ABRicate and the Virulence Factor Database (VFDB). When counted, post-travel MAGs have significantly fewer virulence genes than pre-travel MAGs (p = <0.001; Figure 4.8).

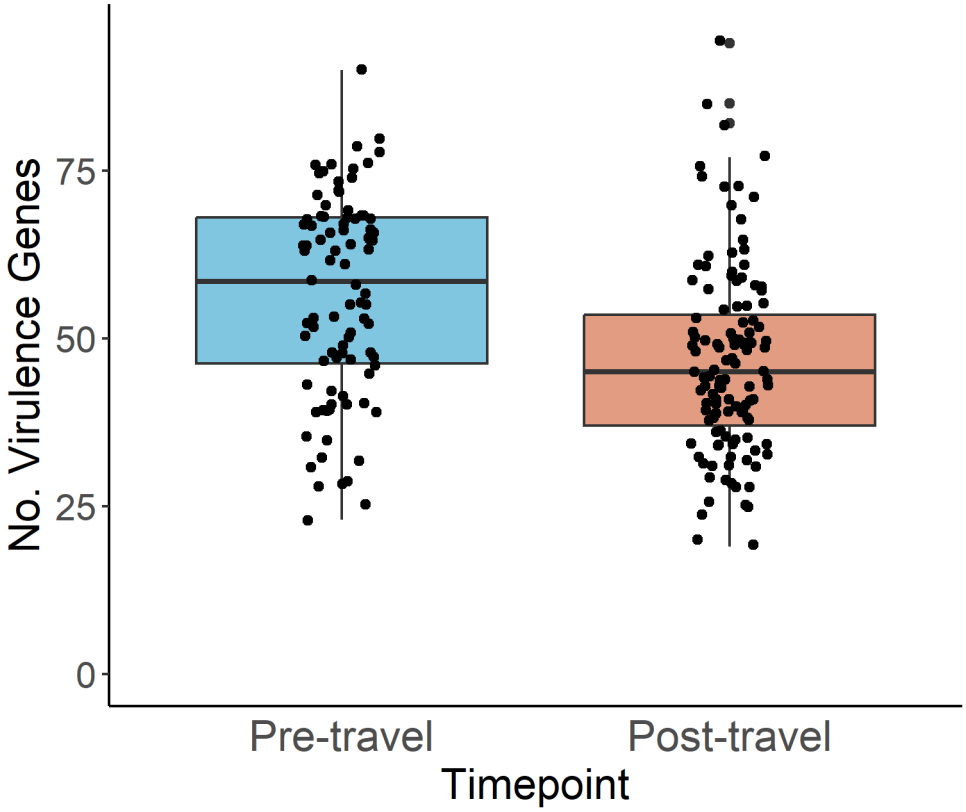

Fig. 4.8: Number of virulence genes screened for in *E. coli* MAGs, as separated by time point.
MAGs were screened against VFDB by ABRicate and the numbers of genes detected were counted. Data was separated by time point and 90 pre-travel MAGs are represented as blue, 123 post-travel MAGs as red. Significant p-value of <0.05 is indicated with *.

This is likely explained by the number of virulence genes across the phylogroups. Phylogroups B2 and D have the highest number of virulence genes overall, which is statistically higher than phylogroups A and B1, (Figure 4.9). As Figure 4.1 shows, B2 and D are most prevalent in pre-travel MAGs.

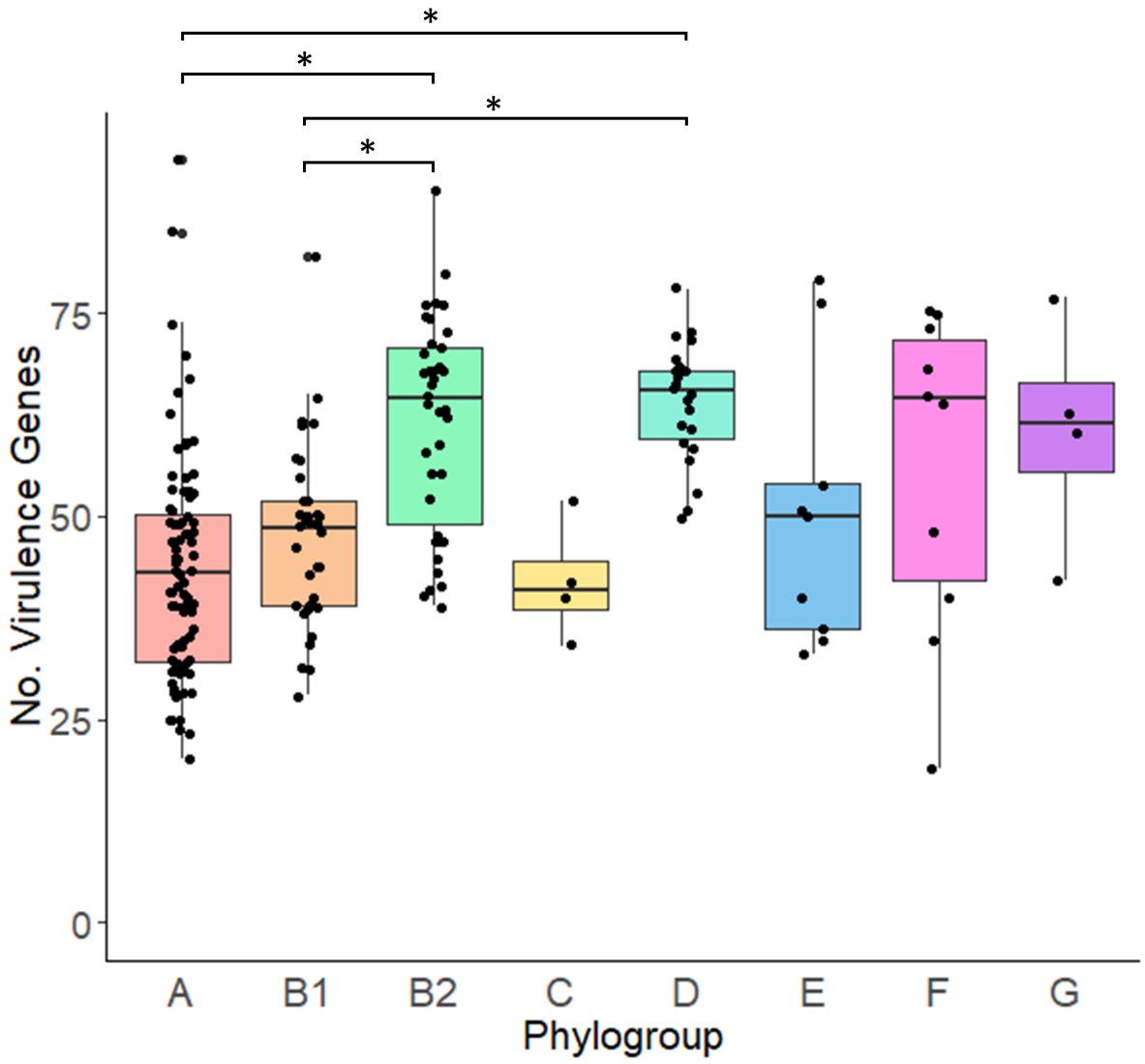

Fig. 4.9: Number of virulence genes screened for in *E. coli* MAGs, as separated by phylogroup.
MAGs were screened against VFDB by ABRicate and the numbers of genes detected were counted. Data was separated into phylogroups: 84 MAGs as phylogroup A (red); 36 as B1 (orange); 38 as B2 (green); 4 as C (yellow); 22 as D (turquoise); 9 as E (blue); 10 as F (pink) and 4 as G (purple). Significant differences ($p < 0.05$) are represented as *.

To check for other confounding factors that could lead to acquiring different amounts of virulence genes, further comparisons were made. The region travelled to does not determine the number of virulence genes in the *E. coli* acquired as they do not significantly differ between groups (Figure 4.10).

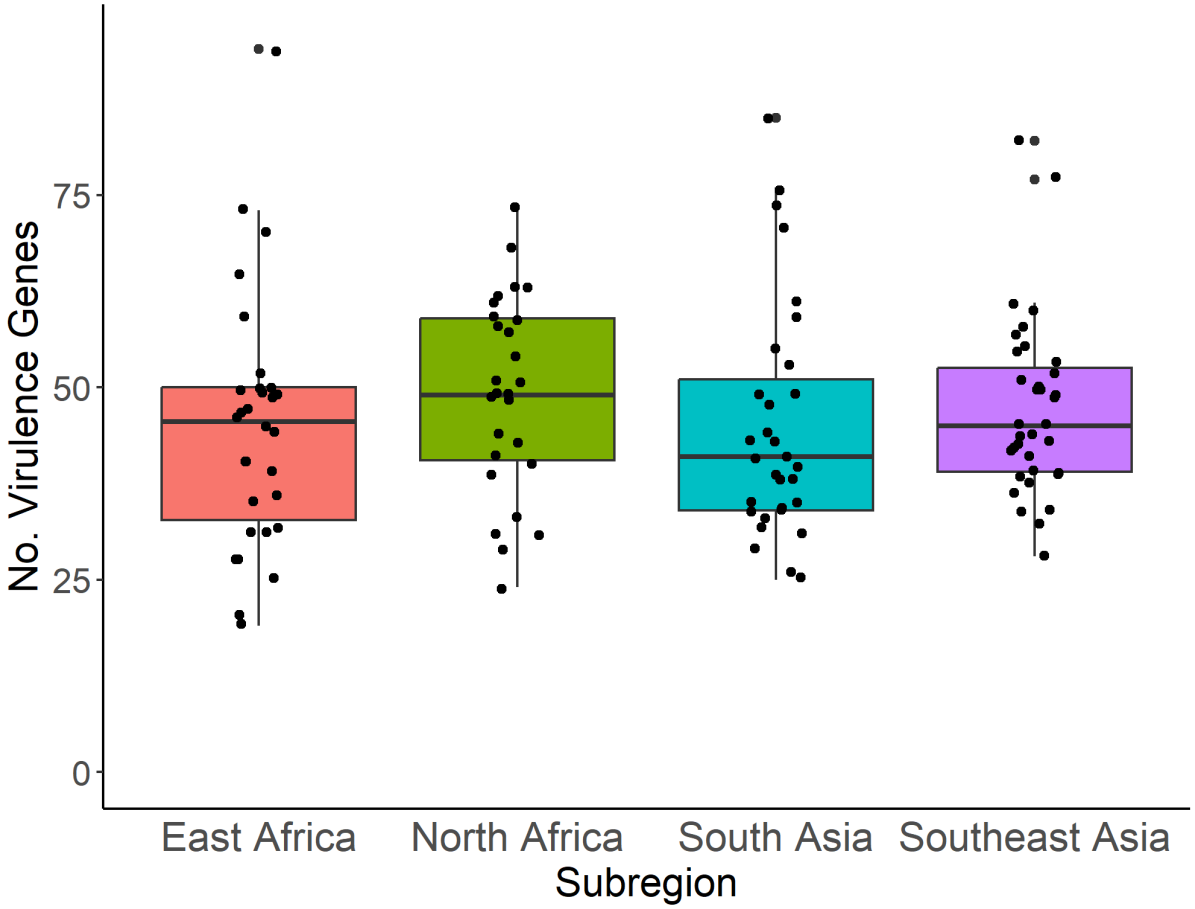

Fig. 4.10: Number of virulence genes screened for in *E. coli* MAGs, as separated by travel destination.
Post-travel MAGs were screened against VFDB by ABRicate and the numbers of genes detected were counted. Data was separated by destination of travel: 30 MAGs from East Africa (blue); 27 from North Africa (red); 31 from South Asia (purple) and 35 from Southeast Asia (green).

Additionally, the post-travel *E. coli* acquired in those with travellers' diarrhoea do not have a significantly different number of virulence genes to those that do not develop diarrhoea (Figure 4.11).

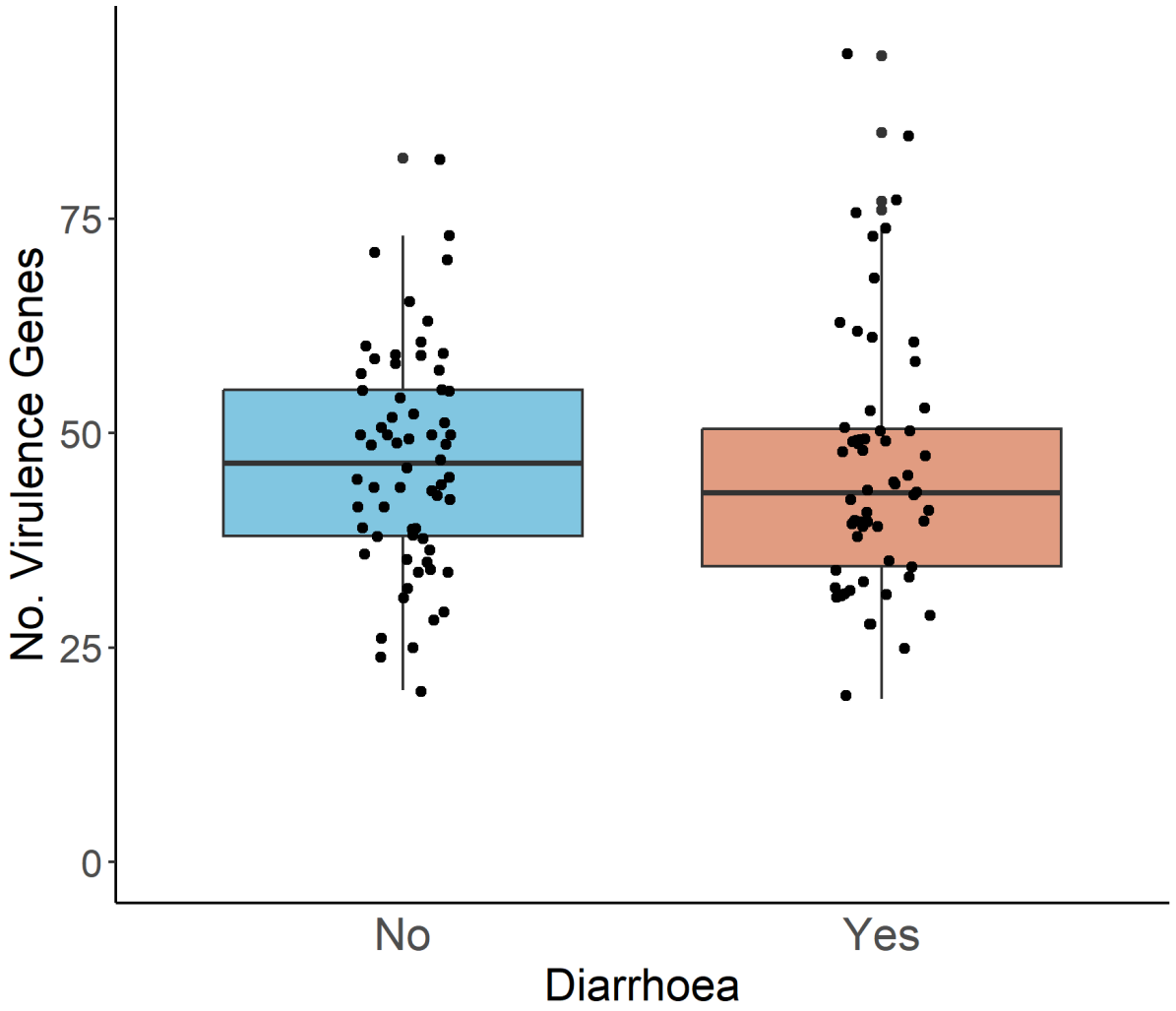

Fig. 4.11: Number of virulence genes screened for in *E. coli* MAGs, as separated by onset of travellers' diarrhoea.
Post-travel MAGs were screened against VFDB by ABRicate and the numbers of genes detected were counted. Data was separated by onset of diarrhoea during travel (Yes/No), and 64 travellers who did not develop diarrhoea are represented in blue whereas 59 travellers who developed travellers' diarrhoea are represented in red.

And finally, the *E. coli* in people colonised with ESBL-E do not differ in number of virulence genes compared to those remaining uncolonised by ESBL-E (Figure 4.12).

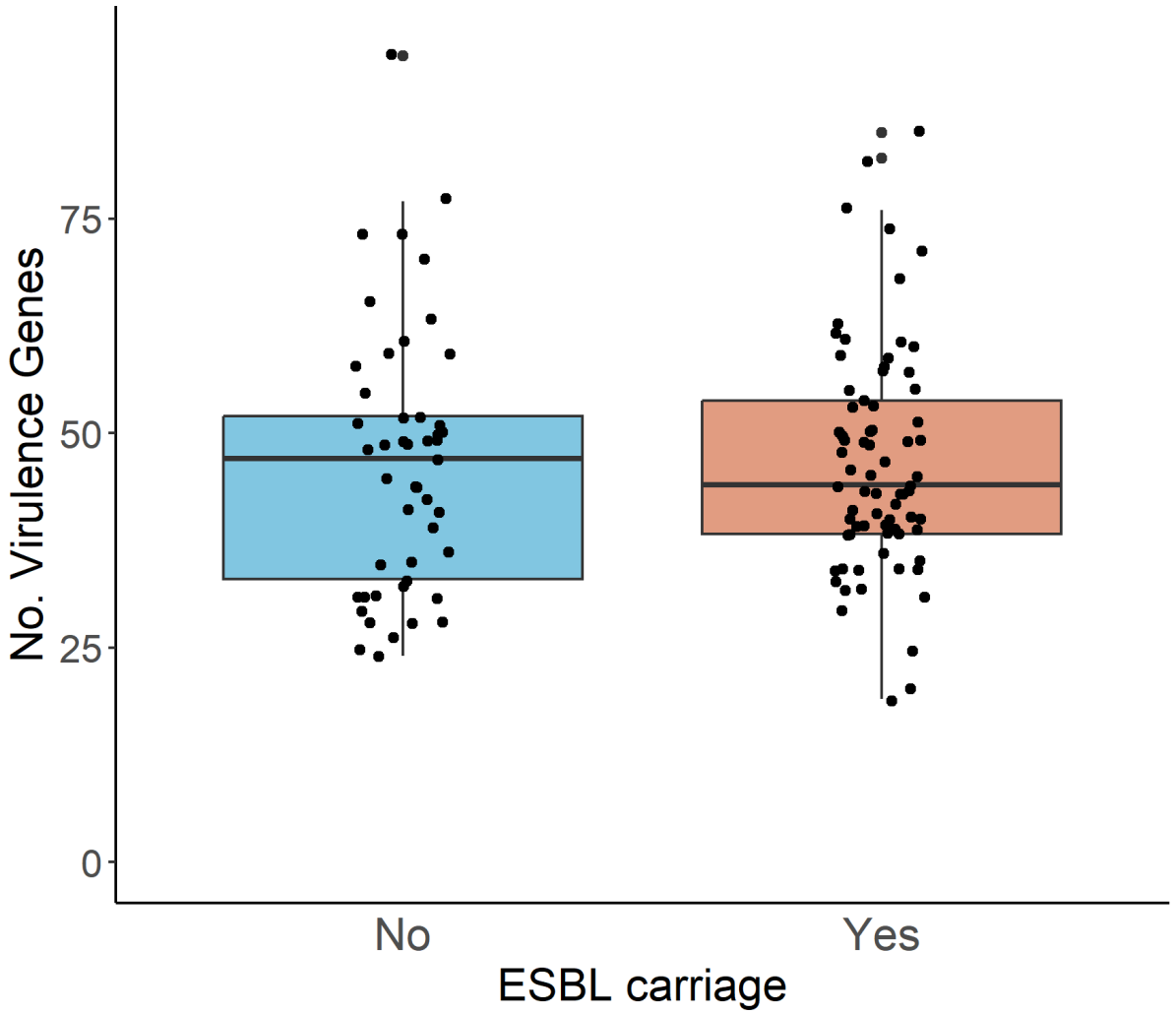

Fig. 4.12: Number of virulence genes screened for in *E. coli* MAGs, as separated by carriage of ESBL-E.
Post-travel MAGs were screened against VFDB by ABRicate and the numbers of genes detected were counted. Data was separated into 49 travellers who were not colonised with ESBL-E (blue) and 74 travellers who were colonised with ESBL-E (red).

4.3.5 Within traveller longitudinal change in *E. coli strain*

Next, we looked at travellers who had an *E. coli* MAG present in both their pre- and post-travel samples. Within these 66 travellers, the most acquired phylogroup was phylogroup A in 33 travellers of whom, before travel, 10 carried phylogroup A, 10 with B2 and 7 with D. Only one or two of the remainder of this group carried phylogroups B1, E, F and G. The next most acquired phylogroup was B1 in 13 travellers and their pre-travel phylogroups were spread between phylogroups A, B1, B2, D and F. No

traveller was colonised during travel by phylogroup B2, except for two travellers already carrying a phylogroup B2 MAG pre-travel (Figure 4.13).

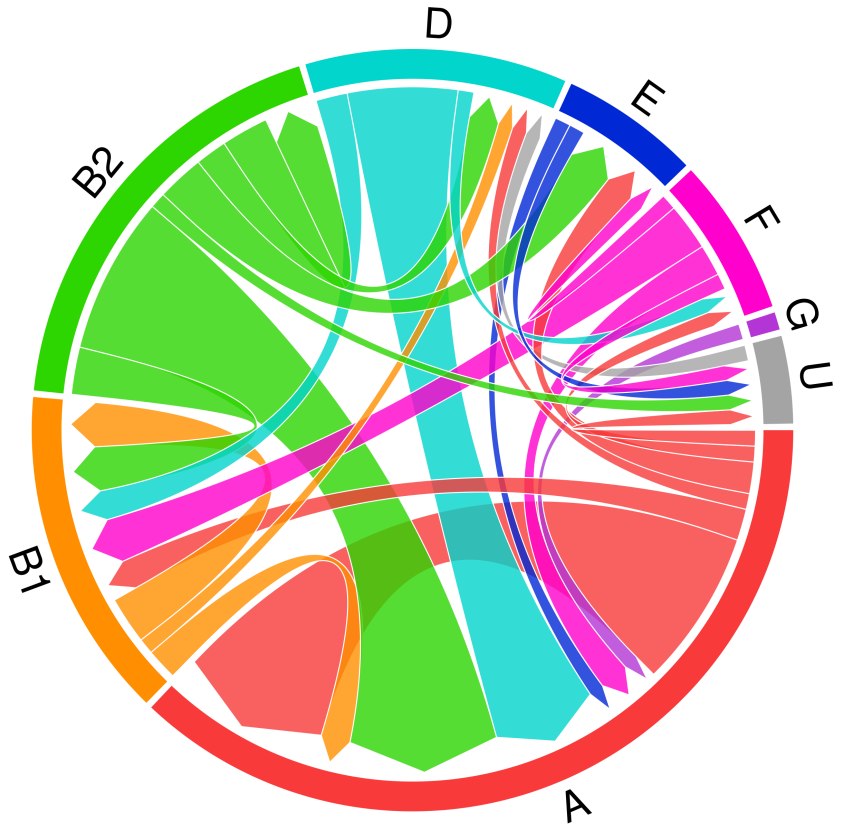

Fig. 4.13: Change in *E. coli* MAG phylogroup during travel.
The longitudinal change in phylogroup of the 66 travellers that contained both a pre- and post-travel *E. coli* MAG. Detected phylogroups were A, B1, B2, D, E, F and G; MAGs labelled 'Unknown' are represented as 'U'. Arrow thickness represents number of MAGs.

Figure 4.13 was separated into travellers who did or did not acquire ESBL-E, to visualise if there are colonisation specific patterns (Figure 4.14). Of the phylogroups present in more than 10% of travellers before travel, their prevalence are comparable between those who did and those who did not acquire ESBL-E - except for phylogroups A and F where 13/17 and 6/7, respectively, go on to acquire ESBL-E during travel. On the other hand, only phylogroups A and B1 were present in more than 10% of travellers post-travel, and these were comparable between those who did or did not acquire ESBL-E.

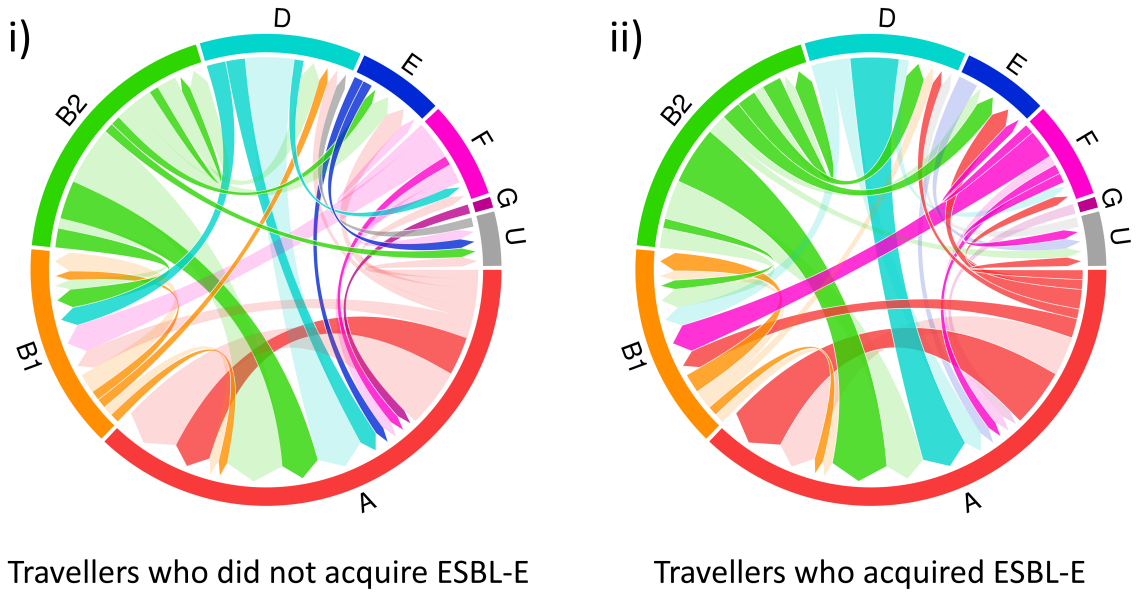

Fig. 4.14: Longitudinal change in *E. coli* MAG phylogroup with or without ESBL-E acquisition.
The longitudinal change in phylogroup of the 66 travellers that contained both pre- and post-travel *E. coli* MAGs, separated into travellers who i) did acquire ESBL-E or ii) did not acquire ESBL-E during travel. Detected phylogroups were A, B1, B2, D, E, F and G; MAGs labelled 'Unknown' are represented as 'U'. Arrow thickness represents number of MAGs.

4.4 Discussion

The travel-associated effects to *E. coli* found in the gut are important to understand as people may travel to countries with drastically different types of *E. coli*. The levels of virulence genes or antibiotic resistance genes may differ so tracking the spread of *E. coli* needs to be monitored. The aim of this study was to classify the commensal *E. coli* population and how this changes during travel, and if this is associated with the acquisition of ESBL-E. The pre-travel commensal population was dominated by phylogroups B2 and A, which was clearly altered during travel where there is a shift to group B1 and an increase in group A. Neither the destination of travel, nor the acquisition of ESBL-E, were associated with this shift.

There is a large sampling bias in the literature and in genomic datasets towards *E. coli* from clinical infections and fewer studies on commensal populations within developed countries. Commensal *E. coli* can be opportunistic pathogens and can act as

a reservoir of AMR genes, so there is definitely a need to focus on this population of bacteria. Also, the scope of the current literature around travel-associated changes to *E. coli* in the gut is limited as there are no large-scale studies that investigate both pre- and post-travel *E. coli*. The use of longitudinal WMGS samples from 179 travellers in this study therefore largely contributes to the field of research.

In this study, phylogroups B2 and A made up most of the commensal *E. coli* population, with phylogroup D being the next most common. These are known to be globally disseminated and commonly detected in commensal populations [218, 244]. The samples from our study were collected from 2012-2013 and the proportions of each phylogroup are consistent with the few studies on commensal *E. coli* from the same time [223]. However, an increasing prevalence of phylogroup B2 may become a problem as the virulence-associated genes carried by its members are largely associated with ExPEC infections [101, 245]. Here, we see higher amounts of virulence-associated genes in phylogroups B2 and D than other phylogroups. However, the virulence factor database is bias towards calling genes associated with long-term colonisation of the gut as virulence genes, so these phylogroups may negatively be considered virulent yet just more adapted to living in the gut. Regardless, the two most common phylogroup B2 STs in this study were ST73 and ST95. Meta-analyses of these STs found higher proportions in samples (recovered from both stool and sites of infection) in Europe and North America when compared to Africa and Asia [245]. It was suggested that there is an overrepresentation due to database bias and isolate selection, yet this study suggests that there is simply a regional difference. The results here are highlighting that there may be a high potential for extraintestinal infections in a healthy population from a developed country.

On the other hand, international travel leads to the replacement of dominant *E. coli* in the gut with phylogroups A and B1, and few instances of phylogroups D and E. This seems to occur independently of whether ESBL-E was acquired during travel, or if the gut microbiome was perturbed. The loss of phylogroup B2 could be beneficial when replaced by phylogroup B1, which is largely considered as commensal [144], especially as there is a reduction of virulence genes after travel. However, it is more complex with other phylogroups, as phylogroup D is associated with infection and high levels of virulence factors [246, 247] and phylogroup E has a structure that mimics the entire *E. coli* population structure including varying rates of infectious strains and AMR [248]. In this study, travellers are acquiring new strains of *E. coli* and experience large scale displacement of the commensal population. This is consistent with real-time sampling of

travellers that showed an acquisition of new strains replacing the initially acquired strains [155] and occurs without drastically impacting the gut microbiome. As phylogroup A is commonly the cause of diarrhoeal diseases, this may explain why they are the most commonly acquired lineages during travel, and non-diarrhoeal strains are largely due to their global prevalence.

A major strength of this study is the longitudinal nature of the dataset as the within-host change to the *E. coli* population can be determined. The large sample size can correct for confounding factors and as the sample collection was the same for all travellers, comparisons between four different destinations of travel can be made. The binning method employed in this study are likely only detecting the most abundant *E. coli* per sample and is losing plasmid DNA information. We therefore cannot claim that the MAGs and the ESBL-producing *E. coli* are the same, but within-host *E. coli* diversity is usually low and dominated by a single strain [144], so this is likely captured as a MAG.

International travel could be considered as a factor causing a shift to the *E. coli* population, but to understand how permanent the changes are would require more longitudinal samples. Most ESBL-producing *E. coli* acquired during travel are lost one month after return [47], so additional samples over time would confirm if this is a loss of AMR genes or a change to the *E. coli* population. Whole genome sequencing of *E. coli* isolates would strengthen the conclusions of each ST and aid with confirming which *E. coli* are producing ESBLs.

This study shows how international travel causes a drastic shift to the within-host *E. coli* population. It is epidemiologically important to further understand commensal *E. coli* and how they link to infection causing *E. coli*, and this study contributes to this by highlighting the role of international travel.

Chapter Five

General Discussion

The increase in levels of AMR bacteria, infections and mortality across the globe is concerning. The risk of an AMR infection poses a danger for one's health, especially in immunocompromised people or those undergoing surgery [137, 138]. An increase in globalisation and international travel is driving the spread of AMR between countries [136, 149], with the gut microbiome acting as a reservoir through which travellers are acquiring and importing AMR bacteria [151]. Perturbations to the gut microbiome are some of the factors most associated with this acquisition [47], but the extent to which the microbiota is involved is unclear. However, if a change to the gut microbiome is associated with an increase in ESBL-E acquisition, then this suggests that the microbiome plays a vital role in creating an effective colonisation resistance to ESBL-E. In this thesis, we explored whether the gut microbiome and individual taxa help prevent the acquisition of ESBL-E.

5.1 How ESBL-E interacts with the gut microbiome as a whole

As the gut microbiome acts as a reservoir for AMR bacteria, the potential for onwards transmission of AMR genes to commensal organisms, opportunistic pathogens or other hosts highlights a concerning problem. In **Chapter 2** we showed how the structure of the microbiome – its bacterial richness and d"iversity – was not correlated to the acquisition of ESBL-E, as it does not significantly differ between travellers acquiring ESBL-E and those remaining uncolonised. It did also not differ amongst the pre-travel microbiomes and travellers who did not go on to acquire ESBL-E did not cluster

separately from those who did, meaning that there is no common microbiome structure that can preclude the invasion of ESBL-E.

Previous studies on the gut microbiome of travellers also found that the microbiome is not correlated with the acquisition of MDR *Enterobacteriaceae* [249, 139, 140], meaning that our results are in concordance with the current literature. However, these studies are limited either by a small sample size [140], or by using 16S rRNA gene sequencing [249] or meta-transcriptomics [139] which have a lack of resolution to species level. These were the main reasons why our WMGS approach on a large data set from longitudinal samples was required. In **Chapter 2**, we also analysed the functional potential of the microbiome and calculated the abundance of genes involved in SCFA production, which are a proxy of good intestinal health [67]. These results were consistent with the microbiome structure analysis as we did not find any groups of metabolic pathways associated with the acquisition of ESBL-E. As we found that ESBL-E can invade the microbiome without causing a significant shift to its structure, it is not unlikely that the metabolic profile of the microbiome also remains stable within our detection limit. The functional composition has been shown to remain relatively stable, even reducing the efficacy of donor microbiome engraftment during faecal microbiome transplantation in recipients with a higher functional redundancy in their gut microbiome [58]. A limitation of the analyses in **Chapter 2** is that we calculated gene abundance, which represents metabolic *potential*; so a follow-up study of transcriptomics or metabolomics analysis could be carried out to determine if there is a higher functional redundancy or specific protective metabolic pathways in travellers who did not acquire ESBL-E than in those who did. A strength of our study, however, was the ability to separate the statistical analysis by geographical region and maintain a sufficiently large sample size, so that the results are applicable per region. Rates of acquisition of ESBL-E can vary quite drastically by country [47], so it was important to analyse at both a global and a regional scale to see if the conclusions are consistent, or if there were regional variations that lead to a different interaction between the gut microbiome and invading ESBL-E, which there were not. Even though travellers will interact with different factors and perturbations that vary by country, the gut microbiome seemingly responds in a similar manner within the scope of our study.

Consistently throughout **Chapter 2**, the onset of travellers' diarrhoea was significantly associated with the acquisition of ESBL-E. Others have already determined that diarrhoea and invasion of MDR bacteria are closely linked [47], for example a reduced species richness in the inflamed gut affecting colonisation resistance [89], or

because *Enterobacteriaceae* are adapted to colonising and thriving in inflammation [94, 95]. Of the traveller microbiome studies, one study found that although the microbiome was not associated with the acquisition of MDR *Enterobacteriaceae*, the pre-travel microbiota did predispose to the onset of travellers' diarrhoea [139]. The analyses in **Chapter 2** could be repeated on pre-travel samples but to instead determine if the onset of travellers' diarrhoea could be predicted. As the onset of travellers' diarrhoea drives an increased acquisition of ESBL-E, better understanding what drives the onset of diarrhoea could then, by proxy, be used to understand the acquisition of ESBL-E. As discussed in **Chapter 2**, ESBL-E is still able to colonise in those without travellers' diarrhoea, but not at the same rate in those with diarrhoea, meaning this is viable future research.

5.2 How ESBL-E interacts with individual species

Given that the microbiome as a whole does not differ between people who do and do not acquire ESBL-E during travel, the question is if individual species or strains have important interactions with ESBL-E. Commensal species are known to contribute to a functioning colonisation resistance to invading pathogens [183, 188, 189], but pathogenic bacteria may utilise their virulence mechanisms to overcome or bypass host defences and colonise the environment [78]. Discovering if there are specific species or a collection of species that ESBL-E are not able to outcompete with could be exploited in the development of probiotics for travellers.

In **Chapter 2** we investigated the taxa that, longitudinally, were significantly associated with the acquisition of ESBL-E. The results show that members of *Bacteroides* and *Citrobacter* had an increase in prevalence and/or abundance in travellers who remain negative for ESBL-E colonisation, suggesting that these species may be protective against ESBL-E. *Bacteroides* species are known to generally be metabolically beneficial for the intestinal health [187, 188, 189], and in a study comparing the gut microbiome of ESBL-E carriers vs non-carriers in Thailand, *Bacteroides uniformis* was found to be positively associated with ESBL-E non-carriers [186], supporting our conclusion that it may be protective. However, an intestinal inflammatory host response, as inducible by ESBL-E, can decrease abundance of *Bacteroides* [94]. Therefore, *Bacteroides* may appear protective in those not colonised with ESBL-E and not experiencing a host immune response, simply as its abundance in

those invaded by ESBL-E was decreased. We show that for *B. cellulosilyticus* and *B. ovatus* there is both an increase in abundance in travellers remaining negative for ESBL-E as well as a decrease in abundance in those who acquired ESBL-E. Together, these suggest that *Bacteroides* may offer protection against ESBL-E but may also have an abundance decrease *in response to* ESBL-E colonisation. This does highlight a limitation of our study, however, as the faecal samples were only collected before travel and immediately upon return, so the chronology of species interactions during travel cannot be determined. To ensure strong conclusions, additional sampling during travel should be taken.

Microcins are known to have strong antibacterial properties and are produced by *Enterobacteriaceae* to be used against other members of the family [141, 83]. It was especially interesting, therefore, that *Citrobacter*, a genus in the family *Enterobacteriaceae*, was also strongly associated with travellers remaining negative for ESBL-E. In **Chapter 3**, we tested the ability of *C. amalonaticus* – the *Citrobacter* species most frequently acquired during international travel to Laos [155] – to compete with commensal and ESBL-producing *E. coli in vitro*. *C. amalonaticus* has been shown to outcompete *C. rodentium*, a model bacterium used in mice to closely represent human enteropathogenic *E. coli* infections in humans, in a mouse infection experiment [143], and is therefore a promising organism to research. Although in **Chapter 2** the *Citrobacter* species of interest was identified as *C. freundii*, the taxonomic identification was limited as *C. freundii* is more present in reference databases as it is more frequently the cause of clinical infections. Our method of taxonomic identification was therefore bias by more present species, so equal numbers of reference genomes should have instead been used to ensure a fairer approach.

We found that *C. amalonaticus* was not able to outcompete either antimicrobial-sensitive or ESBL-producing *E. coli* in nutrient broth *in vitro*. A follow up experiment of growing *E. coli* in the presence of supernatant from *C. amalonaticus* (and other *Citrobacter*) could be performed to confirm the absence of microcins, as this would not be confounded with the effects that *E. coli* may have on the growth of *C. amalonaticus*. There are many metabolic process and host-bacterial interactions that occur *in vivo*, for example the production of immunomodulatory molecules like SCFAs [67], tryptophan [71] or secondary bile acids [69], that were not considered in the *in vitro* experiment. Considering these points, we conclude that *C. amalonaticus* does not directly inhibit the growth of *E. coli* through the use of targeted molecules like microcins, but as there were many metabolic process left unexplored there is potential

for further experiments. An interesting result to note, however, is that one of the *E. coli* isolates from phylogroup A was by far the least able to outcompete *C. amalonaticus*. Discussed further below, members of phylogroup A were the most frequently acquired *E. coli* strains during international travel. If *in vivo* the strains of *Citrobacter* are able to compete with *E. coli*, this may explain how *Citrobacter* is positively associated with travellers who remain uncolonised by ESBL-E in **Chapter 2** as the acquired strains of *E. coli* are the easiest to outcompete.

5.3 How ESBL-E interacts with commensal *E. coli*

It was interesting to note that in **Chapter 2**, we did not detect a significant change in relative abundance of *E. coli* during travel, regardless of ESBL-E acquisition, suggesting that invading *E. coli* have the ability to displace commensal *E. coli*. A study on the metagenomes from 22 independent cohorts of FMT recipients shows a large amount of variability in strain displacement, and that this occurs in a strain-specific manner [250]. Many species, typically the common commensals, frequently showed a donor-recipient strain coexistence, whereas other species showed either a donor or recipient dominance with less frequent coexistence. The outcomes of *E. coli* were not summarised due to a lack in statistical power, but this supports the idea of invading *E. coli* interacting with resident *E. coli*. Strain displacement is a promising avenue for decolonisation of ESBL-E as a preliminary study on the effect of FMT on ESBL-E found a reduction in its abundance and the total amount of ESBL genes [251]. Although the current literature has a strong focus on entire gut microbiome replacement, research into how individual species may interact is still warranted.

In **Chapter 4**, we describe the commensal population of *E. coli* and show how there is a clear temporal change to the detected strains. The commensal population consisted of predominantly phylogroups A and B2 which, even though phylogroup B2 is most frequently detected in extra-intestinal infections [101, 102], was consistent with the limited literature that focused on commensal populations of *E. coli* in wealthy countries [218, 223, 228]. However, we show that during travel there is a clear temporal shift to the *E. coli* phylogenetic distribution towards phylogroups A and B1. Studies on commensal *E. coli* in low- and middle-income countries (LMICs) are very limited, but a high proportion of phylogroups A and B1 have been found in both Mali and Benin in West Africa [229]. If other LMICs were to contain a similar proportion of phylogroups,

it could explain the shift we detected in the travellers. Our results indicate that the invading *E. coli* frequently displace commensal *E. coli* but not in a discernible way between MDR and non-MDR *E. coli*.

Invading *E. coli* employ virulence mechanisms to overcome resident bacteria and host defences in order to colonise [78]. In **Chapter 4**, we investigated whether invading ESBL-producing *E. coli* possessed more virulence genes that potentially contribute to their displacement of commensal *E. coli*. Interestingly, we found that post-travel *E. coli* had significantly fewer virulence genes than pre-travel *E. coli*, independently of ESBL-E acquisition. In a study tracking commensal *E. coli* in a Parisian population between 1980 and 2010, an increase in virulence potential was attributed to both an increase in STs of phylogroup B2, which typically contains more virulence genes [223], and in virulence gene frequency independently of ST frequency [228]. We determined that this significant reduction in virulence is attributed to the temporal loss of phylogroups B2 and D as these contained significantly more virulence genes than phylogroups A and B1. The interpretability is limited, however, by the virulence factor database containing bias towards annotating genes that are required for long term colonisation as virulence genes, which, when calculating gene count, are indistinguishable from genes actually involved in strain-strain competition. Only 10 travellers used antibiotics during travel, so there is no selecting pressure for *E. coli* harbouring AMR genes that could explain the strain displacement. As it is seemingly not dependent on total number of virulence genes, further analyses should be carried out to determine what traits invading ESBL-E possess that allow them to displace resident strains, for example pangenomic analysis to identify genes that cluster strongly with invading *E. coli* and not in the strains that were displaced.

5.4 Future perspectives

Presented in this thesis is an implementation of WMGS in the study of the gut microbiome and its role in the acquisition of AMR during international travel. Despite the foundation of this thesis being a study with one of the most comprehensive traveller gut microbiome data sets, there are still many unanswered questions surrounding microbiome data analysis.

With the increased accessibility to generating sequencing data, traveller study designs must better capture how dynamic a microbiome can be on a day to day basis. Sampling during travel ensures that we do not disregard transient microorganisms, especially as these still have the capability to transmit their AMR genetic material to the resident members of the gut microbiome. Additionally, colonisation resistance to invading pathogens may not always come in the form of total exclusion, but rather a speedy decolonisation triggered by the commensal microbiome, and without consistent sampling these relevant interactions may be missed. In this thesis, the acquisition of ESBL-E is strongly associated with the onset of travellers' diarrhoea, but further sampling during travel could clarify the order in which this occurs. Understanding how frequently ESBL-E elicits an immune response in the healthy intestinal tract, or simply takes advantage of the already inflamed environment could, for example, be used to alter guidance towards traveller behaviour or discover if some gut microbiomes are more resilient to diarrhoea and in turn protected against ESBL-E colonisation.

The future of metagenomics research is bright, especially with the introduction of new, accessible sequencing technologies like long-read sequencing. Long-read sequencing allows for an easier construction of genomes from a metagenomic sample, which can help increase the resolution to which microbiomes are typically analysed. Our findings in chapter 4 exemplify a high level of variation within a single species and that these strains correlate with travel, yet current metagenomics research, due to technical limitations, typically predict the quantity of the entire species. Combining short-read and long-read sequencing to produce high quality genomes of individual strains will shape the future of microbiome research by improving taxonomic identification and better recognising dynamics like strain replacement. A challenge for future research, however, is to keep up the momentum in software for accurate bioinformatics analysis. Statistically analysing microbiome data is notoriously complicated, so new tools must consider the inherent complexity.

Comprehensive microbiome research needs to expand further than metagenomics alone. The complexity of a microbiome is more than just the presence of microorganisms, which is why efforts to integrate metagenomics, transcriptomics and metabolomics are pivotal to the field. Standardising the processes in which these data are produced would harmonise research across institutes, allowing for comparability and meta-analyses. The sheer complexity of the gut microbiome means that discovering a species' role in an environment is not easily replicated in the laboratory, but a multi-omics approach may pinpoint genes of interest and metabolic pathways that lead

to the observed phenotype. The interplay between a single species and the entire gut microbiome means that discovering its role may, in practice, be difficult to replicate due to the exponential knock on effects, but a multi-omics approach would better understand its central role and allow researchers to correctly manipulate the gut microbiome to protect against invasion from pathogenic bacteria.

As the world moves to a more connected and interdependent place, the rates at which international travellers contribute to the spread of AMR will continually rise. We must urgently minimise the global transmission of AMR in order to ensure the clinical use of current antibiotics. The research in this thesis highlights a potential avenue in which to achieve this – colonisation resistance against the acquisition of ESBL-E. We have shown that this is not clearly achievable via the gut microbiome in its entirety, rather individual species that are similar to the invading species. As the technologies involved in microbiome research improve, new targets for interventions to alter one's susceptibility to invading pathogens may be discovered. However, due to the sheer complexity of the gut microbiome, the amount of sampling must be scaled up to incorporate how dynamic the host-bacteria and bacteria-bacteria interactions are on a day by day basis.

Appendix One

Appendix

Supplementary data from the published version [252] of **Chapter 2** is available at https://doi.org/10.1080/19490976.2022.2060676 and further directions are given below in Table A.1.

Table A.1: Location of Chapter 2 supplementary data

Feature statistics	**Supplementary table**
Species richness	1
Chao1	1
Effective Shannon	1
Bray Curtis distance	2
Bacteroides taxonomic classification	3

Table A.2: Description of *E. coli* MAG phylogenetic identification

Traveller ID	**Time Point**	**Phylogroup**	**Closest Branch Phylogroup**	**ST**	**Closest Branch ST**
R301633	Post-travel	A	A	4	-
R102767	Post-travel	-	A	10	-
R100887	Post-travel	A	A	10	-
R100889	Post-travel	A	A	10	-
R101285	Post-travel	A	A	10	-
R101362	Pre-travel	A	A	10	-
R102767	Pre-travel	A	A	10	-
R102852	Post-travel	A	A	10	-
R102946	Post-travel	A	A	10	-
R200595	Post-travel	A	A	10	-
R201153	Post-travel	A	A	10	-
R201960	Pre-travel	A	A	10	-
R202347	Pre-travel	A	A	10	-
R202347	Post-travel	A	A	10	-
R202459	Post-travel	A	A	10	-

R202785	Post-travel	A	A	10	-
R202835	Post-travel	A	A	10	-
R300500	Pre-travel	A	A	10	-
R302776	Post-travel	A	A	10	-
R200109	Post-travel	B1	B1	13	-
R102496	Post-travel	D	D	31	-
R100887	Pre-travel	A	A	34	-
R101018	Post-travel	A	A	34	-
R201517	Post-travel	A	A	34	-
R202056	Post-travel	A	A	34	-
R202702	Post-travel	A	A	34	-
R101601	Post-travel	A	A	43	-
R301773	Post-travel	A	A	43	-
R102045	Pre-travel	A	A	48	-
R102557	Post-travel	A	A	48	-
R202130	Pre-travel	A	A	48	-
R201211	Post-travel	A	A	50	-
R200595	Pre-travel	B1	B1	56	-
R102514	Pre-travel	B1	B1	58	-
R200384	Post-travel	B1	B1	58	-
R200651	Post-travel	B1	B1	58	-
R100956	Pre-travel	F	F	59	-
R200247	Pre-travel	F	F	59	-
R201198	Pre-travel	F	F	59	-
R302008	Pre-travel	F	F	59	-
R101601	Pre-travel	D	D	69	-
R200574	Pre-travel	D	D	69	-
R201348	Pre-travel	D	D	69	-
R202158	Pre-travel	D	D	69	-
R202517	Post-travel	D	D	69	-
R301773	Pre-travel	D	D	69	-
R102887	Pre-travel	B2	B2	73	-
R200323	Pre-travel	B2	B2	73	-
R202472	Pre-travel	B2	B2	73	-
R301633	Pre-travel	B2	B2	73	-
R102496	Pre-travel	B2	B2	80	-
R300442	Pre-travel	A	A	93	-
R200947	Pre-travel	B2	B2	95	-
R201915	Pre-travel	B2	B2	95	-
R502498	Pre-travel	B2	B2	95	-
R201299	Post-travel	G	G	117	-
R202403	Post-travel	G	G	117	-
R100241	Post-travel	C	B1	120	-
R200907	Pre-travel	B2	B2	126	-
R200908	Pre-travel	B2	B2	126	-
R302528	Pre-travel	B2	B2	127	-
R201875	Pre-travel	B2	B2	131	-
R202702	Pre-travel	B2	B2	131	-
R202185	Pre-travel	B2	B2	141	-
R301186	Post-travel	B2	B2	141	-
R200650	Post-travel	B1	B1	155	-
R201062	Pre-travel	G	G	174	-
R301631	Post-travel	A	B1	191	-
R201465	Post-travel	A	A	206	-
R202015	Post-travel	A	A	206	-
R202483	Post-travel	A	A	227	-

R101081	Post-travel	A	B1	278	-
R201711	Post-travel	B1	B1	278	-
R200055	Post-travel	B1	B1	295	-
R302393	Post-travel	B1	B1	295	-
R100661	Post-travel	B1	B1	316	-
R200860	Post-travel	B1	B1	316	-
R201345	Pre-travel	B2	B2	357	-
R101285	Pre-travel	D	D	362	-
R502780	Pre-travel	D	D	394	-
R302950	Post-travel	A	A	398	-
R202785	Pre-travel	D	D	405	-
R200649	Post-travel	C	C	410	-
R201746	Post-travel	C	C	410	-
R300437	Pre-travel	B2	B2	421	-
R201198	Post-travel	B1	B1	448	-
R101897	Pre-travel	B2	B2	452	-
R202509	Pre-travel	B2	B2	452	-
R300140	Pre-travel	B2	B2	452	-
R200114	Post-travel	B2	B2	476	-
R101578	Post-travel	B1	B1	517	-
R202483	Pre-travel	A	A	540	-
R200660	Pre-travel	E	E	543	-
R201196	Pre-travel	B2	B2	550	-
R201517	Pre-travel	A	A	757	-
R200752	Post-travel	B1	B1	906	-
R201921	Post-travel	B1	B1	1249	-
R401016	Pre-travel	A	E	1277	-
R200573	Post-travel	A	A	1638	-
R300659	Post-travel	B1	B1	1800	-
R202334	Pre-travel	D	D	1884	-
R101265	Pre-travel	A	A	2349	-
R202028	Post-travel	A	A	2349	-
R102889	Post-travel	A	A	2624	-
R101440	Post-travel	A	A	2705	-
R202130	Post-travel	E	E	3018	-
R201196	Post-travel	E	E	3036	-
R201159	Post-travel	A	A	3075	-
R200977	Pre-travel	B2	B2	3672	-
R202771	Pre-travel	B1	B1	3765	-
R201615	Pre-travel	-	A	3877	-
R202518	Pre-travel	B1	B1	4054	-
R101578	Pre-travel	A	A	4093	-
R300659	Pre-travel	B1	B1	5528	-
R202771	Post-travel	B1	B1	5614	-
R200360	Post-travel	B1	B1	5869	-
R202932	Pre-travel	B2	B2	6159	-
R502498	Post-travel	A	A	6420	-
R102887	Post-travel	A	A	6431	-
R102045	Post-travel	A	A	9538	-
R101093	Pre-travel	A	A	10912	-
R101288	Post-travel	B2	B2	10976	-
R302784	Pre-travel	B2	B2	11777	-
R100956	Post-travel	A	A	-	Unclear
R101284	Post-travel	A	A	-	Unclear
R201159	Pre-travel	A	A	-	Unclear
R202879	Post-travel	A	A	-	48

R302988	Post-travel	A	A	-	1312
R100665	Post-travel	A	A	-	10
R100695	Post-travel	A	A	-	685
R101897	Post-travel	A	A	-	1312
R200565	Post-travel	A	A	-	4
R201921	Pre-travel	A	A	-	1312
R202407	Post-travel	A	A	-	Unclear
R202802	Pre-travel	A	A	-	685
R401639	Post-travel	A	A	-	685
R502730	Pre-travel	A	A	-	Unclear
R102667	Pre-travel	A	A	-	10
R102901	Pre-travel	A	A	-	10
R200351	Post-travel	A	A	-	Unclear
R201062	Post-travel	A	A	-	Unclear
R201915	Post-travel	A	A	-	685
R202279	Post-travel	A	A	-	685
R202721	Post-travel	A	A	-	Unclear
R202802	Post-travel	A	A	-	Unclear
R502730	Post-travel	A	A	-	Unclear
R102341	Post-travel	A	A	-	Unclear
R200060	Post-travel	A	A	-	Unclear
R200735	Post-travel	A	A	-	757
R202055	Pre-travel	A	A	-	10
R200356	Post-travel	A	B1	-	Unclear
R101265	Post-travel	A	B1	-	Unclear
R301139	Post-travel	A	E	-	Unclear
R202569	Pre-travel	A	D	-	38
R102852	Pre-travel	A	Unclear	-	Unclear
R200929	Post-travel	A	Unclear	-	Unclear
R302528	Post-travel	A	Unclear	-	Unclear
R101984	Post-travel	B1	A	-	685
R102341	Pre-travel	B1	B1	-	678
R200384	Pre-travel	B1	B1	-	4213
R201345	Post-travel	B1	B1	-	40
R100696	Post-travel	B1	B1	-	223
R101533	Post-travel	B1	B1	-	Unclear
R102666	Post-travel	B1	B1	-	1086
R201060	Post-travel	B1	B1	-	2354
R201961	Post-travel	B1	B1	-	Unclear
R200908	Post-travel	B1	B1	-	441
R200978	Post-travel	B1	B1	-	678
R202158	Post-travel	B1	B1	-	2328
R302008	Post-travel	B1	B1	-	200
R102276	Post-travel	B1	B1	-	1086
R102886	Post-travel	B1	Unclear	-	Unclear
R101440	Pre-travel	B2	B2	-	14
R200894	Pre-travel	B2	B2	-	95
R201465	Pre-travel	B2	B2	-	95
R300479	Pre-travel	B2	B2	-	28
R200114	Pre-travel	B2	B2	-	Unclear
R200755	Pre-travel	B2	B2	-	144
R202185	Post-travel	B2	B2	-	410
R300437	Post-travel	B2	B2	-	131
R300501	Pre-travel	B2	B2	-	95
R302988	Pre-travel	B2	B2	-	95
R201961	Pre-travel	B2	B2	-	141

R302906	Pre-travel	B2	B2	-	Unclear
R202783	Post-travel	C	C	-	410
R200668	Pre-travel	D	D	-	38
R200977	Post-travel	D	D	-	38
R201211	Pre-travel	D	D	-	Unclear
R201615	Post-travel	D	D	-	Unclear
R201711	Pre-travel	D	D	-	69
R202879	Pre-travel	D	D	-	38
R202569	Post-travel	D	D	-	Unclear
R202518	Post-travel	D	D	-	Unclear
R400994	Post-travel	D	D	-	Unclear
R202015	Pre-travel	D	D	-	Unclear
R200781	Post-travel	D	D	-	69
R300500	Post-travel	E	E	-	Unclear
R201463	Post-travel	E	E	-	Unclear
R201875	Post-travel	E	E	-	Unclear
R200218	Pre-travel	E	E	-	Unclear
R302776	Pre-travel	E	E	-	Unclear
R200323	Post-travel	E	Unclear	-	Unclear
R200574	Post-travel	F	B2	-	12
R201940	Post-travel	F	B2	-	131
R200929	Pre-travel	F	F	-	Unclear
R201463	Pre-travel	F	F	-	Unclear
R101093	Post-travel	F	F	-	648
R101984	Pre-travel	F	Unclear	-	Unclear
R200375	Post-travel	G	D	-	38
R200247	Post-travel	-	A	-	716
R300140	Post-travel	-	A	-	Unclear
R200218	Post-travel	-	B1	-	Unclear
R200734	Post-travel	-	Unclear	-	Unclear

Table A.3: Description of *E. coli* MAG MLST alleles

Traveller ID	**Time Point**	**Closest Branch ST**	***adk* Allele**	***fumc* Allele**	***gyrB* Allele**	***icd* Allele**	***mdh* Allele**	***purA* Allele**	***recA* Allele**
R301633	Post	-	6	5	4	8	8	8	2
R102767	Post	-	10	11	4	8	8	8	2
R100887	Post	-	10	11	4	8	8	8	2
R100889	Post	-	10	11	4	8	8	8	2
R101285	Post	-	10	11	4	8	8	8	2
R101362	Pre	-	10	11	4	8	8	8	2
R102767	Pre	-	10	11	4	8	8	8	2
R102852	Post	-	10	11	4	8	8	8	2
R102946	Post	-	10	11	4	8	8	8	2
R200595	Post	-	10	11	4	8	8	8	2
R201153	Post	-	10	11	4	8	8	8	2
R201960	Pre	-	10	11	4	8	8	8	2
R202347	Pre	-	10	11	4	8	8	8	2
R202347	Post	-	10	11	4	8	8	8	2
R202459	Post	-	10	11	4	8	8	8	2
R202785	Post	-	10	11	4	8	8	8	2
R202835	Post	-	10	11	4	8	8	8	2
R300500	Pre	-	10	11	4	8	8	8	2

R302776	Post	-	10	11	4	8	8	8	2
R200109	Post	-	6	6	5	9	9	8	2
R102496	Post	-	18	22	17	6	5	5	4
R100887	Pre	-	10	11	4	1	8	8	2
R101018	Post	-	10	11	4	1	8	8	2
R201517	Post	-	10	11	4	1	8	8	2
R202056	Post	-	10	11	4	1	8	8	2
R202702	Post	-	10	11	4	1	8	8	2
R101601	Post	-	24	11	4	8	8	8	2
R301773	Post	-	24	11	4	8	8	8	2
R102045	Pre	-	6	11	4	8	8	8	2
R102557	Post	-	6	11	4	8	8	8	2
R202130	Pre	-	6	11	4	8	8	8	2
R201211	Post	-	26	7	4	1	8	8	6
R200595	Pre	-	6	4	4	18	24	5	14
R102514	Pre	-	6	4	4	16	24	8	14
R200384	Post	-	6	4	4	16	24	8	14
R200651	Post	-	6	4	4	16	24	8	14
R100956	Pre	-	27	32	24	29	26	19	22
R200247	Pre	-	27	32	24	29	26	19	22
R201198	Pre	-	27	32	24	29	26	19	22
R302008	Pre	-	27	32	24	29	26	19	22
R101601	Pre	-	21	35	27	6	5	5	4
R200574	Pre	-	21	35	27	6	5	5	4
R201348	Pre	-	21	35	27	6	5	5	4
R202158	Pre	-	21	35	27	6	5	5	4
R202517	Post	-	21	35	27	6	5	5	4
R301773	Pre	-	21	35	27	6	5	5	4
R102887	Pre	-	36	24	9	13	17	11	25
R200323	Pre	-	36	24	9	13	17	11	25
R202472	Pre	-	36	24	9	13	17	11	25
R301633	Pre	-	36	24	9	13	17	11	25
R102496	Pre	-	13	24	19	14	23	1	10
R300442	Pre	-	6	11	4	10	7	8	6
R200947	Pre	-	37	38	19	37	17	11	26
R201915	Pre	-	37	38	19	37	17	11	26
R502498	Pre	-	37	38	19	37	17	11	26
R201299	Post	-	20	45	41	43	5	32	2
R202403	Post	-	20	45	41	43	5	32	2
R100241	Post	-	49	4	44	9	11	35	7
R200907	Pre	-	13	43	19	37	17	25	25
R200908	Pre	-	13	43	19	37	17	25	25
R302528	Pre	-	13	14	19	36	23	11	10
R201875	Pre	-	53	40	47	13	36	28	29
R202702	Pre	-	53	40	47	13	36	28	29
R202185	Pre	-	13	52	10	14	17	25	17
R301186	Post	-	13	52	10	14	17	25	17
R200650	Post	-	6	4	14	16	24	8	14
R201062	Pre	-	20	67	56	64	5	50	45
R301631	Post	-	6	29	5	18	11	8	41
R201465	Post	-	6	7	5	1	8	18	2
R202015	Post	-	6	7	5	1	8	18	2
R202483	Post	-	10	11	4	1	8	9	2
R101081	Post	-	9	23	64	18	11	8	6
R201711	Post	-	9	23	64	18	11	8	6
R200055	Post	-	6	4	12	1	9	2	7

R302393	Post	-	6	4	12	1	9	2	7
R100661	Post	-	6	23	33	18	11	8	6
R200860	Post	-	6	23	33	18	11	8	6
R201345	Pre	-	13	40	13	13	23	25	66
R101285	Pre	-	62	100	17	31	5	5	4
R502780	Pre	-	21	35	61	52	5	5	4
R302950	Post	-	64	7	1	1	8	8	6
R202785	Pre	-	35	37	29	25	4	5	73
R200649	Post	-	6	4	12	1	20	18	7
R201746	Post	-	6	4	12	1	20	18	7
R300437	Pre	-	37	38	19	37	17	8	26
R201198	Post	-	6	6	5	16	11	8	7
R101897	Pre	-	76	43	19	37	30	1	25
R202509	Pre	-	76	43	19	37	30	1	25
R300140	Pre	-	76	43	19	37	30	1	25
R200114	Post	-	13	13	9	13	16	82	9
R101578	Post	-	109	65	5	1	9	13	14
R202483	Pre	-	6	7	57	1	8	8	2
R200660	Pre	-	83	136	110	117	80	1	2
R201196	Pre	-	14	14	10	14	17	92	10
R201517	Pre	-	136	11	4	1	8	8	2
R200752	Post	-	6	4	3	16	11	8	6
R201921	Post	-	6	23	5	16	24	18	6
R401016	Pre	-	127	232	4	12	160	162	17
R200573	Post	-	10	11	4	12	8	45	2
R300659	Post	-	9	19	33	18	9	122	6
R202334	Pre	-	62	3	58	6	5	16	4
R101265	Pre	-	6	11	4	10	8	8	2
R202028	Post	-	6	11	4	10	8	8	2
R102889	Post	-	6	370	4	8	8	78	2
R101440	Post	-	10	11	4	10	7	8	2
R202130	Post	-	12	58	54	344	1	2	47
R201196	Post	-	207	297	260	254	1	261	2
R201159	Post	-	10	23	109	8	270	8	2
R200977	Pre	-	76	13	190	13	17	30	25
R202771	Pre	-	6	4	4	16	312	8	14
R201615	Pre	-	10	11	369	8	8	8	2
R202518	Pre	-	6	4	4	16	336	8	14
R101578	Pre	-	6	7	386	1	8	8	2
R300659	Pre	-	6	6	5	16	9	13	6
R202771	Post	-	6	6	5	16	11	26	6
R200360	Post	-	80	95	3	18	11	8	333
R202932	Pre	-	552	40	239	13	36	28	30
R502498	Post	-	24	11	4	8	230	8	2
R102887	Post	-	10	11	4	8	25	8	2
R102045	Post	-	136	11	831	12	7	8	2
R101093	Pre	-	10	541	4	8	970	8	2
R101288	Post	-	76	1577	9	36	17	14	10
R302784	Pre	-	37	38	19	37	17	911	26
R100956	Post	Unclear	10	687	4	8	8	8	2
R101284	Post	Unclear	10	11	22	8	8	8	2
R201159	Pre	Unclear	10	11	4	8	6	8	6
R202879	Post	48	10	70	57	8	8	1	2
R302988	Post	1312	6	11	4	104	8	7	2
R100665	Post	10	24	11	4	8	-	8	2
R100695	Post	685	10	-	64	1	6	8	6

R101897	Post	1312	6	-	4	113	8	1	2
R200565	Post	4	6	5	4	8	-	8	2
R201921	Pre	1312	6	11	95	104	8	8	-
R202407	Post	Unclear	-	11	5	8	8	8	2
R202802	Pre	685	-	11	4	8	8	8	2
R401639	Post	685	-	4	96	1	24	8	6
R502730	Pre	Unclear	10	-	4	8	509	695	381
R102667	Pre	10	10	-	4	-	8	8	2
R102901	Pre	10	10	-	4	-	8	8	2
R200351	Post	Unclear	-	-	22	8	8	8	25
R201062	Post	Unclear	10	-	4	-	8	8	2
R201915	Post	685	673	-	109	-	270	8	2
R202279	Post	685	-	11	22	-	8	7	2
R202721	Post	Unclear	6	-	4	10	42	8	-
R202802	Post	Unclear	59	-	4	1	12	8	-
R502730	Post	Unclear	10	-	5	8	8	7	-
R102341	Post	Unclear	10	-	4	-	-	8	2
R200060	Post	Unclear	-	11	22	-	-	8	2
R200735	Post	757	6	-	-	1	8	-	-
R202055	Pre	10	-	-	-	-	99	8	-
R200356	Post	Unclear	6	4	56	-	11	8	6
R101265	Post	Unclear	-	7	5	-	773	8	6
R301139	Post	Unclear	12	136	-	629	80	8	2
R202569	Pre	38	4	26	2	-	5	5	19
R102852	Pre	Unclear	6	-	5	8	8	8	381
R200929	Post	Unclear	1	4	4	522	7	8	-
R302528	Post	Unclear	-	-	4	-	12	8	2
R101984	Post	685	-	-	261	91	91	78	6
R102341	Pre	678	6	420	32	16	522	8	156
R200384	Pre	4213	6	704	15	18	9	8	156
R201345	Post	40	6	1418	5	26	20	8	14
R100696	Post	223	6	4	4	-	522	8	14
R101533	Post	Unclear	6	-	5	9	9	8	2
R102666	Post	1086	-	6	44	45	433	13	14
R201060	Post	2354	6	4	4	-	7	8	14
R201961	Post	Unclear	6	6	5	-	9	8	2
R200908	Post	441	6	23	-	54	-	695	2
R200978	Post	678	-	6	5	56	522	26	-
R202158	Post	2328	127	-	848	-	11	122	34
R302008	Post	200	6	-	5	-	91	8	381
R102276	Post	1086	-	6	5	-	7	-	381
R102886	Post	Unclear	-	-	4	-	7	8	381
R101440	Pre	14	14	-	10	14	17	92	10
R200894	Pre	95	37	38	19	-	17	11	26
R201465	Pre	95	37	38	19	37	17	8	-
R300479	Pre	28	13	-	19	22	17	14	15
R200114	Pre	Unclear	-	14	13	13	17	11	13
R200755	Pre	144	13	-	9	-	30	64	268
R202185	Post	410	6	4	25	-	7	-	6
R300437	Post	131	-	-	47	10	36	28	29
R300501	Pre	95	37	-	86	-	17	11	26
R302988	Pre	95	37	-	286	-	17	11	26
R201961	Pre	141	-	38	-	13	17	-	34
R302906	Pre	Unclear	-	107	494	-	17	23	-
R202783	Post	410	6	4	60	1	20	7	7
R200668	Pre	38	21	26	27	-	5	5	19

R200977	Post	38	4	26	2	-	5	5	19
R201211	Pre	Unclear	21	-	61	52	5	5	4
R201615	Post	Unclear	35	35	61	-	5	318	4
R201711	Pre	69	21	35	27	-	5	5	4
R202879	Pre	38	4	26	2	-	5	5	19
R202569	Post	Unclear	34	-	39	87	67	16	4
R202518	Post	Unclear	21	-	61	52	5	5	4
R400994	Post	Unclear	54	-	211	-	224	16	4
R202015	Pre	Unclear	-	-	22	6	-	5	181
R200781	Post	69	18	-	-	-	12	-	4
R300500	Post	Unclear	6	31	5	-	1	1	2
R201463	Post	Unclear	207	297	260	-	1	261	2
R201875	Post	Unclear	127	161	-	-	91	1	2
R200218	Pre	Unclear	-	-	-	117	80	1	2
R302776	Pre	Unclear	-	136	-	117	80	-	2
R200323	Post	Unclear	-	-	329	269	216	17	-
R200574	Post	12	13	-	9	-	16	10	6
R201940	Post	131	13	-	47	-	-	-	29
R200929	Pre	Unclear	27	32	24	-	26	19	22
R201463	Pre	Unclear	28	33	25	29	7	11	-
R101093	Post	648	-	4	-	-	70	214	-
R101984	Pre	Unclear	63	-	27	-	7	8	216
R200375	Post	38	-	-	2	-	-	5	19
R200247	Post	716	10	7	4	-	8	8	2
R300140	Post	Unclear	10	11	4	8	-	8	2
R200218	Post	Unclear	6	-	7	1	-	193	56
R200734	Post	Unclear	10	-	4	-	6	8	-

References

1. Sarowska J, Futoma-Koloch B, Jama-Kmiecik A, Frej-Madrzak M, Ksiazczyk M, Bugla-Ploskonska G, and Choroszy-Krol I. Virulence factors, prevalence and potential transmission of extraintestinal pathogenic Escherichia coli isolated from different sources: recent reports. Gut Pathogens. 2019 Feb 21; 11:10. DOI: 10.1186/s13099-019-0290-0
2. Croxen MA and Finlay BB. Molecular mechanisms of Escherichia coli pathogenicity. Nature Reviews Microbiology. 2010 Jan; 8:26–38. DOI: 10.1038/nrmicro2265
3. Lu S, Jin D, Wu S, Yang J, Lan R, Bai X, Liu S, Meng Q, Yuan X, Zhou J, Pu J, Chen Q, Dai H, Hu Y, Xiong Y, Ye C, and Xu J. Insights into the evolution of pathogenicity of Escherichia coli from genomic analysis of intestinal E. coli of Marmota himalayana in Qinghai–Tibet plateau of China. Emerging Microbes & Infections. 2016 Jan 1; 5:1–9. DOI: 10.1038/emi.2016.122
4. Clermont O, Christenson JK, Denamur E, and Gordon DM. The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environmental Microbiology Reports. 2013; 5:58–65. DOI: 10.1111/1758-2229.12019
5. Walk ST, Alm EW, Gordon DM, Ram JL, Toranzos GA, Tiedje JM, and Whittam TS. Cryptic Lineages of the Genus Escherichia. Applied and Environmental Microbiology. 2009 Oct 15; 75:6534–44. DOI: 10.1128/AEM.01262-09
6. Walk ST. The "Cryptic" Escherichia. EcoSal Plus. 2015 Jul 30; 6. DOI: 10.1128/ecosalplus.ESP-0002-2015
7. Reid SD, Herbelin CJ, Bumbaugh AC, Selander RK, and Whittam TS. Parallel evolution of virulence in pathogenic Escherichia coli. Nature. 2000 Jul; 406:64–7. DOI: 10.1038/35017546

8. Larsen MV, Cosentino S, Rasmussen S, Friis C, Hasman H, Marvig RL, Jelsbak L, Sicheritz-Pontén T, Ussery DW, Aarestrup FM, and Lund O. Multilocus sequence typing of total-genome-sequenced bacteria. Journal of Clinical Microbiology. 2012 Apr; 50:1355–61. DOI: 10.1128/JCM.06094-11

9. Mare AD, Ciurea CN, Man A, Tudor B, Moldovan V, Decean L, and Toma F. Enteropathogenic Escherichia coli—A Summary of the Literature. Gastroenterology Insights. 2021 Mar; 12:28–40. DOI: 10.3390/gastroent12010004

10. Chan YS and Ng TB. Shiga toxins: from structure and mechanism to applications. Applied Microbiology and Biotechnology. 2016 Feb 1; 100:1597–610. DOI: 10.1007/s00253-015-7236-3

11. Karch H, Tarr PI, and Bielaszewska M. Enterohaemorrhagic Escherichia coli in human medicine. International Journal of Medical Microbiology. 2005 Oct 5; 295:405–18. DOI: 10.1016/j.ijmm.2005.06.009

12. Qadri F, Svennerholm AM, Faruque ASG, and Sack RB. Enterotoxigenic Escherichia coli in Developing Countries: Epidemiology, Microbiology, Clinical Features, Treatment, and Prevention. Clinical Microbiology Reviews. 2005 Jul; 18:465–83. DOI: 10.1128/CMR.18.3.465-483.2005

13. Wiles TJ, Kulesus RR, and Mulvey MA. Origins and virulence mechanisms of uropathogenic Escherichia coli. Experimental and Molecular Pathology. Special Issue: Molecular Pathology and Molecular Diagnostics 2008 Aug 1; 85:11–9. DOI: 10.1016/j.yexmp.2008.03.007

14. Smith JL, Fratamico PM, and Gunther NW. Extraintestinal Pathogenic Escherichia coli. Foodborne Pathogens and Disease. 2007 Jun; 4:134–63. DOI: 10.1089/fpd.2007.0087

15. Holt DE, Halket S, Louvois Jd, and Harvey D. Neonatal meningitis in England and Wales: 10 years on. Archives of Disease in Childhood - Fetal and Neonatal Edition. 2001 Mar 1; 84:F85–F89. DOI: 10.1136/fn.84.2.F85

16. Watanabe H, Wada A, Inagaki Y, Itoh Ki, and Tamura K. Outbreaks of enterohaemorrhagic Escherichia colt O157:H7 infection by two different genotype strains in Japan, 1996. The Lancet. 1996 Sep 21; 348:831–2. DOI: 10.1016/S0140-6736(05)65257-9

17. Bell BP, Goldoft M, Griffin PM, Davis MA, Gordon DC, Tarr PI, Bartleson CA, Lewis JH, Barrett TJ, Wells JG, Baron R, and Kobayashi J. A Multistate Outbreak of Escherichia coli O157:H7—Associated Bloody Diarrhea and Hemolytic Uremic Syndrome From Hamburgers: The Washington Experience. JAMA. 1994 Nov 2; 272:1349–53. DOI: 10.1001/jama.1994.03520170059036

18. Dundas S, Todd WTA, Stewart AI, Murdoch PS, Chaudhuri AKR, and Hutchinson SJ. The Central Scotland Escherichia coli O157:H7 Outbreak: Risk Factors for the Hemolytic Uremic Syndrome and Death among Hospitalized Patients. Clinical Infectious Diseases. 2001 Oct 1; 33:923–31. DOI: 10.1086/322598

19. Buchholz U, Bernard H, Werber D, Böhmer MM, Remschmidt C, Wilking H, Deleré Y, Heiden M an der, Adlhoch C, Dreesman J, Ehlers J, Ethelberg S, Faber M, Frank C, Fricke G, Greiner M, Höhle M, Ivarsson S, Jark U, Kirchner M, Koch J, Krause G, Luber P, Rosner B, Stark K, and Kühne M. German Outbreak of Escherichia coli O104:H4 Associated with Sprouts. New England Journal of Medicine. 2011 Nov 10; 365:1763–70. DOI: 10.1056/NEJMoa1106482

20. Petty NK, Ben Zakour NL, Stanton-Cook M, Skippington E, Totsika M, Forde BM, Phan MD, Gomes Moriel D, Peters KM, Davies M, Rogers BA, Dougan G, Rodriguez-Baño J, Pascual A, Pitout JDD, Upton M, Paterson DL, Walsh TR, Schembri MA, and Beatson SA. Global dissemination of a multidrug resistant Escherichia coli clone. Proceedings of the National Academy of Sciences. 2014 Apr 15; 111:5694–9. DOI: 10.1073/pnas.1322678111

21. Coque TM, Novais Â, Carattoli A, Poirel L, Pitout J, Peixe L, Baquero F, Cantón R, and Nordmann P. Dissemination of Clonally Related Escherichia coli Strains Expressing Extended-Spectrum beta-Lactamase CTX-M-15. Emerging Infectious Diseases. 2008 Feb; 14:195–200. DOI: 10.3201/eid1402.070350

22. Garcillán-Barcia MP, Alvarado A, and Cruz F de la. Identification of bacterial plasmids based on mobility and plasmid population biology. FEMS Microbiology Reviews. 2011 Sep 1; 35:936–56. DOI: 10.1111/j.1574-6976.2011.00291.x

23. Wozniak RAF and Waldor MK. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. Nature Reviews Microbiology. 2010 Aug; 8:552–63. DOI: 10.1038/nrmicro2382

24. Johnston C, Martin B, Fichant G, Polard P, and Claverys JP. Bacterial transformation: distribution, shared mechanisms and divergent control. Nature Reviews Microbiology. 2014 Mar; 12:181–96. DOI: 10.1038/nrmicro3199

25. Charpentier X, Kay E, Schneider D, and Shuman HA. Antibiotics and UV Radiation Induce Competence for Natural Transformation in Legionella pneumophila. Journal of Bacteriology. 2011 Mar; 193:1114–21. DOI: 10.1128/JB.01146-10

26. Colomer-Lluch M, Jofre J, and Muniesa M. Antibiotic Resistance Genes in the Bacteriophage DNA Fraction of Environmental Samples. PLOS ONE. 2011 Mar 3; 6:e17549. DOI: 10.1371/journal.pone.0017549

27. Mena A, Plasencia V, García L, Hidalgo O, Ayestarán JI, Alberti S, Borrell N, Pérez JL, and Oliver A. Characterization of a Large Outbreak by CTX-M-1-Producing Klebsiella pneumoniae and Mechanisms Leading to In Vivo Carbapenem Resistance Development. Journal of Clinical Microbiology. 2006 Aug; 44:2831–7. DOI: 10.1128/JCM.00418-06

28. Webber MA and Piddock LJV. The importance of efflux pumps in bacterial antibiotic resistance. Journal of Antimicrobial Chemotherapy. 2003 Jan 1; 51:9–11. DOI: 10.1093/jac/dkg050

29. Nolivos S, Cayron J, Dedieu A, Page A, Delolme F, and Lesterlin C. Role of AcrAB-TolC multidrug efflux pump in drug-resistance acquisition by plasmid transfer. Science. 2019 May 24; 364:778–82. DOI: 10.1126/science.aav6390

30. Dolejska M, Villa L, Poirel L, Nordmann P, and Carattoli A. Complete sequencing of an IncHI1 plasmid encoding the carbapenemase NDM-1, the ArmA 16S RNA methylase and a resistance–nodulation–cell division/multidrug efflux pump. Journal of Antimicrobial Chemotherapy. 2013 Jan 1; 68:34–9. DOI: 10.1093/jac/dks357

31. Aldred KJ, Kerns RJ, and Osheroff N. Mechanism of Quinolone Action and Resistance. Biochemistry. 2014 Mar 18; 53:1565–74. DOI: 10.1021/bi5000564

32. Dixon RA and Chopra I. Leakage of periplasmic proteins from Escherichia coli mediated by polymyxin B nonapeptide. Antimicrobial Agents and Chemotherapy. 1986 May; 29:781–8. DOI: 10.1128/AAC.29.5.781

33. Poirel L, Jayol A, and Nordmann P. Polymyxins: Antibacterial Activity, Susceptibility Testing, and Resistance Mechanisms Encoded by Plasmids or Chromosomes. Clinical Microbiology Reviews. 2017 Apr; 30:557–96. DOI: 10.1128/CMR.00064-16

34. Liu YY, Wang Y, Walsh TR, Yi LX, Zhang R, Spencer J, Doi Y, Tian G, Dong B, Huang X, Yu LF, Gu D, Ren H, Chen X, Lv L, He D, Zhou H, Liang Z, Liu JH, and Shen J. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. The Lancet Infectious Diseases. 2016 Feb 1; 16:161–8. DOI: 10.1016/S1473-3099(15)00424-7

35. Ramirez MS and Tolmasky ME. Aminoglycoside modifying enzymes. Drug Resistance Updates. 2010 Dec 1; 13:151–71. DOI: 10.1016/j.drup.2010.08.003

36. Bush K and Jacoby GA. Updated Functional Classification of beta-Lactamases. Antimicrobial Agents and Chemotherapy. 2010 Mar; 54:969–76. DOI: 10.1128/AAC.01009-09

37. Canton R, Gonzalez-Alba JM, and Galán JC. CTX-M Enzymes: Origin and Diffusion. Frontiers in Microbiology. 2012; 3

38. Humeniuk C, Arlet G, Gautier V, Grimont P, Labia R, and Philippon A. beta-Lactamases of Kluyvera ascorbata, Probable Progenitors of Some Plasmid-Encoded CTX-M Types. Antimicrobial Agents and Chemotherapy. 2002 Sep; 46:3045–9. DOI: 10.1128/AAC.46.9.3045-3049.2002

39. Rodriguez-Villalobos H, Bogaerts P, Berhin C, Bauraing C, Deplano A, Montesinos I, Mendonça R de, Jans B, and Glupczynski Y. Trends in production of extended-spectrum beta-lactamases among Enterobacteriaceae of clinical interest: results of a nationwide survey in Belgian hospitals. Journal of Antimicrobial Chemotherapy. 2011 Jan 1; 66:37–47. DOI: 10.1093/jac/dkq388

40. Kim J, Bae IK, Jeong SH, Chang CL, Lee CH, and Lee K. Characterization of IncF plasmids carrying the blaCTX-M-14 gene in clinical isolates of Escherichia coli from Korea. Journal of Antimicrobial Chemotherapy. 2011 Jun 1; 66:1263–8. DOI: 10.1093/jac/dkr106

41. Dhanji H, Patel R, Wall R, Doumith M, Patel B, Hope R, Livermore DM, and Woodford N. Variation in the genetic environments of blaCTX-M-15 in Escherichia coli from the faeces of travellers returning to the United Kingdom. Journal of Antimicrobial Chemotherapy. 2011 May 1; 66:1005–12. DOI: 10.1093/jac/dkr041

42. Bevan ER, Jones AM, and Hawkey PM. Global epidemiology of CTX-M beta-lactamases: temporal and geographical shifts in genotype. Journal of Antimicrobial Chemotherapy. 2017 Aug 1; 72:2145–55. DOI: 10.1093/jac/dkx146

43. Poirel L, Decousser JW, and Nordmann P. Insertion Sequence ISEcp1B Is Involved in Expression and Mobilization of a blaCTX-M beta-Lactamase Gene. Antimicrobial Agents and Chemotherapy. 2003 Sep; 47:2938–45. DOI: 10.1128/AAC.47.9.2938-2945.2003

44. Novais Â, Cantón R, Moreira R, Peixe L, Baquero F, and Coque TM. Emergence and Dissemination of Enterobacteriaceae Isolates Producing CTX-M-1-Like Enzymes in Spain Are Associated with IncFII (CTX-M-15) and Broad-Host-Range (CTX-M-1, -3, and -32) Plasmids. Antimicrobial Agents and Chemotherapy. 2007 Feb; 51:796–9. DOI: 10.1128/AAC.01070-06

45. Falgenhauer L, Yao Y, Fritzenwanker M, Schmiedel J, Imirzalioglu C, and Chakraborty T. Complete Genome Sequence of Phage-Like Plasmid pECOH89, Encoding CTX-M-15. Genome Announcements. 2014 May; 2:e00356–14. DOI: 10.1128/genomeA.00356-14

46. Been Md, Lanza VF, Toro Md, Scharringa J, Dohmen W, Du Y, Hu J, Lei Y, Li N, Tooming-Klunderud A, Heederik DJJ, Fluit AC, Bonten MJM, Willems RJL, Cruz Fdl, and Schaik Wv. Dissemination of Cephalosporin Resistance Genes between Escherichia coli Strains from Farm Animals and Humans by Specific Plasmid Lineages. PLOS Genetics. 2014 Dec 18; 10:e1004776. DOI: 10.1371/journal.pgen.1004776

47. Arcilla MS, Hattem JM van, Haverkate MR, Bootsma MCJ, Genderen PJJ van, Goorhuis A, Grobusch MP, Lashof AMO, Molhoek N, Schultsz C, Stobberingh EE, Verbrugh HA, Jong MD de, Melles DC, and Penders J. Import and spread of extended-spectrum beta-lactamase-producing Enterobacteriaceae by international travellers (COMBAT study): a prospective, multicentre cohort study. Lancet Infectious Diseases. 2017; 17:78–85. DOI: 10.1016/s1473-3099(16)30319-x

48. Evans JM, Morris LS, and Marchesi JR. The gut microbiome: the role of a virtual organ in the endocrinology of the host. The Journal of Endocrinology. 2013 Sep; 218:R37–47. DOI: 10.1530/JOE-13-0131

49. Lozupone CA, Stombaugh JI, Gordon JI, Jansson JK, and Knight R. Diversity, stability and resilience of the human gut microbiota. Nature. 2012 Sep; 489:220–30. DOI: 10.1038/nature11550

50. Gritz EC and Bhandari V. The Human Neonatal Gut Microbiome: A Brief Review. Frontiers in Pediatrics. 2015; 3

51. O'Neill I, Schofield Z, and Hall LJ. Exploring the role of the microbiota member Bifidobacterium in modulating immune-linked diseases. Emerging Topics in Life Sciences. 2017 Nov 30; 1. Ed. by Marchesi JR:333–49. DOI: 10.1042/ETLS20170058

52. Lawson MAE, O'Neill IJ, Kujawska M, Gowrinadh Javvadi S, Wijeyesekera A, Flegg Z, Chalklen L, and Hall LJ. Breast milk-derived human milk oligosaccharides promote Bifidobacterium interactions within a single ecosystem. The ISME Journal. 2020 Feb; 14:635–48. DOI: 10.1038/s41396-019-0553-2

53. Lin AE, Autran CA, Szyszka A, Escajadillo T, Huang M, Godula K, Prudden AR, Boons GJ, Lewis AL, Doran KS, Nizet V, and Bode L. Human milk oligosaccharides inhibit growth of group B Streptococcus. Journal of Biological Chemistry. 2017 Jul 7; 292:11243–9. DOI: 10.1074/jbc.M117.789974

54. Coppa GV, Zampini L, Galeazzi T, Facinelli B, Ferrante L, Capretti R, and Orazio G. Human Milk Oligosaccharides Inhibit the Adhesion to Caco-2 Cells of Diarrheal Pathogens: Escherichia coli, Vibrio cholerae, and Salmonella fyris.

Pediatric Research. 2006 Mar; 59:377–82. DOI: 10.1203/01.pdr.0000200805.45593.17

55. He Y, Liu S, Kling DE, Leone S, Lawlor NT, Huang Y, Feinberg SB, Hill DR, and Newburg DS. The human milk oligosaccharide 2-fucosyllactose modulates CD14 expression in human enterocytes, thereby attenuating LPS-induced inflammation. Gut. 2016 Jan 1; 65:33–46. DOI: 10.1136/gutjnl-2014-307544

56. Galazzo G, Best N van, Bervoets L, Dapaah IO, Savelkoul PH, Hornef MW, Hutton EK, Morrison K, Holloway AC, McDonald H, Ratcliffe EM, Stearns JC, Schertzer JD, Surette MG, Thabane L, Mommers M, Lau S, Hamelmann E, and Penders J. Development of the Microbiota and Associations With Birth Mode, Diet, and Atopic Disorders in a Longitudinal Analysis of Stool Samples, Collected From Infancy Through Early Childhood. Gastroenterology. 2020 May 1; 158:1584–96. DOI: 10.1053/j.gastro.2020.01.024

57. Eckburg PB, Bik EM, Bernstein CN, Purdom E, Dethlefsen L, Sargent M, Gill SR, Nelson KE, and Relman DA. Diversity of the Human Intestinal Microbial Flora. Science. 2005 Jun 10; 308:1635–8. DOI: 10.1126/science.1110591

58. Tian L, Wang XW, Wu AK, Fan Y, Friedman J, Dahlin A, Waldor MK, Weinstock GM, Weiss ST, and Liu YY. Deciphering functional redundancy in the human microbiome. en. Nature Communications 2020 Dec; 11. Number: 1 Publisher: Nature Publishing Group:6217. DOI: 10.1038/s41467-020-19940-1

59. Kriss M, Hazleton KZ, Nusbacher NM, Martin CG, and Lozupone CA. Low diversity gut microbiota dysbiosis: drivers, functional implications and recovery. Current Opinion in Microbiology. Microbiota 2018 Aug 1; 44:34–40. DOI: 10.1016/j.mib.2018.07.003

60. Dethlefsen L and Relman DA. Incomplete recovery and individualized responses of the human distal gut microbiota to repeated antibiotic perturbation. Proceedings of the National Academy of Sciences. 2011 Mar 15; 108(supplement_1):4554–61. DOI: 10.1073/pnas.1000087107

61. Xie H, Guo R, Zhong H, Feng Q, Lan Z, Qin B, Ward KJ, Jackson MA, Xia Y, Chen X, Chen B, Xia H, Xu C, Li F, Xu X, Al-Aama JY, Yang H, Wang J, Kristiansen K, Wang J, Steves CJ, Bell JT, Li J, Spector TD, and Jia H. Shotgun Metagenomics of 250 Adult Twins Reveals Genetic and Environmental Impacts on the Gut Microbiome. Cell Systems. 2016 Dec 21; 3:572–584.e3. DOI: 10.1016/j.cels.2016.10.004

62. Dominianni C, Sinha R, Goedert JJ, Pei Z, Yang L, Hayes RB, and Ahn J. Sex, Body Mass Index, and Dietary Fiber Intake Influence the Human Gut Microbiome. PLOS ONE. 2015 Apr 15; 10:e0124599. DOI: 10.1371/journal.pone.0124599

63. Markle JGM, Frank DN, Mortin-Toth S, Robertson CE, Feazel LM, Rolle-Kampczyk U, Bergen M von, McCoy KD, Macpherson AJ, and Danska JS. Sex Differences in the Gut Microbiome Drive Hormone-Dependent Regulation of Autoimmunity. Science. 2013 Mar; 339:1084–8. DOI: 10.1126/science.1233521

64. Singh RK, Chang HW, Yan D, Lee KM, Ucmak D, Wong K, Abrouk M, Farahnik B, Nakamura M, Zhu TH, Bhutani T, and Liao W. Influence of diet on the gut microbiome and implications for human health. Journal of Translational Medicine. 2017 Apr 8; 15:73. DOI: 10.1186/s12967-017-1175-y

65. Shen J, Zuo ZX, and Mao AP. Effect of probiotics on inducing remission and maintaining therapy in ulcerative colitis, Crohn's disease, and pouchitis: meta-analysis of randomized controlled trials. Inflammatory Bowel Diseases. 2014 Jan; 20:21–35. DOI: 10.1097/01.MIB.0000437495.30052.be

66. Palleja A, Mikkelsen KH, Forslund SK, Kashani A, Allin KH, Nielsen T, Hansen TH, Liang S, Feng Q, Zhang C, Pyl PT, Coelho LP, Yang H, Wang J, Typas A, Nielsen MF, Nielsen HB, Bork P, Wang J, Vilsbøll T, Hansen T, Knop FK, Arumugam M, and Pedersen O. Recovery of gut microbiota of healthy adults following antibiotic exposure. Nature Microbiology. 2018 Nov; 3:1255–65. DOI: 10.1038/s41564-018-0257-9

67. Morrison DJ and Preston T. Formation of short chain fatty acids by the gut microbiota and their impact on human metabolism. Gut Microbes. 2016 May 3; 7:189–200. DOI: 10.1080/19490976.2015.1134082

68. Shreiner AB, Kao JY, and Young VB. The gut microbiome in health and in disease. Current opinion in gastroenterology. 2015 Jan; 31:69–75. DOI: 10.1097/MOG.0000000000000139

69. Inagaki T, Moschetta A, Lee YK, Peng L, Zhao G, Downes M, Yu RT, Shelton JM, Richardson JA, Repa JJ, Mangelsdorf DJ, and Kliewer SA. Regulation of antibacterial defense in the small intestine by the nuclear bile acid receptor. Proceedings of the National Academy of Sciences of the United States of America. 2006 Mar 7; 103:3920–5. DOI: 10.1073/pnas.0509592103

70. Vavassori P, Mencarelli A, Renga B, Distrutti E, and Fiorucci S. The bile acid receptor FXR is a modulator of intestinal innate immunity. Journal of Immunology (Baltimore, Md: 1950). 2009 Nov 15; 183:6251–61. DOI: 10.4049/jimmunol.0803978

71. Monteleone I, Rizzo A, Sarra M, Sica G, Sileri P, Biancone L, MacDonald TT, Pallone F, and Monteleone G. Aryl Hydrocarbon Receptor-Induced Signals Up-regulate IL-22 Production and Inhibit Inflammation in the Gastrointestinal Tract. Gastroenterology. 2011 Jul 1; 141:237–248.e1. DOI: 10.1053/j.gastro.2011.04.007

72. Qiu J, Heller JJ, Guo X, Chen ZmE, Fish K, Fu YX, and Zhou L. The Aryl Hydrocarbon Receptor Regulates Gut Immunity through Modulation of Innate Lymphoid Cells. Immunity. 2012 Jan 27; 36:92–104. DOI: 10.1016/j.immuni.2011.11.011

73. Carlson AL, Xia K, Azcarate-Peril MA, Goldman BD, Ahn M, Styner MA, Thompson AL, Geng X, Gilmore JH, and Knickmeyer RC. Infant Gut Microbiome Associated With Cognitive Development. Biological Psychiatry. Early Life Environments and Later Vulnerabilities 2018 Jan 15; 83:148–59. DOI: 10.1016/j.biopsych.2017.06.021

74. Dinan TG and Cryan JF. Gut instincts: microbiota as a key regulator of brain development, ageing and neurodegeneration. The Journal of Physiology. 2017; 595:489–503. DOI: 10.1113/JP273106

75. Morais LH, Schreiber HL, and Mazmanian SK. The gut microbiota–brain axis in behaviour and brain disorders. Nature Reviews Microbiology. 2021 Apr; 19:241–55. DOI: 10.1038/s41579-020-00460-0

76. Bohnhoff M, Drake BL, and Miller CP. EFFECT OF STREPTOMYCIN ON SUSCEPTIBILITY OF INTESTINAL TRACT TO EXPERIMENTAL SALMONELLA INFECTION. Proceedings of the Society for Experimental Biology and Medicine. 1954; 86:132–7

77. Caballero-Flores G, Pickard JM, Fukuda S, Inohara N, and Núñez G. An Enteric Pathogen Subverts Colonization Resistance by Evading Competition for Amino Acids in the Gut. Cell Host & Microbe. 2020 Oct 7; 28:526–533.e5. DOI: 10.1016/j.chom.2020.06.018

78. Kamada N, Kim YG, Sham HP, Vallance BA, Puente JL, Martens EC, and Nunez G. Regulated Virulence Controls the Ability of a Pathogen to Compete with the Gut Microbiota. Science. 2012; 336:1325–9. DOI: 10.1126/science.1222195

79. Deriu E, Liu JZ, Pezeshki M, Edwards RA, Ochoa RJ, Contreras H, Libby SJ, Fang FC, and Raffatellu M. Probiotic Bacteria Reduce Salmonella Typhimurium Intestinal Colonization by Competing for Iron. Cell Host & Microbe. 2013 Jul 17; 14:26–37. DOI: 10.1016/j.chom.2013.06.007

80. Koenigsknecht MJ, Theriot CM, Bergin IL, Schumacher CA, Schloss PD, and Young VB. Dynamics and Establishment of Clostridium difficile Infection in the Murine Gastrointestinal Tract. Infection and Immunity. 2015 Mar; 83:934–41. DOI: 10.1128/IAI.02768-14

81. G. Arnison P, J. Bibb M, Bierbaum G, A. Bowers A, S. Bugni T, Bulaj G, A. Camarero J, J. Campopiano D, L. Challis G, Clardy J, D. Cotter P, J. Craik D, Dawson M, Dittmann E, Donadio S, C. Dorrestein P, Entian KD, A. Fischbach M, S. Garavelli J, Göransson U, W. Gruber C, H. Haft D, K. Hemscheidt T, Hertweck C, Hill C, R. Horswill A, Jaspars M, L. Kelly W, P. Klinman J, P. Kuipers O, James Link A, Liu W, A. Marahiel M, A. Mitchell D, N. Moll G, S. Moore B, Müller R, K. Nair S, F. Nes I, E. Norris G, M. Olivera B, Onaka H, L. Patchett M, Piel J, T. Reaney MJ, Rebuffat S, Paul Ross R, Sahl HG, W. Schmidt E, E. Selsted M, Severinov K, Shen B, Sivonen K, Smith L, Stein T, D. Süssmuth R, R. Tagg J, Tang GL, W. Truman A, C. Vederas J, T. Walsh C, D. Walton J, C. Wenzel S, M. Willey J, and Donk WAvd. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. Natural Product Reports. 2013; 30:108–60. DOI: 10.1039/C2NP20085F

82. Teng K, Huang F, Liu Y, Wang Y, Xia T, Yun F, and Zhong J. Food and gut originated bacteriocins involved in gut microbe-host interactions. Critical Reviews in Microbiology. 2022 Jun 17; 0:1–13. DOI: 10.1080/1040841X.2022.2082860

83. Duquesne S, Destoumieux-Garzón D, Peduzzi J, and Rebuffat S. Microcins, gene-encoded antibacterial peptides from enterobacteria. Natural Product Reports. 2007; 24:708–34. DOI: 10.1039/B516237H

84. Birchenough GMH, Nyström EEL, Johansson MEV, and Hansson GC. A sentinel goblet cell guards the colonic crypt by triggering Nlrp6-dependent Muc2 secretion. Science. 2016 Jun 24; 352:1535–42. DOI: 10.1126/science.aaf7419

85. Burger-van Paassen N, Vincent A, Puiman PJ, Sluis M van der, Bouma J, Boehm G, Goudoever JB van, Seuningen I van, and Renes IB. The regulation of intestinal mucin MUC2 expression by short-chain fatty acids: implications for epithelial protection. Biochemical Journal. 2009 May 13; 420:211–9. DOI: 10.1042/BJ20082222

86. Everard A, Belzer C, Geurts L, Ouwerkerk JP, Druart C, Bindels LB, Guiot Y, Derrien M, Muccioli GG, Delzenne NM, Vos WM de, and Cani PD. Cross-talk between Akkermansia muciniphila and intestinal epithelium controls diet-induced obesity. Proceedings of the National Academy of Sciences. 2013 May 28; 110:9066–71. DOI: 10.1073/pnas.1219451110

87. Nood E van, Vrieze A, Nieuwdorp M, Fuentes S, Zoetendal EG, Vos WM de, Visser CE, Kuijper EJ, Bartelsman JF, Tijssen JG, Speelman P, Dijkgraaf MG, and Keller JJ. Duodenal Infusion of Donor Feces for Recurrent Clostridium difficile. New England Journal of Medicine. 2013 Jan 31; 368:407–15. DOI: 10.1056/NEJMoa1205037

88. Duan H, Yu L, Tian F, Zhai Q, Fan L, and Chen W. Antibiotic-induced gut dysbiosis and barrier disruption and the potential protective strategies. Critical Reviews in Food Science and Nutrition. 2022 Mar 26; 62:1427–52. DOI: 10.1080/10408398.2020.1843396

89. Manichanh C, Rigottier-Gois L, Bonnaud E, Gloux K, Pelletier E, Frangeul L, Nalin R, Jarrin C, Chardon P, Marteau P, Roca J, and Dore J. Reduced diversity of faecal microbiota in Crohn's disease revealed by a metagenomic approach. Gut. 2006 Feb 1; 55:205–11. DOI: 10.1136/gut.2005.073817

90. Gevers D, Kugathasan S, Denson LA, Vázquez-Baeza Y, Van Treuren W, Ren B, Schwager E, Knights D, Song SJ, Yassour M, Morgan XC, Kostic AD, Luo C, González A, McDonald D, Haberman Y, Walters T, Baker S, Rosh J, Stephens M, Heyman M, Markowitz J, Baldassano R, Griffiths A, Sylvester F, Mack D, Kim S, Crandall W, Hyams J, Huttenhower C, Knight R, and Xavier RJ. The Treatment-Naive Microbiome in New-Onset Crohn's Disease. Cell Host & Microbe. 2014 Mar 12; 15:382–92. DOI: 10.1016/j.chom.2014.02.005

91. Lamas B, Richard ML, Leducq V, Pham HP, Michel ML, Da Costa G, Bridonneau C, Jegou S, Hoffmann TW, Natividad JM, Brot L, Taleb S, Couturier-Maillard A, Nion-Larmurier I, Merabtene F, Seksik P, Bourrier A, Cosnes J, Ryffel B, Beaugerie L, Launay JM, Langella P, Xavier RJ, and Sokol H. CARD9 impacts colitis by altering gut microbiota metabolism of tryptophan into aryl hydrocarbon receptor ligands. Nature Medicine. 2016 Jun; 22:598–605. DOI: 10.1038/nm.4102

92. Henao-Mejia J, Elinav E, Jin C, Hao L, Mehal WZ, Strowig T, Thaiss CA, Kau AL, Eisenbarth SC, Jurczak MJ, Camporez JP, Shulman GI, Gordon JI, Hoffman HM, and Flavell RA. Inflammasome-mediated dysbiosis regulates progression of NAFLD and obesity. Nature. 2012 Feb; 482:179–85. DOI: 10.1038/nature10809

93. Vallianou N, Stratigou T, Christodoulatos GS, and Dalamaga M. Understanding the Role of the Gut Microbiome and Microbial Metabolites in Obesity and Obesity-Associated Metabolic Disorders: Current Evidence and Perspectives. Current Obesity Reports. 2019 Sep 1; 8:317–32. DOI: 10.1007/s13679-019-00352-2

94. Lupp C, Robertson ML, Wickham ME, Sekirov I, Champion OL, Gaynor EC, and Finlay BB. Host-Mediated Inflammation Disrupts the Intestinal Microbiota and Promotes the Overgrowth of Enterobacteriaceae. Cell Host & Microbe. 2007 Aug 16; 2:119–29. DOI: 10.1016/j.chom.2007.06.010

95. Carvalho FA, Koren O, Goodrich JK, Johansson MEV, Nalbantoglu I, Aitken JD, Su Y, Chassaing B, Walters WA, González A, Clemente JC, Cullender TC, Barnich N, Darfeuille-Michaud A, Vijay-Kumar M, Knight R, Ley RE, and Gewirtz AT. Transient Inability to Manage Proteobacteria Promotes Chronic Gut Inflammation in TLR5-Deficient Mice. Cell Host & Microbe. 2012 Aug 16; 12:139–52. DOI: 10.1016/j.chom.2012.07.004

96. Huang YL, Chassard C, Hausmann M, Itzstein M von, and Hennet T. Sialic acid catabolism drives intestinal inflammation and microbial dysbiosis in mice. Nature Communications. 2015 Aug 25; 6:8141. DOI: 10.1038/ncomms9141

97. Stecher B, Robbiani R, Walker AW, Westendorf AM, Barthel M, Kremer M, Chaffron S, Macpherson AJ, Buer J, Parkhill J, Dougan G, Mering C von, and Hardt WD. Salmonella enterica Serovar Typhimurium Exploits Inflammation to Compete with the Intestinal Microbiota. PLoS Biology. 2007 Oct; 5:e244. DOI: 10.1371/journal.pbio.0050244

98. Vale A do, Cabanes D, and Sousa S. Bacterial Toxins as Pathogen Weapons Against Phagocytes. Frontiers in Microbiology 2016; 7

99. Klemm P and Schembri MA. Bacterial adhesins: function and structure. en. International Journal of Medical Microbiology 2000 Mar; 290:27–35. DOI: 10.1016/S1438-4221(00)80102-2

100. Jayamani E and Mylonakis E. Effector triggered manipulation of host immune response elicited by different pathotypes of Escherichia coli. Virulence 2014 Oct; 5. Publisher: Taylor & Francis _eprint: https://doi.org/10.4161/viru.29948:733–9. DOI: 10.4161/viru.29948

101. Starčič Erjavec M and Žgur-Bertok D. Virulence potential for extraintestinal infections among commensal Escherichia coli isolated from healthy humans—the Trojan horse within our gut. FEMS Microbiology Letters. 2015 Mar 1; 362:fnu061. DOI: 10.1093/femsle/fnu061

102. Micenková L, Bosák J, Vrba M, Ševčíková A, and Šmajs D. Human extraintestinal pathogenic Escherichia coli strains differ in prevalence of virulence factors, phylogroups, and bacteriocin determinants. BMC Microbiology. 2016 Sep 20; 16:218. DOI: 10.1186/s12866-016-0835-z

103. Winter SE, Winter MG, Xavier MN, Thiennimitr P, Poon V, Keestra AM, Laughlin RC, Gomez G, Wu J, Lawhon SD, Popova IE, Parikh SJ, Adams LG, Tsolis RM, Stewart VJ, and Bäumler AJ. Host-Derived Nitrate Boosts Growth of E. coli in the Inflamed Gut. Science. 2013 Feb 8; 339:708–11. DOI: 10.1126/science.1232467

104. Jones SA, Chowdhury FZ, Fabich AJ, Anderson A, Schreiner DM, House AL, Autieri SM, Leatham MP, Lins JJ, Jorgensen M, Cohen PS, and Conway T. Respiration of Escherichia coli in the Mouse Intestine. Infection and Immunity. 2007 Oct; 75:4891–9. DOI: 10.1128/IAI.00484-07

105. Hartman AL, Lough DM, Barupal DK, Fiehn O, Fishbein T, Zasloff M, and Eisen JA. Human gut microbiome adopts an alternative state following small bowel transplantation. Proceedings of the National Academy of Sciences. 2009 Oct 6; 106:17187–92. DOI: 10.1073/pnas.0904847106

106. Singh V, Yeoh BS, Xiao X, Kumar M, Bachman M, Borregaard N, Joe B, and Vijay-Kumar M. Interplay between enterobactin, myeloperoxidase and lipocalin 2 regulates E. coli survival in the inflamed gut. Nature Communications. 2015 May 12; 6:7113. DOI: 10.1038/ncomms8113

107. Nedialkova LP, Denzler R, Koeppel MB, Diehl M, Ring D, Wille T, Gerlach RG, and Stecher B. Inflammation Fuels Colicin Ib-Dependent Competition of Salmonella Serovar Typhimurium and E. coli in Enterobacterial Blooms. PLOS Pathogens. 2014 Jan 2; 10:e1003844. DOI: 10.1371/journal.ppat.1003844

108. Schaik W van. The human gut resistome. Philosophical Transactions of the Royal Society B: Biological Sciences. 2015 Jun 5; 370:20140087. DOI: 10.1098/rstb.2014.0087

109. Poirel L, Rodriguez-Martinez JM, Mammeri H, Liard A, and Nordmann P. Origin of Plasmid-Mediated Quinolone Resistance Determinant QnrA. Antimicrobial Agents and Chemotherapy. 2005 Aug; 49:3523–5. DOI: 10.1128/AAC.49.8.3523-3525.2005

110. Poirel L, Potron A, and Nordmann P. OXA-48-like carbapenemases: the phantom menace. Journal of Antimicrobial Chemotherapy. 2012 Jul 1; 67:1597–606. DOI: 10.1093/jac/dks121

111. Stecher B, Denzler R, Maier L, Bernet F, Sanders MJ, Pickard DJ, Barthel M, Westendorf AM, Krogfelt KA, Walker AW, Ackermann M, Dobrindt U, Thomson NR, and Hardt WD. Gut inflammation can boost horizontal gene transfer between pathogenic and commensal Enterobacteriaceae. Proceedings of the National Academy of Sciences. 2012 Jan 24; 109:1269–74. DOI: 10.1073/pnas.1113246109

112. Kent AG, Vill AC, Shi Q, Satlin MJ, and Brito IL. Widespread transfer of mobile antibiotic resistance genes within individual gut microbiomes revealed through bacterial Hi-C. Nature Communications. 2020 Sep 1; 11:4379. DOI: 10.1038/s41467-020-18164-7

113. Venturini C, Zingali T, Wyrsch ER, Bowring B, Iredell J, Partridge SR, and Djordjevic SP. Diversity of P1 phage-like elements in multidrug resistant Escherichia coli. Scientific Reports. 2019 Dec 11; 9:18861. DOI: 10.1038/s41598-019-54895-4

114. Quirós P, Colomer-Lluch M, Martínez-Castillo A, Miró E, Argente M, Jofre J, Navarro F, and Muniesa M. Antibiotic Resistance Genes in the Bacteriophage DNA Fraction of Human Fecal Samples. Antimicrobial Agents and Chemotherapy. 2014 Jan; 58:606–9. DOI: 10.1128/AAC.01684-13

115. Modi SR, Lee HH, Spina CS, and Collins JJ. Antibiotic treatment expands the resistance reservoir and ecological network of the phage metagenome. Nature. 2013 Jul; 499:219–22. DOI: 10.1038/nature12212

116. International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. Nature. 2004 Oct; 431:931–45. DOI: 10.1038/nature03001

117. Trevino V, Falciani F, and Barrera-Saldaña HA. DNA Microarrays: a Powerful Genomic Tool for Biomedical and Clinical Research. Molecular Medicine. 2007 Sep; 13:527–41. DOI: 10.2119/2006-00107.Trevino

118. Lagier JC, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, Levasseur A, Rolain JM, Fournier PE, and Raoult D. Culturing the human microbiota and culturomics. Nature Reviews Microbiology. 2018 Sep; 16:540–50. DOI: 10.1038/s41579-018-0041-0

119. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. Journal of Molecular Evolution. 1980 Dec; 16:111–20. DOI: 10.1007/BF01731581

120. Johnson JS, Spakowicz DJ, Hong BY, Petersen LM, Demkowicz P, Chen L, Leopold SR, Hanson BM, Agresta HO, Gerstein M, Sodergren E, and Weinstock GM. Evaluation of 16S rRNA gene sequencing for species and strain-level microbiome analysis. Nature Communications. 2019 Nov 6; 10:5029. DOI: 10.1038/s41467-019-13036-1

121. Treangen TJ and Salzberg SL. Repetitive DNA and next-generation sequencing: computational challenges and solutions. Nature Reviews Genetics. 2011 Nov 29; 13:36–46. DOI: 10.1038/nrg3117

122. Jain M, Olsen HE, Paten B, and Akeson M. The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community. Genome Biology. 2016 Nov 25; 17:239. DOI: 10.1186/s13059-016-1103-0

123. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, deWinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, and Turner S. Real-Time DNA Sequencing from Single Polymerase Molecules. Science. 2009 Jan 2; 323:133–8. DOI: 10.1126/science.1162986

124. Pollard MO, Gurdasani D, Mentzer AJ, Porter T, and Sandhu MS. Long reads: their purpose and place. Human Molecular Genetics. 2018 Aug 1; 27(R2):R234–R241. DOI: 10.1093/hmg/ddy177

125. De Maio N, Shaw LP, Hubbard A, George S, Sanderson ND, Swann J, Wick R, AbuOun M, Stubberfield E, Hoosdally SJ, Crook DW, Peto TEA, Sheppard AE, Bailey MJ, Read DS, Anjum MF, Walker AS, and Stoesser N. Comparison of long-read sequencing technologies in the hybrid assembly of complex bacterial genomes. Microbial Genomics. 2019 Sep 4; 5:e000294. DOI: 10.1099/mgen.0.000294

126. Wick RR, Judd LM, Gorrie CL, and Holt KE. Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Computational Biology. 2017 Jun 8; 13:e1005595. DOI: 10.1371/journal.pcbi.1005595

127. Karlin S, Mrázek J, and Campbell AM. Compositional biases of bacterial genomes and evolutionary implications. Journal of Bacteriology. 1997 Jun; 179:3899–913. DOI: 10.1128/jb.179.12.3899-3913.1997

128. Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K, Suzuki Y, Dudek N, Relman DA, Finstad KM, Amundson R, Thomas BC, and Banfield JF. A new view of the tree of life. Nature Microbiology. 2016 Apr 11; 1:1–6. DOI: 10.1038/nmicrobiol.2016.48

129. Segata N, Waldron L, Ballarini A, Narasimhan V, Jousson O, and Huttenhower C. Metagenomic microbial community profiling using unique clade-specific marker genes. Nature Methods. 2012 Aug; 9:811–4. DOI: 10.1038/nmeth.2066

130. Proctor LM, Creasy HH, Fettweis JM, Lloyd-Price J, Mahurkar A, Zhou W, Buck GA, Snyder MP, Strauss JF, Weinstock GM, White O, Huttenhower C, and The Integrative HMP (iHMP) Research Network Consortium. The Integrative Human Microbiome Project. Nature. 2019 May; 569:641–8. DOI: 10.1038/s41586-019-1238-8

131. Wang Z, Gerstein M, and Snyder M. RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics. 2009 Jan; 10:57–63. DOI: 10.1038/nrg2484

132. Lin H, He QY, Shi L, Sleeman M, Baker MS, and Nice EC. Proteomics and the microbiome: pitfalls and potential. Expert Review of Proteomics. 2019 Jun 3; 16:501–11. DOI: 10.1080/14789450.2018.1523724

133. Woerther PL, Burdet C, Chachaty E, and Andremont A. Trends in Human Fecal Carriage of Extended-Spectrum beta-Lactamases in the Community: Toward the Globalization of CTX-M. Clinical Microbiology Reviews. 2013 Oct; 26:744–58. DOI: 10.1128/CMR.00023-13

134. Östholm-Balkhed Å, Tärnberg M, Nilsson M, Nilsson LE, Hanberger H, Hällgren A, and on behalf of the Travel Study Group of Southeast Sweden. Travel-associated faecal colonization with ESBL-producing Enterobacteriaceae: incidence and risk factors. Journal of Antimicrobial Chemotherapy. 2013 Sep 1; 68:2144–53. DOI: 10.1093/jac/dkt167

135. Kantele A, Lääveri T, Mero S, Vilkman K, Pakkanen SH, Ollgren J, Antikainen J, and Kirveskari J. Antimicrobials Increase Travelers' Risk of Colonization by Extended-Spectrum Betalactamase-Producing Enterobacteriaceae. Clinical Infectious Diseases. 2015 Mar 15; 60:837–46. DOI: 10.1093/cid/ciu957

136. Ruppé E, Armand-Lefèvre L, Estellat C, Consigny PH, El Mniai A, Boussadia Y, Goujon C, Ralaimazava P, Campa P, Girard PM, Wyplosz B, Vittecoq D, Bouchaud O, Le Loup G, Pialoux G, Perrier M, Wieder I, Moussa N, Esposito-Farèse M, Hoffmann I, Coignard B, Lucet JC, Andremont A, and Matheron S. High Rate of Acquisition but Short Duration of Carriage of Multidrug-Resistant Enterobacteriaceae After Travel to the Tropics. Clinical Infectious Diseases. 2015 Aug 15; 61:593–600. DOI: 10.1093/cid/civ333

137. Struelens MJ. The epidemiology of antimicrobial resistance in hospital acquired infections: problems and possible solutions. British Medical Journal. 1998; 317:652–4. DOI: 10.1136/bmj.317.7159.652

138. Epelboin L, Robert J, Tsyrina-Kouyoumdjian E, Laouira S, Meyssonnier V, Caumes E, and Grp MGTW. High Rate of Multidrug-Resistant Gram-Negative Bacilli Carriage and Infection in Hospitalized Returning Travelers: A Cross-Sectional Cohort Study. Journal of Travel Medicine. 2015; 22:292–9. DOI: 10.1111/jtm.12211

139. Leo S, Lazarevic V, Gaia N, Estellat C, Girard M, Matheron S, Armand-Lefevre L, Andremont A, Schrenzel J, Ruppe E, Bouchaud O, Boussadia Y, Campa P, Coignard B, Consigny PH, El Mniai A, Esposito-Farese M, Girard PM, Goujon C, Hoffmann I, Le Loup G, Lucet JC, Moussa N, Perrier M, Pialoux G, Ralaimazava P, Vittecoq D, Wieder I, Wyplosz B, and Grp VRS. The intestinal microbiota predisposes to traveler's diarrhea and to the carriage of multidrug-resistant Enterobacteriaceae after traveling to tropical regions. Gut Microbes. 2019; 10:631–41. DOI: 10.1080/19490976.2018.1564431

140. Langelier C, Graves M, Kalantar K, Caldera S, Durrant R, Fisher M, Backman R, Tanner W, DeRisi JL, and Leung DT. Microbiome and Antimicrobial Resistance Gene Dynamics in International Travelers. Emerging Infectious Diseases. 2019; 25:1380–3. DOI: 10.3201/eid2507.181492

141. Sassone-Corsi M, Nuccio SP, Liu H, Hernandez D, Vu CT, Takahashi AA, Edwards RA, and Raffatellu M. Microcins mediate competition among Enterobacteriaceae in the inflamed gut. Nature. 2016; 540:280–+. DOI: 10.1038/nature20557

142. Kitamoto S, Alteri CJ, Rodrigues M, Nagao-Kitamoto H, Sugihara K, Himpsl SD, Bazzi M, Miyoshi M, Nishioka T, Hayashi A, Morhardt TL, Kuffa P, Grasberger H, El-Zaatari M, Bishu S, Ishii C, Hirayama A, Eaton KA, Dogan B, Simpson KW, Inohara N, Mobley HLT, Kao JY, Fukuda S, Barnich N, and Kamada N. Dietary L-serine confers a competitive fitness advantage to Enterobacteriaceae in the inflamed gut. Nature Microbiology. 2020; 5:116–+. DOI: 10.1038/s41564-019-0591-6

143. Mullineaux-Sanders C, Carson D, Hopkins EGD, Glegola-Madejska I, Escobar-Zepeda A, Browne HP, Lawley TD, and Frankel G. Citrobacter amalonaticus Inhibits the Growth of Citrobacter rodentium in the Gut Lumen. mBio. 2021 Oct 26; 12:e02410–21. DOI: 10.1128/mBio.02410-21

144. Escobar-Páramo P, Grenet K, Le Menac'h A, Rode L, Salgado E, Amorin C, Gouriou S, Picard B, Rahimy MC, Andremont A, Denamur E, and Ruimy R. Large-Scale Population Structure of Human Commensal Escherichia coli Isolates. Applied and Environmental Microbiology. 2004 Sep; 70:5698–700. DOI: 10.1128/AEM.70.9.5698-5700.2004

145. Micenková L, Bosák J, Štaudová B, Kohoutová D, Čejková D, Woznicová V, Vrba M, Ševčíková A, Bureš J, and Šmajs D. Microcin determinants are associated with B2 phylogroup of human fecal Escherichia coli isolates. MicrobiologyOpen. 2016 Mar 14; 5:490–8. DOI: 10.1002/mbo3.345

146. UNWTO W. Tourism highlights. Retrieved September. 2014

147. Ruppe E, Armand-Lefevre L, Estellat C, El-Mniai A, Boussadia Y, Consigny PH, Girard PM, Vittecoq D, Bouchaud O, Pialoux G, Esposito-Farese M, Coignard B, Lucet JC, Andremont A, and Matheron S. Acquisition of carbapenemase-producing Enterobacteriaceae by healthy travellers to India, France, February 2012 to March 2013. Eurosurveillance. 2014; 19:2–5. DOI: 10.2807/1560-7917.es2014.19.14.20768

148. Kuenzli E, Jaeger VK, Frei R, Neumayr A, DeCrom S, Haller S, Blum J, Widmer AF, Furrer H, Battegay M, Endimiani A, and Hatz C. High colonization rates of extended-spectrum beta-lactamase (ESBL)-producing Escherichia coli in Swiss Travellers to South Asia- a prospective observational multicentre cohort study looking at epidemiology, microbiology and risk factors. Bmc Infectious Diseases. 2014; 14:10. DOI: 10.1186/1471-2334-14-528

149. Frost I, Van Boeckel TP, Pires J, Craig J, and Laxminarayan R. Global geographic trends in antimicrobial resistance: the role of international travel. Journal of Travel Medicine. 2019; 26:13. DOI: 10.1093/jtm/taz036

150. Huang YS, Lai LC, Chen YA, Lin KY, Chou YH, Chen HC, Wang SS, Wang JT, and Chang SC. Colonization With Multidrug-Resistant Organisms Among Healthy Adults in the Community Setting: Prevalence, Risk Factors, and Composition of Gut Microbiome. Frontiers in Microbiology. 2020 Jun 24; 11:1402. DOI: 10.3389/fmicb.2020.01402

151. Baquero F. Metagenomic epidemiology: a public health need for the control of antimicrobial resistance. Clinical Microbiology and Infection. 2012; 18:67–73. DOI: 10.1111/j.1469-0691.2012.03860.x

152. Paterson DL. Resistance in gram-negative bacteria: Enterobacteriaceae. American Journal of Infection Control. 2006; 34:S20–S28. DOI: 10.1016/j.ajic.2006.05.238

153. Haverkate MR, Platteel TN, Fluit AC, Stuart JWC, Hall M, Thijsen SFT, Scharringa J, Kloosterman RC, Bonten MJM, and Bootsma MCJ. Quantifying within-household transmission of extended-spectrum beta-lactamase-producing bacteria. Clinical Microbiology and Infection. 2017; 23. DOI: 10.1016/j.cmi.2016.08.021

154. Kantele A, Mero S, Kirveskari J, and Laaveri T. Increased Risk for ESBL-Producing Bacteria from Co-administration of Loperamide and Antimicrobial Drugs for Travelers' Diarrhea. Emerging Infectious Diseases. 2016; 22:117–20. DOI: 10.3201/eid2201.151272

155. Kantele A, Kuenzli E, Dunn SJ, Dance DAB, Newton PN, Davong V, Mero S, Pakkanen SH, Neumayr A, Hatz C, Snaith A, Kallonen T, Corander J, and McNally A. Dynamics of intestinal multidrug-resistant bacteria colonisation contracted by visitors to a high-endemic setting: a prospective, daily, real-time sampling study. The Lancet Microbe. 2021 Apr 1; 2:e151–e158. DOI: 10.1016/S2666-5247(20)30224-X

156. Rea MC, Sit CS, Clayton E, O'Connor PM, Whittal RM, Zheng J, Vederas JC, Ross RP, and Hill C. Thuricin CD, a posttranslationally modified bacteriocin with a narrow spectrum of activity against Clostridium difficile. Proceedings of the National Academy of Sciences of the United States of America. 2010 May 18; 107:9352–7. DOI: 10.1073/pnas.0913554107

157. Rea MC, Dobson A, O'Sullivan O, Crispie F, Fouhy F, Cotter PD, Shanahan F, Kiely B, Hill C, and Ross RP. Effect of broad- and narrow-spectrum antimicrobials on Clostridium difficile and microbial diversity in a model of the distal colon. Proceedings of the National Academy of Sciences of the United States of America. 2011; 108:4639–44. DOI: 10.1073/pnas.1001224107

158. Ducarmon QR, Zwittink RD, Hornung BVH, Schaik W van, Young VB, and Kuijper EJ. Gut Microbiota and Colonization Resistance against Bacterial Enteric Infection. Microbiology and Molecular Biology Reviews. 2019; 83. DOI: 10.1128/mmbr.00007-19

159. Canani RB, Costanzo MD, Leone L, Pedata M, Meli R, and Calignano A. Potential beneficial effects of butyrate in intestinal and extraintestinal diseases. World Journal of Gastroenterology : WJG 2011 Mar; 17:1519–28. DOI: 10.3748/wjg.v17.i12.1519

160. Hosseini E, Grootaert C, Verstraete W, and Van de Wiele T. Propionate as a health-promoting microbial metabolite in the human gut. Nutrition Reviews 2011 May; 69:245–58. DOI: 10.1111/j.1753-4887.2011.00388.x

161. Nogal A, Louca P, Zhang X, Wells PM, Steves CJ, Spector TD, Falchi M, Valdes AM, and Menni C. Circulating Levels of the Short-Chain Fatty Acid Acetate Mediate the Effect of the Gut Microbiome on Visceral Fat. Frontiers in Microbiology 2021; 12

162. Cash HL and Hooper LV. Commensal bacteria shape intestinal immune system development. Asm News. 2005; 71:77–+

163. Arcilla MS, Hattem JM van, Bootsma MCJ, Genderen PJ van, Goorhuis A, Schultsz C, Stobberingh EE, Verbrugh HA, Jong MD de, Melles DC, and Penders J. The Carriage Of Multiresistant Bacteria After Travel (COMBAT) prospective cohort study: methodology and design. Bmc Public Health. 2014; 14. DOI: 10.1186/1471-2458-14-410

164. NVMM Guideline Laboratory detection of highly resistant microorganisms, version 2.0. 2012

165. Costea PI, Zeller G, Sunagawa S, Pelletier E, Alberti A, Levenez F, Tramontano M, Driessen M, Hercog R, Jung FE, Kultima JR, Hayward MR, Coelho LP, Allen-Vercoe E, Bertrand L, Blaut M, Brown JRM, Carton T, Cools-Portier S, Daigneault M, Derrien M, Druesne A, Vos WM de, Finlay BB, Flint HJ, Guarner F, Hattori M, Heilig H, Luna RA, Vlieg JVH, Junick J, Klymiuk I, Langella P, Le Chatelier E, Mai V, Manichanh C, Martin JC, Mery C, Morita H, O'Toole PW, Orvain C, Patil KR, Penders J, Persson S, Pons N, Popova M, Salonen A, Saulnier D, Scott KP, Singh B, Slezak K, Veiga P, Versalovic J, Zhao LP, Zoetendal EG, Ehrlich SD, Dore J, and Bork P. Towards standards for human fecal sample processing in metagenomic studies. Nature Biotechnology. 2017; 35:1069–+. DOI: 10.1038/nbt.3960

166. Baym M, Kryazhimskiy S, Lieberman TD, Chung H, Desai MM, and Kishony R. Inexpensive Multiplexed Library Preparation for Megabase-Sized Genomes (vol 10, e0128036, 2015). Plos One. 2015; 10. DOI: 10.1371/journal.pone.0131262

167. Kultima JR, Coelho LP, Forslund K, Huerta-Cepas J, Li SS, Driessen M, Voigt AY, Zeller G, Sunagawa S, and Bork P. MOCAT2: a metagenomic assembly, annotation and profiling framework. Bioinformatics. 2016; 32:2520–3. DOI: 10.1093/bioinformatics/btw183

168. Hannon. FASTX-Toolkit. 2010

169. Li RQ, Li YR, Kristiansen K, and Wang J. SOAP: short oligonucleotide alignment program. Bioinformatics. 2008; 24:713–4. DOI: 10.1093/bioinformatics/btn025

170. Zhu WH, Lomsadze A, and Borodovsky M. Ab initio gene identification in metagenomic sequences. Nucleic Acids Research. 2010; 38. DOI: 10.1093/nar/gkq275

171. Nielsen HB, Almeida M, Juncker AS, Rasmussen S, Li JH, Sunagawa S, Plichta DR, Gautier L, Pedersen AG, Le Chatelier E, Pelletier E, Bonde I, Nielsen T, Manichanh C, Arumugam M, Batto JM, Santos MBQ dos, Blom N, Borruel N, Burgdorf KS, Boumezbeur F, Casellas F, Dore J, Dworzynski P, Guarner F, Hansen T, Hildebrand F, Kaas RS, Kennedy S, Kristiansen K, Kultima JR, Leonard P, Levenez F, Lund O, Moumen B, Le Paslier D, Pons N, Pedersen O, Prifti E, Qin

JJ, Raes J, Sorensen S, Tap J, Tims S, Ussery DW, Yamada T, Renault P, Sicheritz-Ponten T, Bork P, Wang J, Brunak S, Ehrlich SD, and Meta HITC. Identification and assembly of genomes and genetic elements in complex metagenomic samples without using reference genomes. Nature Biotechnology. 2014; 32:822–8. DOI: 10.1038/nbt.2939

172. Altschul SF, Gish W, Miller W, Myers EW, and Lipman DJ. BASIC LOCAL ALIGNMENT SEARCH TOOL. Journal of Molecular Biology. 1990; 215:403–10. DOI: 10.1016/s0022-2836(05)80360-2

173. Madeira F, Park YM, Lee J, Buso N, Gur T, Madhusoodanan N, Basutkar P, Tivey ARN, Potter SC, Finn RD, and Lopez R. The EMBL-EBI search and sequence analysis tools APIs in 2019. Nucleic Acids Research. 2019; 47(W1):W636–W641. DOI: 10.1093/nar/gkz268

174. Huerta-Cepas J, Szklarczyk D, Heller D, Hernandez-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, Mering C von, and Bork P. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Research. 2019; 47(D1):D309–D314. DOI: 10.1093/nar/gky1085

175. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Mering C von, and Bork P. Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. Molecular Biology and Evolution. 2017; 34:2115–22. DOI: 10.1093/molbev/msx148

176. Dixon P. VEGAN, a package of R functions for community ecology. Journal of Vegetation Science. 2003; 14:927–30. DOI: 10.1658/1100-9233(2003)014[0927:Vaporf]2.0.Co;2

177. Chen EZ and Li HZ. A two-part mixed-effects model for analyzing longitudinal microbiome compositional data. Bioinformatics. 2016; 32:2611–7. DOI: 10.1093/bioinformatics/btw308

178. W.N Venables BR. Modern Applied Statistics with S. New York: Springer, 2002

179. Kaminski J, Gibson MK, Franzosa EA, Segata N, Dantas G, and Huttenhower C. High-Specificity Targeted Functional Profiling in Microbial Communities with ShortBRED. PLOS Computational Biology. 2015 Dec 18; 11:e1004557. DOI: 10.1371/journal.pcbi.1004557

180. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, and Larsen MV. Identification of acquired antimicrobial resistance genes. The Journal of Antimicrobial Chemotherapy. 2012 Nov; 67:2640–4. DOI: 10.1093/jac/dks261

181. The UniProt Consortium. UniProt: the universal protein knowledgebase in 2021. Nucleic Acids Research. 2021 Jan 8; 49(D1):D480–D489. DOI: 10.1093/nar/gkaa1100

182. D'Argenio V and Salvatore F. The role of the gut microbiome in the healthy adult status. Clinica Chimica Acta. 2015; 451:97–102. DOI: 10.1016/j.cca.2015.01.003

183. Litvak Y, Mon KKZ, Nguyen H, Chanthavixay G, Liou M, Velazquez EM, Kutter L, Alcantara MA, Byndloss MX, Tiffany CR, Walker GT, Faber F, Zhu YH, Bronner DN, Byndloss AJ, Tsolis RM, Zhou HJ, and Baumler AJ. Commensal Enterobacteriaceae Protect against Salmonella Colonization through Oxygen Competition. Cell Host & Microbe. 2019; 25:128–+. DOI: 10.1016/j.chom.2018.12.003

184. Velazquez EM, Nguyen H, Heasley KT, Saechao CH, Gil LM, Rogers AWL, Miller BM, Rolston MR, Lopez CA, Litvak Y, Liou MJ, Faber F, Bronner DN, Tiffany CR, Byndloss MX, Byndloss AJ, and Baumler AJ. Endogenous Enterobacteriaceae underlie variation in susceptibility to Salmonella infection. Nature Microbiology. 2019; 4:1057–64. DOI: 10.1038/s41564-019-0407-8

185. Huttenhower C et al. Structure, function and diversity of the healthy human microbiome. Nature. 2012; 486:207–14. DOI: 10.1038/nature11234

186. Piewngam P, Quinones M, Thirakittiwatthana W, Yungyuen T, Otto M, and Kiratisin P. Composition of the intestinal microbiota in extended-spectrum beta-lactamase-producing Enterobacteriaceae carriers and non-carriers in Thailand. International Journal of Antimicrobial Agents. 2019; 53:435–41. DOI: 10.1016/j.ijantimicag.2018.12.006

187. Brown EM, Ke XB, Hitchcock D, Jeanfavre S, Avila-Pacheco J, Nakata T, Arthur TD, Fornelos N, Heim C, Franzosa EA, Watson N, Huttenhower C, Haiser HJ, Dillow G, Graham DB, Finlay BB, Kostic AD, Porter JA, Vlamakis H, Clish CB, and Xavier RJ. Bacteroides-Derived Sphingolipids Are Critical for Maintaining Intestinal Homeostasis and Symbiosis. Cell Host & Microbe. 2019; 25:668–+. DOI: 10.1016/j.chom.2019.04.002

188. Shen Y, Torchia MLG, Lawson GW, Karp CL, Ashwell JD, and Mazmanian SK. Outer Membrane Vesicles of a Human Commensal Mediate Immune Regulation and Disease Protection. Cell Host & Microbe. 2012; 12:509–20. DOI: 10.1016/j.chom.2012.08.004

189. Round JL and Mazmanian SK. Inducible Foxp(3+) regulatory T-cell development by a commensal bacterium of the intestinal microbiota. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:12204–9. DOI: 10.1073/pnas.0909122107

190. Spees AM, Wangdi T, Lopez CA, Kingsbury DD, Xavier MN, Winter SE, Tsolis RM, and Baumler AJ. Streptomycin-Induced Inflammation Enhances Escherichia coli Gut Colonization Through Nitrate Respiration. Mbio. 2013; 4. DOI: 10.1128/mBio.00430-13

191. Cano PG, Santacruz A, Moya A, and Sanz Y. Bacteroides uniformis CECT 7771 Ameliorates Metabolic and Immunological Dysfunction in Mice with High-Fat-Diet Induced Obesity. Plos One. 2012; 7. DOI: 10.1371/journal.pone.0041079

192. Despres J, Forano E, Lepercq P, Comtet-Marre S, Jubelin G, Chambon C, Yeoman CJ, Miller MEB, Fields CJ, Martens E, Terrapon N, Henrissat B, White BA, and Mosoni P. Xylan degradation by the human gut Bacteroides xylanisolvens XB1A(T) involves two distinct gene clusters that are linked at the transcriptional level. Bmc Genomics. 2016; 17. DOI: 10.1186/s12864-016-2680-8

193. Mendis M and Simsek S. Arabinoxylans and human health. Food Hydrocolloids. 2014; 42:239–43. DOI: 10.1016/j.foodhyd.2013.07.022

194. Hromadkova Z, Paulsen BS, Polovka M, Kost'alova Z, and Ebringerova A. Structural features of two heteroxylan polysaccharide fractions from wheat bran with anti-complementary and antioxidant activities. Carbohydrate Polymers. 2013; 93:22–30. DOI: 10.1016/j.carbpol.2012.05.021

195. Sorbara MT, Dubin K, Littmann ER, Moody TU, Fontana E, Seok R, Leiner IM, Taur Y, Peled JU, Brink MRM van den, Litvak Y, Baumler AJ, Chaubard JL, Pickard AJ, Cross JR, and Pamer EG. Inhibiting antibiotic-resistant Enterobacteriaceae by microbiota-mediated intracellular acidification. Journal of Experimental Medicine. 2019; 216:84–98. DOI: 10.1084/jem.20181639

196. D'Souza AW, Boolchandani M, Patel S, Galazzo G, Hattem JM van, Arcilla MS, Melles DC, Jong MD de, Schultsz C, Bootsma MCJ, Genderen PJ van, Goorhuis A, Grobusch MP, Molhoek N, Oude Lashof AML, Stobberingh EE, Verbrugh HA, Dantas G, Penders J, and COMBAT Consortium. Destination shapes antibiotic resistance gene acquisitions, abundance increases, and diversity changes in Dutch travelers. Genome Medicine. 2021 Jun 7; 13:79. DOI: 10.1186/s13073-021-00893-z

197. Tedjo DI, Jonkers D, Savelkoul PH, Masclee AA, Best N van, Pierik MJ, and Penders J. The Effect of Sampling and Storage on the Fecal Microbiota Composition in Healthy and Diseased Subjects. Plos One. 2015; 10. DOI: 10.1371/journal.pone.0126685

198. Amir A, McDonald D, Navas-Molina JA, Debelius J, Morton JT, Hyde E, Robbins-Pianka A, and Knight R. Correcting for Microbial Blooms in Fecal Samples during Room-Temperature Shipping. Msystems. 2017; 2. DOI: 10.1128/mSystems.00199-16

199. Antonara S and Ardura MI. 141 - Citrobacter Species. *Principles and Practice of Pediatric Infectious Diseases (Fifth Edition)*. Ed. by Long SS, Prober CG, and Fischer M. Elsevier, 2018 Jan 1:827–829.e1. DOI: 10.1016/B978-0-323-40181-4.00141-9

200. Schauer DB and Falkow S. Attaching and effacing locus of a Citrobacter freundii biotype that causes transmissible murine colonic hyperplasia. Infection and Immunity. 1993 Jun; 61:2486–92. DOI: 10.1128/iai.61.6.2486-2492.1993

201. Bolger AM, Lohse M, and Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014 Aug 1; 30:2114–20. DOI: 10.1093/bioinformatics/btu170

202. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, and Pevzner PA. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. Journal of Computational Biology. 2012 May; 19:455–77. DOI: 10.1089/cmb.2012.0021

203. Jolley KA, Bray JE, and Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Research. 2018; 3:124. DOI: 10.12688/wellcomeopenres.14826.1

204. Arredondo-Alonso S, Rogers MRC, Braat JC, Verschuuren TD, Top J, Corander J, Willems RJL, and Schürch AC. mlplasmids: a user-friendly tool to predict plasmid- and chromosome-derived sequences for single species. Microbial Genomics. 2018 Nov; 4. DOI: 10.1099/mgen.0.000224

205. O'Leary NA, Wright MW, Brister JR, Ciufo S, Haddad D, McVeigh R, Rajput B, Robbertse B, Smith-White B, Ako-Adjei D, Astashyn A, Badretdin A, Bao Y, Blinkova O, Brover V, Chetvernin V, Choi J, Cox E, Ermolaeva O, Farrell CM, Goldfarb T, Gupta T, Haft D, Hatcher E, Hlavina W, Joardar VS, Kodali VK, Li W, Maglott D, Masterson P, McGarvey KM, Murphy MR, O'Neill K, Pujar S, Rangwala SH, Rausch D, Riddick LD, Schoch C, Shkeda A, Storz SS,

Sun H, Thibaud-Nissen F, Tolstoy I, Tully RE, Vatsan AR, Wallin C, Webb D, Wu W, Landrum MJ, Kimchi A, Tatusova T, DiCuccio M, Kitts P, Murphy TD, and Pruitt KD. Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. Nucleic Acids Research. 2016 Jan 4; 44(D1):D733–745. DOI: 10.1093/nar/gkv1189

206. Wood DE, Lu J, and Langmead B. Improved metagenomic analysis with Kraken 2. Genome Biology. 2019 Nov 28; 20:257. DOI: 10.1186/s13059-019-1891-0

207. Katz LS, Griswold T, Morrison SS, Caravas JA, Zhang S, Bakker HCd, Deng X, and Carleton HA. Mashtree: a rapid comparison of whole genome sequence files. Journal of Open Source Software. 2019 Dec 10; 4:1762. DOI: 10.21105/joss.01762

208. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014 Jul 15; 30:2068–9. DOI: 10.1093/bioinformatics/btu153

209. Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, Lees JA, Gladstone RA, Lo S, Beaudoin C, Floto RA, Frost SD, Corander J, Bentley SD, and Parkhill J. Producing polished prokaryotic pangenomes with the Panaroo pipeline. Genome Biology. 2020 Jul 22; 21:180. DOI: 10.1186/s13059-020-02090-4

210. Nguyen LT, Schmidt HA, Haeseler A von, and Minh BQ. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Molecular Biology and Evolution. 2015 Jan 1; 32:268–74. DOI: 10.1093/molbev/msu300

211. Miles AA, Misra SS, and Irwin JO. The estimation of the bactericidal power of the blood. The Journal of Hygiene. 1938 Nov; 38:732–49

212. Jung GY, Kim JR, Park JY, and Park S. Hydrogen production by a new chemoheterotrophic bacterium Citrobacter sp. Y19. International Journal of Hydrogen Energy. 2002 Jun 1; 27:601–10. DOI: 10.1016/S0360-3199(01)00176-8

213. Sauvageot N, Gouffi K, Laplace JM, and Auffray Y. Glycerol metabolism in Lactobacillus collinoides: production of 3-hydroxypropionaldehyde, a precursor of acrolein. International Journal of Food Microbiology. 2000 Apr 10; 55:167–70. DOI: 10.1016/S0168-1605(00)00191-4

214. Homann T, Tag C, Biebl H, Deckwer WD, and Schink B. Fermentation of glycerol to 1,3-propanediol by Klebsiella and Citrobacter strains. Applied Microbiology and Biotechnology. 1990 May 1; 33:121–6. DOI: 10.1007/BF00176511

215. Vimont A, Fernandez B, Ahmed G, Fortin HP, and Fliss I. Quantitative antifungal activity of reuterin against food isolates of yeasts and moulds and its potential application in yogurt. International Journal of Food Microbiology. 2019 Jan 16; 289:182–8. DOI: 10.1016/j.ijfoodmicro.2018.09.005

216. Cleusix V, Lacroix C, Vollenweider S, and Le Blay G. Glycerol induces reuterin production and decreases Escherichia coli population in an in vitro model of colonic fermentation with immobilized human feces. FEMS Microbiology Ecology. 2008 Jan 1; 63:56–64. DOI: 10.1111/j.1574-6941.2007.00412.x

217. San Millan A and MacLean RC. Fitness Costs of Plasmids: a Limit to Plasmid Transmission. Microbiology Spectrum. 2017 Sep 7; 5:5.5.02. DOI: 10.1128/microbiolspec.MTBP-0016-2017

218. Tenaillon O, Skurnik D, Picard B, and Denamur E. The population genetics of commensal Escherichia coli. Nature Reviews Microbiology. 2010 Mar; 8:207–17. DOI: 10.1038/nrmicro2298

219. Denamur E, Clermont O, Bonacorsi S, and Gordon D. The population genetics of pathogenic Escherichia coli. Nature Reviews Microbiology. 2021 Jan; 19:37–54. DOI: 10.1038/s41579-020-0416-x

220. Bailey JK, Pinyon JL, Anantham S, and Hall RM2. Commensal Escherichia coli of healthy humans: a reservoir for antibiotic-resistance determinants. Journal of Medical Microbiology. 59:1331–9. DOI: 10.1099/jmm.0.022475-0

221. Anthony WE, Burnham CAD, Dantas G, and Kwon JH. The Gut Microbiome as a Reservoir for Antimicrobial Resistance. The Journal of Infectious Diseases. 2021 Jun 15; 223(Supplement_3):S209–S213. DOI: 10.1093/infdis/jiaa497

222. Thanh Duy P, Thi Nguyen TN, Vu Thuy D, Chung The H, Alcock F, Boinett C, Dan Thanh HN, Thanh Tuyen H, Thwaites GE, Rabaa MA, and Baker S. Commensal Escherichia coli are a reservoir for the transfer of XDR plasmids into epidemic fluoroquinolone-resistant Shigella sonnei. Nature Microbiology. 2020 Feb; 5:256–64. DOI: 10.1038/s41564-019-0645-9

223. Massot M, Daubié AS, Clermont O, Jaureguy F, Couffignal C, Dahbi G, Mora A, Blanco J, Branger C, Mentré F, Eddi A, Picard B, and Denamur E. Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of Escherichia coli from community subjects in the Paris area in 2010 and evolution over 30 years. Microbiology. 2016; 162:642–50. DOI: 10.1099/mic.0.000242

224. World Health Organization. WHO preferred product characteristics for vaccines against enterotoxigenic Escherichia coli. Geneva: World Health Organization, 2021

225. Mentzer A von, Connor TR, Wieler LH, Semmler T, Iguchi A, Thomson NR, Rasko DA, Joffre E, Corander J, Pickard D, Wiklund G, Svennerholm AM, Sjöling Å, and Dougan G. Identification of enterotoxigenic Escherichia coli (ETEC) clades with long-term global distribution. Nature Genetics. 2014 Dec; 46:1321–6. DOI: 10.1038/ng.3145

226. Valenza G, Nickel S, Pfeifer Y, Eller C, Krupa E, Lehner-Reindl V, and Höller C. Extended-Spectrum-beta-Lactamase-Producing Escherichia coli as Intestinal Colonizers in the German Community. Antimicrobial Agents and Chemotherapy. 2014 Feb; 58:1228–30. DOI: 10.1128/AAC.01993-13

227. Overdevest ITMA, Bergmans AMC, Verweij JJ, Vissers J, Bax N, Snelders E, and Kluytmans JAJW. Prevalence of phylogroups and O25/ST131 in susceptible and extended-spectrum beta-lactamase-producing Escherichia coli isolates, the Netherlands. Clinical Microbiology and Infection. 2015 Jun 1; 21:570.e1–570.e4. DOI: 10.1016/j.cmi.2015.02.020

228. Marin J, Clermont O, Royer G, Mercier-Darty M, Decousser JW, Tenaillon O, Denamur E, and Blanquart F. The Population Genomics of Increased Virulence and Antibiotic Resistance in Human Commensal Escherichia coli over 30 Years in France. Applied and environmental microbiology. 2022 Aug 9; 88. Place: United States:e0066422. DOI: 10.1128/aem.00664-22

229. Bailey JK, Pinyon JL, Anantham S, and Hall RM. Distribution of Human Commensal Escherichia coli Phylogenetic Groups. Journal of Clinical Microbiology. 2010 Sep; 48:3455–6. DOI: 10.1128/JCM.00760-10

230. Nurk S, Meleshko D, Korobeynikov A, and Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Research. 2017 May; 27:824–34. DOI: 10.1101/gr.213959.116

231. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, and Wang Z. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. PeerJ. 2019 Jul 26; 7:e7359. DOI: 10.7717/peerj.7359

232. Olm MR, Brown CT, Brooks B, and Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. The ISME Journal. 2017 Dec; 11:2864–8. DOI: 10.1038/ismej.2017.126

233. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, and Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Research. 2015 Jul; 25:1043–55. DOI: 10.1101/gr.186072.114

234. Hyatt D, Chen GL, LoCascio PF, Land ML, Larimer FW, and Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010 Mar 8; 11:119. DOI: 10.1186/1471-2105-11-119

235. Buchfink B, Reuter K, and Drost HG. Sensitive protein alignments at tree-of-life scale using DIAMOND. Nature Methods. 2021 Apr; 18:366–8. DOI: 10.1038/s41592-021-01101-x

236. Zhou Z, Alikhan NF, Mohamed K, Fan Y, Group tAS, Achtman M, Brown D, Chattaway M, Dallman T, Delahay R, Kornschober C, Pietzka A, Malorny B, Petrovska L, Davies R, Robertson A, Tyne W, Weill FX, Accou-Demartin M, and Williams N. The EnteroBase user's guide, with case studies on Salmonella transmissions, Yersinia pestis phylogeny, and Escherichia core genomic diversity. Genome Research. 2020 Jan 1; 30:138–52. DOI: 10.1101/gr.251678.119

237. Connor C. EnterobaseGenomeAssemblyDownload. 2018

238. Beghain J, Bridier-Nahmias A, Le Nagard H, Denamur E, and Clermont O. ClermonTyping: an easy-to-use and accurate in silico method for Escherichia genus strain phylotyping. Microbial Genomics. 2018 Jun 19; 4:e000192. DOI: 10.1099/mgen.0.000192

239. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, and Jin Q. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Research. 2005 Jan 1; 33(Database issue):D325–328. DOI: 10.1093/nar/gki008

240. Mann HB and Whitney DR. On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. The Annals of Mathematical Statistics. 1947 Mar; 18:50–60. DOI: 10.1214/aoms/1177730491

241. Hothorn T, Bretz F, and Westfall P. Simultaneous inference in general parametric models. Biometrical Journal Biometrische Zeitschrift. 2008 Jun; 50:346–63. DOI: 10.1002/bimj.200810425

242. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. Scandinavian Journal of Statistics. 1979; 6:65–70

243. Gu Z, Gu L, Eils R, Schlesner M, and Brors B. circlize Implements and enhances circular visualization in R. Bioinformatics (Oxford, England). 2014 Oct; 30:2811–2. DOI: 10.1093/bioinformatics/btu393

244. Picard B, Garcia JS, Gouriou S, Duriez P, Brahimi N, Bingen E, Elion J, and Denamur E. The Link between Phylogeny and Virulence inEscherichia coli Extraintestinal Infection. Infection and Immunity. 1999 Feb; 67:546–53. DOI: 10.1128/IAI.67.2.546-553.1999

245. Manges AR, Geum HM, Guo A, Edens TJ, Fibke CD, and Pitout JDD. Global Extraintestinal Pathogenic Escherichia coli (ExPEC) Lineages. Clinical Microbiology Reviews. 2019 Jun 19; 32:e00135–18. DOI: 10.1128/CMR.00135-18

246. Chakraborty A, Saralaya V, Adhikari P, Shenoy S, Baliga S, and Hegde A. Characterization of Escherichia coli Phylogenetic Groups Associated with Extraintestinal Infections in South Indian Population. Annals of Medical and Health Sciences Research. 2015; 5:241–6. DOI: 10.4103/2141-9248.160192

247. Sannes MR, Kuskowski MA, Owens K, Gajewski A, and Johnson JR. Virulence Factor Profiles and Phylogenetic Background of Escherichia coli Isolates from Veterans with Bacteremia and Uninfected Control Subjects. The Journal of Infectious Diseases. 2004 Dec 15; 190:2121–8. DOI: 10.1086/425984

248. Clermont O, Condamine B, Dion S, Gordon DM, and Denamur E. The E phylogroup of Escherichia coli is highly diverse and mimics the whole E. coli species population structure. Environmental Microbiology. 2021; 23:7139–51. DOI: 10.1111/1462-2920.15742

249. Pires J, Kraemer JG, Kuenzli E, Kasraian S, Tinguely R, Hatz C, Endimiani A, and Hilty M. Gut microbiota dynamics in travelers returning from India colonized with extended-spectrum cephalosporin-resistant Enterobacteriaceae: A longitudinal study. en. Travel Medicine and Infectious Disease 2019 Jan; 27:72–80. DOI: 10.1016/j.tmaid.2018.10.012

250. Schmidt TSB, Li SS, Maistrenko OM, Akanni W, Coelho LP, Dolai S, Fullam A, Glazek AM, Hercog R, Herrema H, Jung F, Kandels S, Orakov A, Thielemann R, Stetten M von, Van Rossum T, Benes V, Borody TJ, Vos WM de, Ponsioen CY, Nieuwdorp M, and Bork P. Drivers and determinants of strain dynamics following fecal microbiota transplantation. en. Nature Medicine 2022 Sep; 28. Number: 9 Publisher: Nature Publishing Group:1902–12. DOI: 10.1038/s41591-022-01913-0

251. Leo S, Lazarevic V, Girard M, Gaïa N, Schrenzel J, Lastours V de, Fantin B, Bonten M, Carmeli Y, Rondinaud E, Harbarth S, and Huttner BD. Metagenomic Characterization of Gut Microbiota of Carriers of Extended-Spectrum Beta-Lactamase or Carbapenemase-Producing Enterobacteriaceae Following Treatment with Oral Antibiotics and Fecal Microbiota Transplantation: Results from a Multicenter Randomized Trial. Microorganisms 2020 Jun; 8:941. DOI: 10.3390/microorganisms8060941

252. Davies M, Galazzo G, Hattem JM van, Arcilla MS, Melles DC, Jong MD de, Schultsz C, Wolffs P, McNally A, Schaik Wv, and Penders J. Enterobacteriaceae and Bacteroidaceae provide resistance to travel-associated intestinal colonization by multi-drug resistant Escherichia coli. Gut Microbes. 2022 Dec 31; 14:2060676. DOI: 10.1080/19490976.2022.2060676

253. Kampmann C, Dicksved J, Engstrand L, and Rautelin H. Changes to human faecal microbiota after international travel. en. Travel Medicine and Infectious Disease 2021 Nov; 44:102199. DOI: 10.1016/j.tmaid.2021.102199

254. Nussberger AM, Luo L, Celis LE, and Crockett MJ. Public attitudes value interpretability but prioritize accuracy in Artificial Intelligence. Nature Communications. 2022 Oct 3; 13:1–13. DOI: 10.1038/s41467-022-33417-3

255. Building a BLAST database with your (local) sequences. National Center for Biotechnology Information (US), 2021 Jan 7

256. Lundberg S and Lee SI. A Unified Approach to Interpreting Model Predictions. 2017 Nov 24. DOI: 10.48550/arXiv.1705.07874. arXiv: 1705.07874[cs,stat]

257. Chen T and Guestrin C. XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '16. event-place: New York, NY, USA. Association for Computing Machinery, 2016 Aug 13:785–94. DOI: 10.1145/2939672.2939785

258. Cavallaro M, Moran E, Collyer B, McCarthy ND, Green C, and Keeling MJ. Informing antimicrobial stewardship with explainable AI. 2022 Aug 12. DOI: 10.1101/2022.08.12.22278678

259. Moran E, Robinson E, Green C, Keeling M, and Collyer B. Towards personalized guidelines: using machine-learning algorithms to guide antimicrobial selection. Journal of Antimicrobial Chemotherapy. 2020 Sep; 75:2677–80. DOI: 10.1093/jac/dkaa222

260. Weisenberg SA, Mediavilla JR, Chen L, Alexander EL, Rhee KY, Kreiswirth BN, and Jenkins SG. Extended Spectrum Beta-Lactamase-Producing Enterobacteriaceae in International Travelers and Non-Travelers in New York City. Plos One. 2012; 7:4. DOI: 10.1371/journal.pone.0045141

261. Martinson JNV, Pinkham NV, Peters GW, Cho H, Heng J, Rauch M, Broadaway SC, and Walk ST. Rethinking gut microbiome residency and the Enterobacteriaceae in healthy human adults. Isme Journal. 2019; 13:2306–18. DOI: 10.1038/s41396-019-0435-7

262. Korach-Rechtman H, Hreish M, Fried C, Gerassy-Vainberg S, Azzam ZS, Kashi Y, and Berger G. Intestinal Dysbiosis in Carriers of Carbapenem-Resistant Enterobacteriaceae. Msphere. 2020; 5. DOI: 10.1128/mSphere.00173-20.

263. Hafza N, Challita C, Dandachi I, Bousaab M, Dahdouh E, and Daoud Z. Competition assays between ESBL-producing E. coli and K. pneumoniae isolates collected from Lebanese elderly: An additional cost on fitness. Journal of Infection and Public Health. 2018; 11:393–7. DOI: 10.1016/j.jiph.2017.09.010

264. Flores GE, Caporaso JG, Henley JB, Rideout JR, Domogala D, Chase J, Leff JW, Vazquez-Baeza Y, Gonzalez A, Knight R, Dunn RR, and Fierer N. Temporal variability is a personalized feature of the human microbiome. Genome Biology. 2014; 15. DOI: 10.1186/s13059-014-0531-y

265. Angelin M, Forsell J, Granlund M, Evengard B, Palmgren H, and Johansson A. Risk factors for colonization with extended-spectrum beta-lactamase producing Enterobacteriaceae in healthcare students on clinical assignment abroad: A prospective study. Travel Medicine and Infectious Disease. 2015; 13:223–9. DOI: 10.1016/j.tmaid.2015.04.007

266. Zimmer J, Lange B, Frick JS, Sauer H, Zimmermann K, Schwiertz A, Rusch K, Klosterhalfen S, and Enck P. A vegan or vegetarian diet substantially alters the human colonic faecal microbiota. European Journal of Clinical Nutrition. 2012 Jan; 66:53–60. DOI: 10.1038/ejcn.2011.141

267. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B, White O, Kelley ST, Methé B, Schloss PD, Gevers D, Mitreva M, and Huttenhower C. Metabolic Reconstruction for Metagenomic Data and Its Application to the Human Microbiome. PLOS Computational Biology. 2012 Jun 13; 8:e1002358. DOI: 10.1371/journal.pcbi.1002358

268. Clarridge JE. Impact of 16S rRNA Gene Sequence Analysis for Identification of Bacteria on Clinical Microbiology and Infectious Diseases. Clinical Microbiology Reviews. 2004 Oct; 17:840–62. DOI: 10.1128/CMR.17.4.840-862.2004

269. Amarasinghe SL, Su S, Dong X, Zappia L, Ritchie ME, and Gouil Q. Opportunities and challenges in long-read sequencing data analysis. Genome Biology. 2020 Feb 7; 21:30. DOI: 10.1186/s13059-020-1935-5

270. Bentley DR et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008 Nov 6; 456:53–9. DOI: 10.1038/nature07517

271. Shalon D, Smith SJ, and Brown PO. A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. Genome Research. 1996 Jul 1; 6:639–45. DOI: 10.1101/gr.6.7.639

272. Sanger F and Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. Journal of Molecular Biology. 1975 May 25; 94:441–8. DOI: 10.1016/0022-2836(75)90213-2

273. Maxam AM and Gilbert W. A new method for sequencing DNA. Proceedings of the National Academy of Sciences of the United States of America. 1977 Feb; 74:560–4. DOI: 10.1073/pnas.74.2.560

274. Poirel L, Liard A, Rodriguez-Martinez JM, and Nordmann P. Vibrionaceae as a possible source of Qnr-like quinolone resistance determinants. Journal of Antimicrobial Chemotherapy. 2005 Dec 1; 56:1118–21. DOI: 10.1093/jac/dki371

275. Sana TG, Flaugnatti N, Lugo KA, Lam LH, Jacobson A, Baylot V, Durand E, Journet L, Cascales E, and Monack DM. Salmonella Typhimurium utilizes a T6SS-mediated antibacterial weapon to establish in the host gut. Proceedings of the National Academy of Sciences. 2016 Aug 23; 113:E5044–E5051. DOI: 10.1073/pnas.1608858113

276. Agus A, Planchais J, and Sokol H. Gut Microbiota Regulation of Tryptophan Metabolism in Health and Disease. Cell Host & Microbe. 2018 Jun 13; 23:716–24. DOI: 10.1016/j.chom.2018.05.003

277. Magne F, Gotteland M, Gauthier L, Zazueta A, Pesoa S, Navarrete P, and Balamurugan R. The Firmicutes/Bacteroidetes Ratio: A Relevant Marker of Gut Dysbiosis in Obese Patients? Nutrients. 2020 May 19; 12:E1474. DOI: 10.3390/nu12051474

278. Pitout JD and DeVinney R. Escherichia coli ST131: a multidrug-resistant clone primed for global domination. F1000Research. 2017 Feb 28; 6:F1000 Faculty Rev–195. DOI: 10.12688/f1000research.10609.1

279. Cantón R and Coque TM. The CTX-M beta-lactamase pandemic. Current Opinion in Microbiology. Antimicrobials/Genomics 2006 Oct 1; 9:466–75. DOI: 10.1016/j.mib.2006.08.011

280. Royer G, Darty MM, Clermont O, Condamine B, Laouenan C, Decousser JW, Vallenet D, Lefort A, Lastours V de, Denamur E, Wolff M, Alavoine L, Duval X, Skurnik D, Woerther PL, Andremont A, Carbonnelle E, Lortholary O, Nassif X, Abgrall S, Jaureguy F, Picard B, Houdouin V, Aujard Y, Bonacorsi S, Meybeck A, Barnaud G, Branger C, Lefort A, Fantin B, Bellier C, Bert F, Nicolas-Chanoine MH, Page B, Cremniter J, Gaillard JL, Leturdu F, Sollet JP, Plantefève G, Panhard X, Mentré F, Marcault E, Tubach F, Zarrouk V, Bert F, Duprilot M, Leflon-Guibout V, Maataoui N, Armand L, Nguyen LL, Collarino R, Munier AL, Jacquier H, Lecorché E, Coutte L, Gomart C, Fateh OA, Landraud L, Messika J, Aslangul E, Gerin M, Bleibtreu A, Lescat M, Walewski V, Mechaï F, Dollat M, Maherault AC, Wolff M, Mercier-Darty M, Basse B, and COLIBAFI and SEPTICOLI groups. Phylogroup stability contrasts with high within sequence type complex dynamics of Escherichia coli bloodstream infection isolates over a 12-year period. Genome Medicine. 2021 May 5; 13:77. DOI: 10.1186/s13073-021-00892-0

281. Nielsen KL, Stegger M, Kiil K, Godfrey PA, Feldgarden M, Lilje B, Andersen PS, and Frimodt-Møller N. Whole-genome comparison of urinary pathogenic Escherichia coli and faecal isolates of UTI patients and healthy controls. International Journal of Medical Microbiology. 2017 Dec 1; 307:497–507. DOI: 10.1016/j.ijmm.2017.09.007

282. Ranjan A, Scholz J, Semmler T, Wieler LH, Ewers C, Müller S, Pickard DJ, Schierack P, Tedin K, Ahmed N, Schaufler K, and Guenther S. ESBL-plasmid carriage in E. coli enhances in vitro bacterial competition fitness and serum resistance in some strains of pandemic sequence types without overall fitness cost. Gut Pathogens. 2018 Jun 15; 10:24. DOI: 10.1186/s13099-018-0243-z

283. Escobar-Páramo P, Grenet K, Le Menac'h A, Rode L, Salgado E, Amorin C, Gouriou S, Picard B, Rahimy MC, Andremont A, Denamur E, and Ruimy R. Large-Scale Population Structure of Human Commensal Escherichia coli Isolates. Applied and Environmental Microbiology. 2004 Sep; 70:5698–700. DOI: 10.1128/AEM.70.9.5698-5700.2004

284. Abram K, Udaondo Z, Bleker C, Wanchai V, Wassenaar TM, Robeson MS, and Ussery DW. Mash-based analyses of Escherichia coli genomes reveal 14 distinct phylogroups. Communications Biology. 2021 Jan 26; 4:1–12. DOI: 10.1038/s42003-020-01626-5

285. Ducarmon QR, Zwittink RD, Willems RPJ, Verhoeven A, Nooij S, Klis FRMvd, Franz E, Kool J, Giera M, Vandenbroucke-Grauls CMJE, Fuentes S, and Kuijper EJ. Gut colonisation by extended-spectrum beta-lactamase-producing Escherichia coli and its association with the gut microbiome and metabolome in Dutch adults: a matched case-control study. The Lancet Microbe. 2022 Apr 20; 0. DOI: 10.1016/S2666-5247(22)00037-4

286. Cummins EA, Hall RJ, McInerney JO, and McNally A. Prokaryote pangenomes are dynamic entities. Current Opinion in Microbiology. 2022 Apr 1; 66:73–8. DOI: 10.1016/j.mib.2022.01.005

287. Nji E, Kazibwe J, Hambridge T, Joko CA, Larbi AA, Damptey LAO, Nkansa-Gyamfi NA, Stålsby Lundborg C, and Lien LTQ. High prevalence of antibiotic resistance in commensal Escherichia coli from healthy human sources in community settings. Scientific Reports. 2021 Feb 9; 11:3372. DOI: 10.1038/s41598-021-82693-4

Appendix One

Addendum

A.1 Impact

Antimicrobial resistance is recognised as one of the largest current threats to global health. Many Low- and Middle-income countries contain a high degree of AMR bacteria amongst their populations which, alongside the increase in globalisation throughout the years, has contributed to an exceptional rate of transmission between countries. This is especially problematic with ESBL-E which is recognised as one of the predominant highly virulent and AMR pathogens that plague the world. Meanwhile, there is an upwards trend in the amount of focus that the gut microbiome receives from the scientific community as research discovers the many links it has with human health. A known role of the gut microbiome is in creating a colonisation resistance to invading species, but the sheer complexity behind the number of interactions between microorganisms and the host has meant that it is still a field of research with an unquantifiable potential. This research in this thesis aims to better link these two fields and explain how involved the gut microbiome is with the acquisition of AMR bacteria during travel to Low- and Middle-income countries.

The ability to predict the risk of acquiring AMR internationally could aid how guidance is provided to travellers before travel, especially high-risk individuals, in order to minimise the transmission of AMR bacteria. In this thesis, we first studied the impact that the pre-travel gut microbiome content has on the predisposition to acquiring ESBL-E. Our findings show that the microbiome structure and composition are measurements that are either too insignificant with the acquisition of ESBL-E, or too imprecise in making a solid prediction. As this thesis highlights that the bacterial population alone does not distinguish into a protective 'type' of microbiome, there is a

need for additional omics analyses of the bacterial transcription or metabolic profiles. These results are the foundation to this field, but with further analyses, the guidance offered to travellers can be updated to incorporate the healthiest behaviours for one's gut microbiome.

How the gut microbiome is altered in healthy people during invasion of ESBL-E highlights the ways in which some travellers are seemingly resistant to colonisation. Results outlined in this thesis show how acquiring ESBL-E seemingly has little effect on the gut microbiome as a whole, instead the onset of travellers diarrhoea is the most significant perturbation. The large, longitudinal data set used in this study allowed us to carry out robust statistics which are lacking from this field of research, meaning that other research can refer to these results or access the sequencing data that was made publicly available. This thesis explores methods that can be utilised in future research on the pre-travel risk of developing diarrhoea. There is therefore a high potential to discover a gut microbiome structure or composition that is more resilient to diarrhoeal perturbation, and in turn resilient to ESBL-E taking advantage of an inflamed environment. Especially with the increased potential behind faecal microbiome transplantation, discovering a protective microbiome would lead to intervention strategies that minimise transmission of AMR bacteria, and the research in this thesis helps focus future research in this field.

Finally, we increased the resolution on the gut microbiome and focussed on the relationship that individual species have with ESBL-E during travel. Until the work in this thesis, this area of research remained understudied due to the lack of species-level technologies and sufficient number of samples. The outcome from this research highlights that members of Bacteroidaceae or other members of Enterobacteriaceae are potential targets of probiotic research against ESBL-E. We also discovered how the *E. coli* population itself is very dynamic in a travel-dependent manner and this provides strong leads for future research to better understand how this may be utilised to counteract the acquisition of ESBL-E.

To conclude, this thesis contributes to the field of research on the gut microbiome during travel by explaining the roles, of lack thereof, that the gut microbiome or its individual components play in the acquisition of ESBL-E. The results provide the groundwork for various future research endeavours, but also emphasise the need to increase the resolution of the analysis technology and to incorporate more components

of the gut microbiome, in order to fully understand how to exploit this to prevent the spread of AMR.

A.2 Acknowledgements

Behind all of the science of the past few years, there's a myriad a people that supported me through this journey. I will forever be grateful to the people I got to know during my PhD.

It is only fitting that I start with thanking my supervisors **Alan**, **Willem**, **John** and **Petra**. At first I was a little concerned to have four supervisors, I was expecting there to be constant disagreements and me having to choose sides, but you all consistently wanted what was best for me. I couldn't have asked for better guidance and support. I started the PhD at the age of 22 and, with a lot of patience from you all, I managed to come out the other end as a somewhat competent scientist! Thank you from the bottom of my heart for the chance to work with you and all of the support you gave at every step. I look forward to a future in microbiology together as peers.

I would like to thank all of the staff and researchers involved in creating the data used in this thesis. From the staff at the travel clinics, to the researchers in the COMBAT consortium who conceived the study design and acquired funding, this project could not have happened about their input.

A huge thank you of course goes to my family who, ever since I was a kid, have supported my route through education. **Mum**, I attribute so much of my success in life to how attentively and with empathy you raised and loved me, and I hope you know how grateful I am! **Dad**, you always placed so much importance on education and continuing development and I wouldn't be anywhere near where I am today without your influence and guidance.

I started this PhD in Birmingham and want to thank the incredibly warm and welcoming team I first joined. **Steven**, without your endless bioinformatics knowledge I would've floundered at step 1. **Ann**, 9 months of you feeding me creme eggs definitely wasn't good for my health but the energy you brought to work was always so infectious. **Harry**, I don't think Vito ever forgave me for how distracting we were but I wouldn't have changed a thing. And of course thanks to the countless others: **Danai**, **Chris**, **Beth**, **Hannah**, **Hu Ya**, **Emma**, **Greg**, **Ross**, **Lisa** and **Stan**, it's been an honour developing as scientists together!

In my first time moving country, I moved to complete half of the PhD in Maastricht. I don't think I could start by thanking anyone other than **Gianluca**. I developed so much as a person during these years and I attribute so much of this to you. You welcomed me into your group of friends and our weekly board games, Downton Abbey, D&D&D and endless fun were always the best way to end a hard week. **Simone**, you are the loudest gay I have probably ever met, but you always wear your heart on your sleeve. **Ian**, although you throw shade at every chance, you always had so much love to give. **Jorge**, I will never forget sitting in a church next to a Nun whilst crying my eyes out watching you beautifully play the viola. I would like to also thank my dear roommates who were always there at the end of the day. **Pilar**, you brought your Spanish energy and love each day, but were always so deeply empathetic too. **Jane**, I've learned so much from you just watching you grab life by the horns and I appreciate the greater consideration for the planet I developed from you. **Kilian**, although you poisoned me with fennel bread, it was always a delight coming home to your laid back energy.

My great time in Maastricht is of course also attributable to my MUMC friends. **Mayk**, **David** and **Rebekah**, thank you for all the bouldering tips and after work fun. **Casper** and **Heike**, you were both always exceptionally kind and inclusive. And everyone else for making work such a welcoming environment: **Brian**, **Liene**, **Niels**, **Melissa**, **Nader**, **Giang**.

I returned to Birmingham to finish the PhD and expected to stay for 6 months before going back to the Netherlands. 18 months later and I've only just left! I want to thank everyone in Birmingham for making work so incredibly fun. Of course this is largely thanks to my main gals. **Lilly**, it is so precious to me that I can quote the most nonsensical thing to you and we'll end up parroting it to each other for months. I couldn't have asked for a better colleague. **Laura**, my twin, I'm struggling to briefly write how grateful I am for you- I can summarise it by saying everyone needs a friend like you in their life. **Ann**, you get a second acknowledgement! Although Alan baiting you with my return was stressful for you, I knew I was welcome back to Birmingham with open arms. It may be controversial but you're the original schatje patatje. Top bloke **Freya**, when I'm a professional badminton player I'll give you a dedication! **Bex**, the Mum of the group and now of little **Edison**, I genuinely aspire to be as half as good as person and scientist as you. Still not sorry about being a distraction though. **Rob**, you always brought an infectiously lighthearted and laid back energy to work and I will always treasure our shared love of Pokemon. **Ryan**, I knew you for like a week and yet

I've made a lifelong friend, walk that duck. And of course everyone else from Birmingham that I've had the privilege of being friends with and making the PhD a delight to complete: **Emma**, **Sophie**, **James**, **Jeremy**, **Lizzy**, **Sara**, **Ruth**, **Sarah**, **Eric**, **Izzie**.

And last but not least, **Alex**. Words cannot properly express how grateful I am to have you in my life and for the unwavering support you have showed me throughout my PhD, I love you.

A.3 About the author

Matthew Davies was born in Wigan (United Kingdom) on the 16th of January 1996. After finishing college at Winstanley College in 2014, he started his undergraduate masters in biological sciences at the University of Birmingham. In 2018, he graduated with a first class honours with a thesis on the role of periplasmic adapter proteins in Salmonella. He remained at the University of Birmingham to start a PhD programme shared with Maastricht University, which focussed on the role of the gut microbiome during international travel. After the PhD was completed, he started a short-term research project focussed on using machine learning to improve antibiotic prescribing practices for patients admitted to hospital with acute sepsis. Afterwards, he began his role at Amsterdam UMC as a bioinformatician with a specialisation on infections from invasive group A Streptococcus.