FEMS Microbiology Reviews, 46, 1–22
DOI: 10.1093/femsre/fuac008
Advance access publication date: 4 February 2022
Review Article

The microbial ecology of *Escherichia coli* in the vertebrate gut

Ebenezer Foster-Nyarko [1,2,3,*] and Mark J. Pallen [1,4,5]

[1]Quadram Institute Bioscience, Norwich Research Park, Norwich, NR4 7UQ, United Kingdom
[2]Norwich Medical School, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom
[3]Department of Infection Biology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom
[4]School of Veterinary Medicine, University of Surrey, Guildford, Surrey, GU2 7AL, United Kingdom
[5]School of Biological Sciences, University of East Anglia, Norwich Research Park, Norwich, NR4 7TU, United Kingdom
*Corresponding author: Department of Infection Biology, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, United Kingdom. E-mail: Ebenezer.Foster-Nyarko2@lshtm.ac.uk
One sentence summary: This review presents an overview of *E. coli* diversity studies encompassing human and other nonhuman vertebrate hosts in the 137 years since this organism was first described and outstanding gaps in our knowledge.
Editor: Suzana Salcedo

Abstract

Escherichia coli has a rich history as biology's 'rock star', driving advances across many fields. In the wild, *E. coli* resides innocuously in the gut of humans and animals but is also a versatile pathogen commonly associated with intestinal and extraintestinal infections and antimicrobial resistance—including large foodborne outbreaks such as the one that swept across Europe in 2011, killing 54 individuals and causing approximately 4000 infections and 900 cases of haemolytic uraemic syndrome. Given that most *E. coli* are harmless gut colonizers, an important ecological question plaguing microbiologists is what makes *E. coli* an occasionally devastating pathogen? To address this question requires an enhanced understanding of the ecology of the organism as a commensal. Here, we review how our knowledge of the ecology and within-host diversity of this organism in the vertebrate gut has progressed in the 137 years since *E. coli* was first described. We also review current approaches to the study of within-host bacterial diversity. In closing, we discuss some of the outstanding questions yet to be addressed and prospects for future research.

Keywords: *Escherichia coli*, resident *E. coli*, ecology, within-host diversity, vertebrate gut

Introduction

Escherichia coli: a versatile organism

'Tout ce qui est vrai pour le Colibacille est vrai pour l'éléphant" (All that is true of *E. coli* is also true of the elephant; Friedmann 2004). This widely quoted epigram reflects the exalted status of *E. coli* in microbiology as the favourite model organism. Many advances in biology have been driven by studies with *E. coli*, particularly, using the strain designated K-12 and its derivatives (Tenaillon *et al.* 2010, Friedmann 2014; Table 1).

Theodor Escherich, a Bavarian doctor, first described *E. coli* on 14 July 1885 in a lecture to the Society for Morphology and Physiology in Munich (Escherich 1988). Over 15 months, Escherich observed and isolated 19 different bacteria in Otto von Bollinger's bacteriology laboratory (Escherich 1988), including the *Bacterium coli commune* ('the common colon bacillus'; Escherich and Pfaundler 1903, Shulman *et al.* 2007)—which we now know as *E. coli*—and *Bacterium lactis aërogenes* (now *Klebsiella pneumoniae*). Aided by Christian Gram's new staining technique (Escherich 1888, Bettelheim 1986, Friedmann 2014), Escherich described the organism as a Gram-negative bacillus of approximately 1.1–1.5 μm × 2.0–6.0 μm. The name *E. coli* was proposed by Castellani and Chalmers in 1919 to honor Escherich, and was officially adopted in 1958 by the Judicial Commission of the International Committee on Systematic Bacteriology. It was subsequently included in the Approved Lists of Bacterial Names in 1980 (Castellani and Chalmers 1919, Judicial Commission of the International Committee on Bacterial Nomenclature 1958, Skerman *et al.* 1980, Bettelheim 1986).

In the wild, *E. coli* exists as a common resident of the vertebrate gut and nonhost associated habitats such as water, soil, manure, and food (Macfarlane and Macfarlane 1997, Alm *et al.* 2011, van Elsas *et al.* 2011, Blount 2015). Due to its remarkable metabolic and regulatory abilities, *E. coli* can also survive under prolonged periods of nongrowth (Leimbach *et al.* 2013). Thus, *E. coli* represents a highly versatile species, capable of adapting to many different ecological habitats (Leimbach *et al.* 2013). *Escherichia coli* is also a versatile pathogen, eliciting a broad spectrum of diseases and responsible for at least 2 million human deaths per year (Kosek *et al.* 2003). The organism's role in intestinal and extraintestinal disease was recognized not long after its discovery (Friedmann 2014). Ørskov and Ørskov note, 'Any *E. coli* strain can probably cause invasive disease given the right opportunities, and *E. coli* has therefore aptly been called an opportunistic pathogen' (Orskov and Orskov 1992).

Escherichia coli is a facultative anaerobe, meaning that it can grow in aerobic and anaerobic conditions. It is the most common aerobe in the lower intestine of mammals (Tenaillon *et al.* 2010); however, it typically constitutes only 0.1%–5% of a gut microbial community comprised of over 500 other bacterial species (Tenail-

Received: June 18, 2021. **Revised:** January 31, 2022. **Accepted:** February 1, 2022

Table 1. The contribution of studies using *E. coli* across various fields.

Field(s)	Contribution of studies with *E. coli*	Reference(s)
Molecular biology, physiology, and genetics	The explanation of the genetic code	Crick *et al.* 1961
	DNA replication	Lehman *et al.* 1958
	Transcription	Stevens 1960
	The life cycle of bacterial viruses	Ellis and Delbrück 1939, Lwoff 1953
	The elucidation of the molecular basis of antibiotic tolerance	Hu and Coates 2005
	The discovery of restriction enzymes	Linn and Arber 1968, Meselson and Yuan 1968
	Swarming motility	Harshey and Matsuyama 1994, Inoue *et al.* 2007
	Gene regulation	Jacob *et al.* 1960
	Elucidation of the structure and function of ATP synthase	Capaldi *et al.* 2000
Pharmaceuticals	The synthesis of recombinant proteins *in vivo*, such as insulin, which is used to treat millions of people with diabetes worldwide	Swartz 2001
	The synthesis of several other biopharmaceuticals, such as human interferon-β, interleukin-2, human growth hormone, and human blood clotting factors	Huang *et al.* 2012, Baeshen *et al.* 2015
Evolution	*Escherichia coli* is the model organism of choice in experimental evolution studies; e.g. in Lenski's long-term evolution experiment, on-going since February 1988 and spanning over 60 000 generations	Good *et al.* 2017
	The demonstration of the stochastic nature of mutations	Luria and Delbrück 1943, Lederberg and Lederberg 1952
	Mapping the trajectory of long-term fitness	Wiser *et al.* 2013
	The elucidation of how sexual recombination influences adaptation	Cooper 2007
	Insights into predator–prey interactions	Lenski 1988
	The evolution of a novel trait, aerobic citrate utilization	Blount *et al.* 2012
Genetic engineering and biotechnology	The development of recombinant DNA techniques and molecular cloning, the production of biofuels and industrial chemicals such as phenol	Kim et al. 2014
	Mannitol production	Kaup *et al.*
	Ethanol production	Hildebrand *et al.* 2013, Liu and Khosla 2010

lon *et al.* 2010). This community is typically dominated by obligate anaerobes, including members of Firmicutes and Bacteroidetes—which make up at least 90% of the gut microbial population (Bäckhed *et al.* 2005, Claesson *et al.* 2009, Tenaillon *et al.* 2010). Nevertheless, *E. coli* can hold its ground in this highly competitive, ever-changing niche, existing in a life-long relationship with its host.

Escherichia coli pathotypes

Pathogenic *E. coli* are classified into 'pathotypes' or 'pathovars' (Kaper *et al.* 2004, Croxen and Finlay 2010; Table 2) based on several criteria, including the site of infection (e.g. uropathogenic *E. coli* (UPEC), named for their impact on the urinary tract, extraintestinal pathogenic *E. coli* (ExPEC), which cause infections in organs outside the gut), host (e.g. avian pathogenic *E. coli* (APEC), named after infections in avian species), site and host (e.g. neonatal meningitis *E. coli* (NMEC), which infect the cerebrospinal fluid in new-borns), and pathogenesis (e.g. Shiga-toxigenic *E. coli* (STEC)).

The pathogenic strains of *E. coli* have acquired specific virulence factors that enable them to adapt to new niches and cause a wide variety of diseases (Kaper *et al.* 2004, Denamur *et al.* 2020). These include adhesion/colonization factors, toxins, and effectors enabling pathogenic strains to colonize sites such as the urethra and small intestine and affect various fundamental eukaryotic processes (Kaper *et al.* 2004; Tables 3 and 4). For example, pathogenic *E. coli* can adhere to host epithelial cells using adhesion molecules known as intimin (Dutta *et al.* 2015). Pathogenic *E. coli* are also known to induce inflammation in host cells, resulting in the production of inducible nitric oxide synthase by the host innate immune cells. This results in the accumulation of nitrate, which the pathogenic strains can utilize to produce energy (Kamada *et al.* 2013). UPEC strains are equipped with type I

Table 2. Classification and examples of the pathotypes of *E. coli*.

Pathotype/pathovar	Virulence mechanism(s)	Host range	Reference
InPEC	Locus of enterocyte effacement; pathogenicity island 1	Humans and all mammals	Robins-Browne *et al.* (2016)
InPEC	Small fimbrial adhesins; toxins; transcriptional activator gene; aggregative adhesion	HumansRobins-Browne *et al.* (2016)	Robins-Browne *et al.* (2016)
InPEC	Shiga toxin or verotoxins; afimbrial and fimbrial adhesins	Humans and piglets	Robins-Browne *et al.* (2016)
InPEC	Heat labile and heat-stable enterotoxins	Humans, ruminants, pigs, and dogs	Robins-Browne *et al.* (2016)
InPEC	Invasion and multiplication in enterocytes	Humans and primates	Robins-Browne *et al.* (2016)
InPEC	Adhesins	Humans and animals	Robins-Browne *et al.* (2016)
InPEC	Shiga toxin	Humans	[323]
InPEC	Adherent invasive phenotype	Humans and animals	Robins-Browne *et al.* (2016)
ExPEC	Adhesins, secretion and iron uptake systems, increased serum survival and cytotoxic proteins	Birds	Sheldon *et al.* 2010
ExPEC	Fimbrial adhesins; siderophores, resistance to complement	Humans and animals (especially dogs and cats)	Mainil 2013
ExPEC	Iron acquisition systems, degradation of interferon-gamma, and cleavage of the human defensin LL-37	Humans	Nicholson *et al.* (2016)
ExPEC	Fimbrial adhesins; siderophores; resistance to complement	All mammals and birds (especially poultry)	Mainil 2013
ExPEC	Unknown	Animals	Blum *et al.* 2015
ExPEC	Type II, IV, and VI secretion systems, long polar fimbriae (lpfA) and iron acquisition	Animals	Dogan *et al.* 2012
ExPEC	Cytotoxic Necrotizing Factors 1 or 2 and α haemolysin; fimbrial and/or afimbrial adhesins; siderophores; resistance to complement	Humans, animals, and ruminants	Mainil 2013

InPEC, Intestinal pathogenic *E. coli*; ExPEC, Extraintestinal pathogenic *E. coli*; EPEC, Enteropathogenic *E. coli*; EAEC, Enteroaggregative *E. coli*; EHEC, Enterohaemorrhagic *E. coli*; ETEC, Enterotoxigenic *E. coli*; EIEC, Enteroinvasive *E. coli*; DAEC, Diffusely adherent *E. coli*; STEAEC, Shiga-toxigenic Enteraggregative *E. coli*; AIEC, Adherent invasive *E. coli*; APEC, Avian pathogenic *E. coli*; UPEC, Uropathogenic *E. coli*; NMEC, Neonatal meningitis *E. coli*; SePEC, Human sepsis-associated *E. coli*; MPEC, Mammary pathogenic *E. coli*; EnPEC, Endometrial pathogenic *E. coli*; and NTEC, Necrotoxigenic *E. coli*.

fimbriae, *AfA/Dr* adhesins and pyelonephritis-associated pili (PAP) that enable them to colonize and infect the urinary tract. NMEC and sepsis-associated *E. coli* are armed with the K1 polysaccharide capsule that facilitates their evasion from host complement-mediated killing (Pitout 2012, Leimbach *et al.* 2013). Differential utilization of nutrients is another way in which pathogenic strains of *E. coli* can overcome the resistance mounted by the commensal strains. For example, pathogenic EHEC can use the sugars galactose, mannose, ribose, and hexuronates as carbon sources, while commensal *E. coli* cannot (Fabich *et al.* 2008, Le Bouguénec and Schouler 2011). Besides, EHEC possesses the *eut* operon, which enables it to assimilate nutrients such as ethanolamine, whereas commensal *E. coli* lack this operon, and are therefore, unable to utilize ethanolamine (Perna *et al.* 2001, Fabich *et al.* 2008, Bertin *et al.* 2011).

Despite its designation as a separate genus, *Shigella* is classified as an intestinal pathogenic *E. coli* (InPEC) pathotype (InPEC encompasses Adherent Invasive *E. coli* (AIEC) and diarrheagenic *E. coli*). It possesses virulence traits and pathogenicity that closely resemble enteroinvasive *E. coli* (EIEC) and is, thus regarded as an EIEC pathotype (Pupo *et al.* 2000, Escobar-Páramo *et al.* 2003). Phylogenomic data support the classification of *Shigella* as an *E. coli* pathotype (Lan *et al.* 2004, The *et al.* 2016), even though its nomenclature has been retained solely for historical reasons and to avoid confusion in the clinical setting. Chaudhuri and Henderson (2012) advocate that the clinical and academic community working on *E. coli/Shigella* adopts a similar approach as was used in the classification of species such as *Salmonella typhimurium, Salmonella enteritidis,* and *Salmonella typhi* as *Salmonella enterica* subgroups so that the *Shigella* species are re-designated as subspecies within *E. coli* to avoid their neglect in *E. coli* studies.

Despite the importance of pathotyping to the epidemiology and pathogenesis of strains, the recent discovery of complex hybrid pathotypes has challenged some of the classical pathotype definitions (Denamur *et al.* 2020), as illustrated in the following examples.

The strain that caused an outbreak of foodborne illness in Germany in 2011 (Frank *et al.* 2011), which swept across most of Europe, killing 54 individuals and causing approximately 4000 infections including 900 cases of haemolytic uraemic syndrome (HUS), was, in fact, an EHEC-EAEC hybrid strain (Karch *et al.* 2012, Denamur *et al.* 2020) belonging to serotype O104:H4 and sequence type ST678. This strain combined virulence characteristics of EHEC (Shiga toxin production) and adherence typical of EAEC strains, despite lacking the Type III secretion and *tir*/intimin system. It also possessed virulence factors commonly found in ExPEC strains,

Table 3. Colonization and fitness factors (adapted from Kaper *et al.* (2004)).

Virulence factor	Pathotype/pathovar	Effect(s)
icsA/virG	EIEC	Nucleates actin filaments
Intimin	EPEC/EHEC	Adhesin, inducing TH1 response
Dr adhesins	DAEC/UPEC	Adhesin, binds to decay-accelerating factor and activates phosphatidylinositol 3-kinase, induces MHC class I chain-related gene A
P (*Pap*) fimbriae	UPEC	Adhesin, also induces cytokine expression
Colonization factor antigens	ETEC	Adhesin
S fimbriae	UPEC/NMEC	Adhesin
Bundle-forming pili (BFP)	EPEC	Type IV pili
Aggregative adherence fimbriae	EAEC	Adhesin
paa	EPEC/EHEC	Adhesin
toxB	EHEC	Adhesin
Efa-1/LifA	EHEC	Adhesin
Long polar filaments	EHEC/EPEC	Adhesin
saa	EHEC	Adhesin
ompA	NMEC/EHEC	Adhesin
Curli	Various	Adhesin, binds to fibronectin
ibeA/B/C	NMEC	Stimulates invasion
aslA	NMEC	Stimulates invasion
Dispersin	EAEC	Stimulates colonization; facilitates mucous penetration
K antigen capsules	MNEC	Antiphagocytic activity
Aerobactin	EIEC	Siderophore, iron acquisition
Yersiniabactin	Various	Siderophore, iron acquisition
ireA	UPEC	Siderophore, iron acquisition
iroN	UPEC	Siderophore, iron acquisition
***chu* (*shu*)**	EIEC/UPEC/NMEC	Siderophore, iron acquisition
Flagellin	All	Motility, inducing cytokine expression through Toll-like receptors
Liposaccharides	All	Inducing cytokine expression through Toll-like receptors

Please see Table 2 for the full names of the pathovars/pathotypes.

such as yersiniabactin and aerobactin (iron acquisition factors) and demonstrated expanded-spectrum beta-lactamase (ESBL) resistance (Karch *et al.* 2012).

A hybrid clone belonging to serotype O80:H2 and sequence type ST301 (clonal complex 165) has emerged in France, Belgium, and Switzerland, capable of causing HUS and bacteraemia (Mariani-Kurkdjian *et al.* 2014, Soysal *et al.* 2016, Thiry *et al.* 2017, Nüesch-Inderbinen *et al.* 2018, De Rauw *et al.* 2019). This strain possesses all the virulence factors typical of EHEC strains, such as intimin, Shiga toxin production, and enterohaemolysin, yet it belongs to the phylogenetic group A, unlike other EHEC strains (Mariani-Kurkdjian *et al.* 2014). It possesses a large plasmid (> 100 kb), which encodes a resistance cassette, providing resistance characteristics to a wide range of antimicrobials, including cotrimoxazole, tetracyclines, streptomycin, and penicillin (Soysal *et al.* 2016, Cointe *et al.* 2020). The carriage of large plasmids is commonly associated with ExPEC strains (Peigne *et al.* 2009, Nicholson *et al.* 2016).

Escherichia coli are classified into eight major phylogenetic groups (phylotypes or phylogroups) linked with the ecological niche, lifestyle, and pathogenic potential of strains (Walk *et al.* 2009, Alm *et al.* 2011). Phylogroups B2, D, F, and H appear to be the most basal taxa, with phylogroup E emerging before phylogroups C, B1, and A, which are thought to be the most recently diverged lineages (Lu *et al.* 2016, Beghain *et al.* 2018). The most anciently diverged encompass mostly extraintestinal pathogenic strains. In contrast, the most recently diverged lineages span strains that are associated with life-threatening intestinal diseases such as dysentery and HUS (Picard *et al.* 1999, Escobar-Páramo *et al.* 2004a). Hybrid clones exhibiting characteristics of the B2 phylogenetic backbone typical of ExPEC strains and EPEC and STEC attributes have been recently reported in the literature, further blurring the lines of pathotype boundaries that have been used to define *E. coli* pathogens for so long. Isolates belonging to serotype O153:H10 and harboring *eae* (depicting an atypical EPEC-ExPEC hybrid pathotype) have been detected in meat, poultry farms, human diarrhoeagenic samples, and wildlife from northwest Spain (Díaz-Jiménez *et al.* 2020). Similarly, isolates belonging to serotype O137:H6 and ST2678 and positive for *eae*, *bfpA*, and *stx2f* genes—demonstrating typical EPEC-STEC hybrid pathotype—have been isolated in exotic psittacine birds (Gioia-Di Chiacchio *et al.* 2018). In addition, strains belonging to serotype O2:H4 and ST141 and demonstrating uropathogenic traits have been found with some characteristics of the EHEC pathotype (Bielaszewska *et al.* 2014, Gati *et al.* 2019). Also, an ST12 strain belonging to serotype O4:H1 was found to harbor the locus of enterocyte effacement pathogenicity island and the bundle forming pili protein (encoded by the *bfp* gene) typical of the EPEC pathotype but simultaneously causing diarrhoea and bacteraemia in the same patient (Kessler *et al.* 2015). Furthermore, some EPEC isolates that harbor the heat-stable enterotoxin produced by ETEC strains have been described (Dutta *et al.* 2015).

These examples amply demonstrate that pathotyping is limited in its capacity to adapt to new strains that fail to respect the currently utilized pathotype boundaries. However, it can be argued that the emergence of pathovars that exhibit traits unchar-

Table 4. *Escherichia coli* toxins and effectors (adapted from Kaper *et al.* (2004)).

Virulence factor	Pathotype/pathovar	Toxin class	Effect(s)
Heat-labile enterotoxin	ETEC	AB subunit/type II effector	ADP ribosylates and activation of adenylate cyclase, leading to ion secretion
Shiga toxin	EHEC	AB subunit	Depurination of rRNA, inhibiting protein synthesis and inducing apoptosis
Cytolethal distending toxin	Various	ABC subunit	DNase activity, blocks mitosin in G2/M phase
Shigella enterotoxin 1	EAEC/EIEC	AB subunit	Ion secretion
Urease	EHEC	ABC subunit	Cleaves urea to NH_3 and CO_3
EspC	EPEC	Autotransporter	Serine protease, cleavage of coagulation
EspP	EHEC	Autotransporter	Serine protease, cleavage of coagulation factor V
Haemoglobin-binding protease	ExPEC and APEC	Autotransporter	Degradation of haemoglobin to release haem/iron
pet	EAEC	Autotransporter	Serine protease; ion secretion and cytotoxicity
pic	UPEC, EAEC, and EIEC	Autotransporter	Protease/mucinase
sat	UPEC	Autotransporter	Vacuolation
sepA	EIEC	Autotransporter	Serine protease
sigA	EIEC	Autotransporter	Ion secretion
Cycle-inhibiting factor	EPEC and EHEC	Type III effector	Blocks mitosis in G2/M phase, resulting in the inactivation of cdk1
espF	EPEC and EHEC	Type III effector	Opens tight junctions and reduces apoptosis
espH	EPEC and EHEC	Type III effector	Modulates filopodia and pedestal formation
map	EPEC and EHEC	Type III effector	Disrupts mitochondrial membrane potential
tir	EPEC and EHEC	Type III effector	Nucleates cytoskeletal proteins, loss of microvilli and GAP-like activity
ipaA	EIEC	Type III effector	Actin depolymerization
ipaB	EIEC	Type III effector	Apoptosis, Interleukin-1 release and membrane insertion
ipaC	EIEC	Type III effector	Actin polymerization
ipaH	EIEC	Type III effector	Modulation of inflammation
ipgD	EIEC	Type III effector	Inositol 4-phosphatase and membrane blebbing
VirA	EIEC	Type III effector	Microtubule destabilization and membrane ruffling
stcE	EHEC	Type II effector	Cleavage of C1-esterase inhibitor and disruption of the complement cascade
hlyA	UPEC	Repeats-in-toxin (RTX) toxin	Cell lysis
ehx	EHEC	RTX toxin	Cell lysis
Cytotoxic necrotizing factors (1 and 2)	NMEC, UPEC, and NTEC	-	Altered cytoskeleton and necrosis
LifA/Efa	EPEC and EHEC	-	Inhibits lymphocyte activation and adhesion
Shigella enterotoxin 2	EIEC and ETEC		Ion secretion
Heat-stable enterotoxin a	ETEC	Heat-stable enterotoxins	Activating guanylate cyclase, leading to ion secretion
Heat-stable enterotoxin b	ETEC	Heat-stable enterotoxins	Ion secretion *via* increasing intracellular calcium
Enteroaggregative *E. coli* heat-stable enterotoxin	Various	Heat-stable enterotoxin	Activating guanylate cyclase, leading to ion secretion

Please see Table 2 for the full names of the pathovars/pathotypes.

acteristic of their genetic background is not unexpected, given that pathotypes are usually defined based on horizontally acquired elements like plasmids and genomic islands. Thus, while the majority of a pathovar's membership may show similarity in pathogenic phenotype, horizontal genetic exchange would result in an elastic continuum within a pathovar's definition.

A further limitation lies in the use of negative criteria to identify pathogens. For instance, a strain is described as a typical EPEC if it presents the locus of enterocyte effacement but does not produce Shiga toxin. On the other hand, if a strain both lacks Shiga toxin and *bfp*, it is classified as an atypical EPEC (Kaper *et al.* 2004). As Robins-Browne *et al.* (2016) point out, 'characterizing pathogens based on their lack of one or more virulence determinants may group several types of distantly related or unrelated bacteria together and cause some distinct pathogenic categories with uncharacterized virulence determinants to be overlooked'.

The taxonomy of *Escherichia*

In the premolecular era, several non-*coli* species were designated under the *Escherichia* genus based on DNA relatedness/hybridization and overall phenotypic similarity: *Escherichia blattae* (1973) (Burgess *et al.* 1973), *Escherichia vulneris* (1982) (Brenner *et al.* 1982a), *Escherichia hermannii* (1982) (Brenner *et al.* 1982b), and *Escherichia fergusonii* (1985) (Farmer et al. 1985).

Lawrence *et al.* (1991) utilized DNA sequences of the slowly evolving genes *gap* and *ompA* (encoding glyceraldehyde-3-phosphate dehydrogenase (GAPDH) and outer membrane protein 3A, respectively), representing conserved genes across 11 species of enteric bacteria, to map the phylogenetic relationships among the above mentioned five species. Their analysis revealed these species to be distantly related, rather than a monophyletic group. Apart from *E. fergusonii*, the other species were more divergent from *E. coli* than *Salmonella*. As *Salmonella* is traditionally consid-

ered most closely related to *E. coli*, the results justified the redesignation of *E. blattae*, *E. hermannii*, and *E. vulneris* in alternative genera (Lawrence *et al.* 1991).

Subsequently, in 1991, Albert *et al.* (1991) isolated a diarrhoeagenic isolate with EPEC-like phenotypic and genetic features from a 9-month-old girl with watery diarrhoea in Bangladesh, initially designated *Hafnia alvei*. Subsequent identification and characterization of five '*H. alvei*-like' isolates by DNA–DNA hybridization, phenotypic characterization, and 16S rDNA sequencing led to their redesignation as *Escherichia albertii* (Huys *et al.* 2003). Members of this species were later found to be closely related to *Shigella boydii* serotype 13—a divergent lineage in *Escherichia*; the *E. albertii*/*S. boydii* serotype 13 lineage estimated to have diverged from a most-recent common ancestor circa 28 million years ago (Hyma *et al.* 2005).

Consequently, the taxonomy of the genus *Escherichia* has been modified with the reclassification of *E. hermannii*, *E. blattae*, and *E. vulneris* to other genera and the description of five cryptic clades within *Escherichia* (Walk *et al.* 2009, Priest and Barker 2010, Hata *et al.* 2016, Jain *et al.* 2018). Until recently, the genus was comprised of three named species: *E. albertii*, *E. coli*, and *E. fergusonii* and five cryptic clades, designated *Escherichia* clades I–V (Denamur *et al.* 2020). Of the three named species within the genus, *E. albertii* is the most divergent species, while *E. fergusonii* is closely related to *Escherichia coli sensu stricto* (Clermont *et al.* 2011).

The name *Escherichia marmotae* has recently been validly published for Clade V (Liu *et al.* 2015, 2019), although the species is not restricted to marmots (Ocejo *et al.* 2020, Gilroy *et al.* 2021). This builds on an earlier suggestion to classify clade V as a novel species and clades III and IV combined as a second novel species, based on digital DNA–DNA association and Average Nucleotide Identity (ANI) data (Walk *et al.* 2009; an ANI of 95% between two genomes is generally taken as the standard for the demarcation of prokaryotic species; Richter and Rosselló-Móra 2009, Jain *et al.* 2018). Furthermore, cryptic clade I strains possess the trademark virulence traits of *E. coli*, besides their potential to cause human infections (Steinsland *et al.* 2010, Clermont *et al.* 2011, Ingle *et al.* 2011). Thus, *E. coli* and *Escherichia* clade I are now designated as *Escherichia coli sensu lato* and the classic *E. coli* (designated phylogroups A–G) as *E. coli sensu stricto* (Clermont *et al.* 2013). Recently, Gilroy *et al.* (2021) assigned the species name *Escherichia whittamii* to clade II, in honor of the American bacteriologist, Thomas S. Whittam, for his contributions to the study of *E. coli*. van der Putten *et al.* have also recently proposed that the cryptic clades III and IV be named *Escherichia ruysiae* sp. nov (van der Putten *et al.* 2021).

Walk *et al.* (2009) estimated the lineages' divergence times that gave rise to each of the cryptic clades using a minimum evolution tree. Based on the assumption that *E. coli* split from *S. enterica* between 100 and 160 million years ago, they estimated that the *Escherichia* lineages shared a common ancestor between 48 and 75 million years ago. They suggested that *E. albertii*, *E. fergusonii*, *Escherichia* clade II, and clade V split between 38 and 75 million years ago, whereas *E. coli*, *Escherichia* clade I, clade III, and clade IV split between 19 and 31 million years ago (Walk *et al.* 2009).

As evidenced by earlier efforts at classifying *E. blattae*, *E. hermanii*, and *E. vulneris*, metabolic characteristics such as the utilization of specific carbon sources or the production/catabolism of certain biochemical compounds are insufficient to delineate strains into species and species groups accurately. First, multiple genes are often required for the expression of a particular phenotype (Lawrence *et al.* 1991). Second, convergent evolution in distantly related species confounds the delineation of species based on these phenetic characteristics. Similarly, DNA hybridization fails to consider the relatedness among congeneric species reliably (Lawrence *et al.* 1991).

The Genome Taxonomy Database Toolkit (GTDB-Tk), a tool for the automatic classification of draft bacterial and archaeal genomes, was published recently (Chaumeil *et al.* 2019). GTDB-Tk inputs genome assemblies in FASTA format and predicts the placements within domains based on identifying a set of 120 bacterial and archaeal marker genes and domain-specific reference trees. Then, species designations are computed using a GTDB reference tree, the relative evolutionary divergence parameter and ANI values. However, GTDB automatically assigns a random alphanumeric designation to novel species, which do not scale well to the increasing number of newly identified novel species and are often confusing and user-unfriendly (Gilroy *et al.* 2021). To address this gap, Pallen *et al.* (2020) recently developed an automated combinatorial approach to creating more than 1 million Linnaean binomials for Bacteria and Archaea.

The emergence of antimicrobial resistance in *E. coli*

Escherichia coli has long been used as an indicator bacterium for monitoring the faecal contamination of food and water and antimicrobial resistance (AMR) in enteric bacteria of animals and humans (Teuber 1999). For example, *E. coli* has been employed as the model organism in determining the level of AMR in bacteria from people in close contact with food animals, such as those who work at abattoirs and veterinarians—with the observation that such people harbor significantly higher levels of resistant *E. coli* compared to the overall community (Bongers *et al.* 1995). The ability to acquire and transfer AMR traits was recognized in *E. coli* as far back as the 1960s: 26%–61% and 50%–76% of human and pig multidrug resistant *E. coli* strains, respectively were found to transfer resistance to lab strains of *E. coli* K-12 in conjugation experiments (Datta 1969, Guinée *et al.* 1970, Linton *et al.* 1972, Gyles *et al.* 1977, Davies and Stewart 1978, Bourque *et al.* 1980, Nijsten *et al.* 1996).

According to the World Health Organization (WHO), the ever-increasing levels of AMR is one of the most significant threats to human health (www.who.int/entity/drugresistance/en), with substantial economic implications (Brown *et al.* 2017). Resistance to antimicrobials represents a daunting challenge to treating many infections, not least those caused by *E. coli*.

As a pathogen, *E. coli* is a prominent cause of urinary tract infections (UTIs), gastroenteritis, and bloodstream infections, among others; and as such, antibiotics are frequently applied to treat *E. coli* infections (Foxman 2010, de Kraker *et al.* 2013, Riley 2020). The use (and abuse) of antibiotics in treatment is linked with antibiotic resistance development (Webber and Piddock 2005, Ventola 2015). Microorganisms exhibit a natural ability to resist antimicrobials' action—a phenomenon referred to as intrinsic resistance (Cheng *et al.* 2016). In addition, AMR can arise due to a gene mutation or the acquisition of resistance determinants via horizontal gene transfer (de la Cruz and Davies 2000). Horizontal gene transfer can take the form of free DNA uptake (transformation), plasmid-mediated transfer of resistance gene traits (conjugation), or *via* phage-mediated transfer (transduction; Burmeister 2015, Lee 2019).

Escherichia coli frequently displays resistance to multiple classes of antimicrobials (MDR)—with observed rates of MDR identified among strains that cause UTI and bacteraemia exceeding 50% (Croxall *et al.* 2011, Alhashash *et al.* 2013, de Kraker *et al.* 2013). The

occurrence of MDR in *E. coli* has been focused on a small number of widely dispersed clones, mediated by ESBL and carbapenemase-encoding plasmids (Dunn *et al.* 2019). Worryingly, the rates of *E. coli* extraintestinal infections such as bacteraemia and UTIs have been rapidly increasing in recent years due to an increase in the number of antibiotic-resistant infections caused by 'superbugs' such as the ST131 clone (Croxall *et al.* 2011, Banerjee and Johnson 2014, McNally *et al.* 2019).

Although several strains have been linked with the global emergence and dissemination of MDR in *E. coli* (e.g. strains belonging to ST88, ST410, ST648, ST405, and ST73), none equal ST131 in their extent (Dunn *et al.* 2019). The MDR ST131 clone belongs to a sub-lineage designated clade C which arose from two predominantly drug-susceptible clades by acquiring large MDR plasmids which confer ESBL and metallo-beta-lactamases encoding resistance to carbapenems and cephalosporins up to third generation; as well as point mutations leading to fluoroquinolone resistance (Ben Zakour *et al.* 2016, Stoesser *et al.* 2016).

McNally *et al.* (2019) recently analysed an extensive collection of more than 1000 ST131 strains. They reported that the acquisition of colonization and fitness factors and the accumulation of genes encoding dehydrogenase enzymes involved in anaerobic metabolism account, at least in part, for the successful expansion of this clone (McNally *et al.* 2019). These results indicate hyper-resistant clones' ability to outcompete resident nonpathogenic strains of the same species, and thus facilitate their long-term colonization in the gut (Dunn *et al.* 2019, Benz *et al.* 2021). The fact that MDR strains such as ST131 are well-suited to colonize the gut competitively suggests that AMR traits can be readily transferred from resistant strains to co-colonizing susceptible strains, with the potential to hamper the treatment of future infections caused by such strains.

Accumulating data shows that ExPEC are frequently isolated from diseased companion animals and livestock—highlighting the potential for zoonotic as well as anthroponotic transmission (Achtman *et al.* 1986, Johnson *et al.* 2008a,b, 2009, Ewers *et al.* 2010, Zogg *et al.* 2018). High rates of AMR among *E. coli* isolates from livestock and poultry have been documented, linked with the agricultural use of antimicrobials (Bennani *et al.* 2020). The use of antimicrobials as growth promoters in poultry feed was banned in Europe in 2006 (Castanon 2007), with a resultant sharp drop in AMR rates among isolates from livestock and poultry (Bennani *et al.* 2020). However, this may be less well controlled in other parts of the world, particularly in low-to-middle income countries. Worryingly, the usage of antimicrobials in developing countries is likely to increase as intensive farming practices are adopted (Van Boeckel *et al.* 2015).

Several studies have investigated the link between antimicrobial usage and AMR in animals and AMR in humans. In 2015, a systematic review (Lazarus *et al.* 2015) of the published literature to quantify the zoonotic transfer of ESBL-encoding extraintestinal *E. coli* infections found that six studies established the zoonotic transfer of AMR by whole bacterial transmission using molecular methods, in particular, *via* poultry in the Netherlands. A further 13 molecular studies suggested the mobile-genetic element-mediated transfer of AMR from resistant organisms in animals to susceptible strains in humans, while four observational epidemiological studies inferred the zoonotic transfer of AMR to humans.

Although the authors cautioned that their conclusions might not be geographically generalizable, it appears that a proportion of human extraintestinal infections caused by ESBL-encoding strains originate from food-producing animals. Further, *E. coli* strains from commercial broilers have recently been reported to share resistance profiles with strains recovered from human extraintestinal infections (Borges *et al.* 2019). This is worrying and highlights the need for further investigation, especially with whole-genome sequence-based studies (Webber and Piddock 2005).

Escherichia coli genomics

Frederick Blattner first conceived the *E. coli* K-12 genome sequence project in 1983 (Blattner 1983); however, due to funding and technological challenges, it was not until 15 years later that the project (which was 6-year-long) was finally completed (Blattner *et al.* 1997). '*E. coli* was the obvious choice for a sequencing effort', noted Frederick Neidhardt in Pennisi (1997), since more was known about the organism than any other (Hayashi *et al.* 2006), and so much information had been gleaned from studies with *E. coli*, that, as Blattner stated, 'Figuring out the microbe's genetic code would help integrate all those years of study' (Pennisi 1997). Blattner and colleagues submitted the final 2.0 Mb of the 4.6 Mb *E. coli* genome to GenBank on 16 January 1997 (Pennisi 1997). This was closely followed by the deposition in GenBank of an incomplete genome sequence of the closely related strain, W3110, 7 days later (Pennisi 1997) (the complete genome sequence of which was published in 2006 (Hayashi *et al.* 2006)). Based on atypical codon usage and base composition, Lawrence and Ochman (1998) inferred 18% of the K-12 genome to have arisen from horizontal transfer (earlier studies had shown that horizontally acquired genes exhibited atypical codon usage, base composition, and dinucleotide frequencies; Muto and Osawa 1987, Lawrence and Ochman 1997). This estimate was later revised to 24.5%, using improved methods (Lawrence and Ochman 2002).

Next, the genome sequences of two EHEC O157:H7 isolates were swiftly published, expanding the number of *E. coli* genomes for comparative studies. First came the EDL933 strain isolated from Michigan ground beef connected to the 1982 multistate outbreak by Perna *et al.* (2001), then RIMD 0508992—the strain that caused a large outbreak in 1996 in Sakai city in Osaka prefecture, Japan, involving at least 6000 school children (Watanabe *et al.* 1996)—the latter sequenced by Hayashi *et al.* (2001). Comparisons with K-12 revealed a shared sequence of 4.1 Mb, representing a common chromosomal 'backbone' of *E. coli*, with the remaining sequence comprising strain-specific clusters varying in size, encoding putative virulence factors, prophages, and prophage-like elements (Hayashi *et al.* 2001, Perna *et al.* 2001). These analyses provided evidence of extensive horizontal gene transfer, with the description of many 'K' and 'O' islands—depicting introgressed DNA present only in K-12 but not in O157:H7 or only in O157:H7, respectively. The existence of a shared *E. coli* backbone was confirmed by a three-way comparison of the K-12 and O157:H7 strains with the genomic sequence of the third *E. coli* strain to be completed: the ExPEC strain CFT073 (Welch *et al.* 2002). A surprising finding was that all three strains shared only 39.2% of the combined nonredundant set of proteins. The role of horizontal gene transfer in the pathogenic strains' evolution was also highlighted by the presence of several pathogenicity islands exhibiting atypical codon usage interrupting the common backbone (Welch *et al.* 2002).

With the availability of more *E. coli* genome sequences, the core and accessory genome concept was defined to represent a conserved set of roughly 2200 genes common to all *E. coli* strains and strain-specific sequences respectively (Rasko *et al.* 2008, Touchon *et al.* 2009, Chaudhuri *et al.* 2010). The core genomic sequence lends itself as a useful tool for the phylogenomic comparison of

isolates, provided the effect of homologous recombination (estimated to affect about a 10th of the *E. coli* core genome; Mau *et al.* 2006) is accounted for (Chaudhuri *et al.* 2010).

The pan-genome concept

The concept of a 'pan-genome' was proposed by Tettelin *et al.* (2005), during an attempt to utilize genome sequence information from Group B streptococcus to predict proteins that might be exposed on the organism's surface and could be exploited as vaccine candidates (Tettelin *et al.* 2005, Tettelin and Medini 2020). By this concept, the pangenome of each bacterial species is defined by three distinct components: namely, its core genome, representing the genes found in each isolate of the species; the accessory genome, depicting the genes present in several but not all isolates of the species, and strain-specific genes detected in one isolate only. Through genomic comparisons of nineteen GBS isolates, Tettelin and colleagues uncovered the first evidence that closely related isolates differed significantly in their gene content. A single isolate of a particular species was insufficient to capture the species' genome (Tettelin *et al.* 2005, Tettelin and Medini 2020).

In a study that compared 61 sequenced genomes of *E. coli*, Lukjancenko *et al.* (2010) predicted a pangenome comprised of 15 741 gene families, with only 993 (6%) of the gene families present in every genome (the core genome)—indicating an accessory genome of more than 90% in *E. coli*. This equates to an accessory or variable gene content of approximately 4/5th of any given *E. coli* genome (Lukjancenko *et al.* 2010). These results corroborate those reported by Rasko *et al.* (2008) and others before them (Ochman and Jones 2000, Fukiya *et al.* 2004). Rasko *et al.* identified a pangenome comprised of more than 13 000 genes *via* comparison of seventeen *E. coli* reference genome sequences encompassing human commensal and distinct clinical groups of *E. coli* and a core gene set of about 2200 genes conserved in all isolates. It has become apparent that the more *E. coli* genomes are sequenced and compared, the more the pangenome continues to increase—what has been referred to as an 'open pangenome' (Rasko *et al.* 2008)—and the core genome shrinks (Robins-Browne *et al.* 2016). Thus, the accessory genome content contributes crucial insights which, coupled with single nucleotide polymorphisms (SNPs) in the core genome, can be employed to track the evolutionary history of natural isolates, as has been recently demonstrated by McNally *et al.* (2016) and Decano and Downing (2019). The accessory genome content has arisen from repeated gene acquisition and the contemporaneous loss of sequences is thought to account for the distinctions between divergent lineages within the same species (Ochman and Jones 2000). They include genes encoding virulence determinants, bacteriophages, virulence factors, and acquired AMR determinants (Robins-Browne *et al.* 2016).

Whole genome sequencing (WGS) offers many advantages for the diagnosis and understanding of the pathobiology of *E. coli* strains, in that it is possible to predict most of the pathotypes of *E. coli* based on the presence of well-recognized virulence factors, as well as elucidating the full array of virulence factors possessed by individual strains within a particular pathotype. Furthermore, typing schemes that combine several genes within the core genome (e.g. core genome MLST available on platforms such as EnteroBase; Zhou *et al.* 2020) and others that would incorporate the accessory genome are expected to become the mainstay of *E. coli* analysis, particularly as the accessibility of whole-genome data continues to increase (Robins-Browne *et al.* 2016).

The ecology of *E. coli* in the gut

Microbial ecology is the study of the diversity, distribution, and abundance of microorganisms, and how microorganisms interact with each other and their environment to generate and preserve such diversities (Xu 2006). There are two areas of focus that have encapsulated microbial ecological studies to date: namely, the examination of microbial diversity—'who is there,' i.e. the identification and characterization and estimation of abundance across a variety of niches; and the study of microbial activity, i.e. what microorganisms are doing and how they are doing it—including their biotic and abiotic interactions and how they impact the observed diversity and the ecosystem.

Members of the gut microbiota have co-diversified with their hosts over millions of years and exist in a mutualistic relationship with their host, in which the gut microbiota carry out vital functions for their hosts and in return occupy a nutrient-rich environment (Round and Mazmanian 2009, Nicholson *et al.* 2012). The gut microbial communities' composition is thought to be influenced by factors such as diet, physiology of the gut, host phylogeny, and diet (Ley *et al.* 2008, Muegge *et al.* 2011, Wu *et al.* 2011, David *et al.* 2014). A healthy gut microbiota plays critical roles in developing the host immune system, and is required for homeostasis in adult life (Hooper and Macpherson 2010, Flint 2012). For example, the gut microbiota cells help maintain the balance between host metabolism and the immune system and in the large intestine, metabolize the indigestible components of the diet (Bäckhed *et al.* 2005, Turnbaugh *et al.* 2009). The gut microbiota also detoxify toxic products and serve as a barrier against the colonization of opportunistic pathogens (termed as colonization resistance; Round and Mazmanian 2009, Claus *et al.* 2016). The mechanisms by which the resident intestinal microbiota elicit resistance against the colonization and invasion of pathogens include: the direct competition for nutrients, the modification of metabolites such as bile salts and short-chain fatty acids that render them toxic to invading pathogens, the alteration of pH and oxygen tension, the induction of host antimicrobial peptides, the expression of a dense mucous, IgA and cellular immunity, and direct attacks through the production of bacteriocins or Type IV secretion (Brown and Valiere 2004, Kamada *et al.* 2013).

Host-directed pathways act to break down complex protein, carbohydrate, and fat in ingested food in eukaryote hosts into simple forms, thereby facilitating the absorption of nutrients. In addition, members of the microbiota contribute to the catabolic process in producing short-chain fatty acids utilized by the host and support the growth of the bacterial community (Sorbara and Pamer 2019). The members of the microbiota have evolved to maximize the available nutrients such that under homeostatic conditions, an exogenous species needs to overcome the intense competition for limited substrates to engraft themselves in the available niches. Resident *E. coli* use sugars such as fucose, mannose, galactose, ribose, arabinose, and *N*-acetyl glucosamine in the mucous layer—which are required by pathogenic strains such as *E. coli* O157—and in this way, mount colonization resistance to invading pathogens (Maltby *et al.* 2013). Besides, commensal *E. coli* employ the preferential utilization of nutrients to out-compete exogenous strains. For example, resistance to EHEC by commensal *E. coli* is due to the competition for nutrients such as amino acids and organic acids (Chang *et al.* 2004, Fabich *et al.* 2008, Leatham *et al.* 2009, Maltby *et al.* 2013).

Competition for trace metals also appears to be significant in colonization resistance, as demonstrated in murine studies,

where the Nissle 1917 *E. coli* strain was found to protect against *Salmonella* infection by competing for iron (Deriu *et al.* 2013).

However, intestinal microorganisms constitute a persistent invasion threat, considering their vast numbers and the large intestinal surface area (Hooper and Macpherson 2010). Emerging evidence suggests that dysbiosis—defined as a shift in the composition of the intestinal microbiota, and thus an alteration of the relationship between the host and the gut microbiota—is linked to the development of various diseases, such as inflammatory bowel disease, obesity, allergy, autoimmune disease, and irritable bowel syndrome (Round and Mazmanian 2009, Codling *et al.* 2010, Kadooka *et al.* 2010, Kang *et al.* 2013, Miyoshi *et al.* 2014, Quince *et al.* 2015, Lynch and Boushey 2016, Tanaka and Nakayama 2017).

Estimates show that anaerobic bacteria outnumber *E. coli* anywhere from 100:1 up to 10 000:1 (Berg 1996). The prevalence of *E. coli* in the various hosts they colonize varies widely (0%–100%), being influenced by host characteristics such as body size, microbiota, diet, and digesta retention times (Tenaillon *et al.* 2010). Over 90% of humans carry *E. coli*, while about 25%–56% of wild mammals appear to be colonized by the organism (Mitsuoka *et al.* 1976, Gordon and Cowling 2003, Penders *et al.* 2006, Tenaillon *et al.* 2010, Lescat *et al.* 2013). The prevalence rate in human-associated animals (such as chickens and cats) is estimated to be above 60% (Lescat *et al.* 2013). We recently reported a prevalence of 56% from a range of nonhuman primate species with varying degrees of contact with humans (Foster-Nyarko *et al.* 2020b).

In the gut, *E. coli* reside in the mucous covering of the epithelial cells and is shed with the degraded mucus components into the intestinal lumen and subsequently excreted in faeces (Poulsen *et al.* 1994, Poulsen *et al.* 1995). Human faeces typically contain between 10^2 and 10^9 colony-forming units (cfu) of *E. coli* per gram (Slanetz and Bartley 1957, Bettelheim *et al.* 1972, Mitsuoka *et al.* 1976, Penders *et al.* 2006), while an estimated 10^4–10^6 cfu can be detected in the faeces of domestic animals (Slanetz and Bartley 1957). Data on the quantity of *E. coli* in stools of wild animals is, however, lacking. As first observed by Escherich (1886), *E. coli* is one of the first bacterial colonizers of the infant gut (Mueller *et al.* 2015), achieving concentrations of up to 10^9 cfu per gram of the stools of infants (Mitsuoka *et al.* 1976, Penders *et al.* 2006). Subsequently, anaerobic members of the microbiota expand and dominate the gut (Syed *et al.* 1970). Given that *E. coli* is a facultative anaerobe, its ability to utilize oxygen probably helps create an anaerobic environment favoring the blooming of strict anaerobes (Mueller *et al.* 2015). As a gut microbiota member, *E. coli* produces vitamin K and mounts resistance against colonization by pathogens (Suvarna *et al.* 1998, Richter *et al.* 2018). Thus, *E. coli* exists in a mutualistic relationship with the human host, although it is mostly described as a commensal (Blount 2015, Martinson and Walk 2020).

Given that most *E. coli* reside innocuously in the gut, an important ecological question that has been plaguing microbiologists is what makes *E. coli* an occasionally devastating pathogen (Lescat *et al.* 2013, Martinson and Walk 2020)? To address such questions requires an enhanced understanding of the ecology of the organism as a commensal. However, nonpathogenic *E. coli* have been traditionally underrepresented in ecological studies of this species (Martinson and Walk 2020). More studies exploring the populations of resident or nonpathogenic *E. coli* within and between hosts and how these populations vary over time are needed to shed light on this evolutionary puzzle.

Humans are exposed to *E. coli* through multiple routes (Reeves *et al.* 2011, Akoachere *et al.* 2014, Nwankwo *et al.* 2014, Pal *et al.* 2015, Ercumen *et al.* 2017, Aijuka *et al.* 2018, Derakhshandeh *et al.* 2018, Abia and Ubomba-Jaswa 2019, Toombs-Ruane *et al.* 2020, Yar 2020): such as the consumption of contaminated food and water; through fomites, e.g. on bank coins and notes (particularly in the cracks of creased notes) and cell phones; pets and domestic animals and the environment. This high level of exposure is reflected in reports of more than one strain in normal stools (Sears *et al.* 1950, Sears and Brownlee 1952, Wiedemann *et al.* 1971, Bettelheim *et al.* 1972, Shooter *et al.* 1977, Chen *et al.* 2013, Dixit *et al.* 2018, Richter *et al.* 2018).

A total of two plausible theories explain the fate of swallowed strains (Sears *et al.* 1950, Sears and Brownlee 1952, Bettelheim *et al.* 1972). The displacement theory suggests that newly ingested strains may fail to establish themselves, in which case they are voided out, or if they succeed in establishing themselves, will displace the 'resident' strains present. Alternatively, the dominant–minor strain theory posits that freshly ingested strains do not replace the established strains within the gut but co-exist as minority or transient strains, albeit in small numbers and may be detected from time to time in the stool along with the dominant strain.

Experimental studies with the Nissle 1917 strain have shown that not all strains are equal in their propensity to establish themselves following immigration (Tannock *et al.* 2011). Evidence suggests that repeated exposure may facilitate immigration and the establishment of strains (Johnson and Clabots 2006, Johnson *et al.* 2008a). Host-related factors appear to contribute to the pattern of phylogroup distribution in the gut. In human studies, phylogroups E and F have been found to encourage a more heterogeneous population of *E. coli* following their establishment in the gut. Conversely, the establishment of phylogroup B2 strains in the gut appears to limit the immigration and establishment of other phylogroups (Gordon *et al.* 2015). Also, locale-specific differences in the distribution of *E. coli* phylotypes have been reported, linked with differences in climate and diet (Duriez *et al.* 2001, Escobar-Páramo *et al.* 2004b). For instance, strains belonging to phylogroups A and B1 are often prevalent in stools sourced from people living in developing countries. On the other hand, stools from people in developed countries are frequently dominated by *E. coli* belonging to phylogroups D and B2 (Skurnik *et al.* 2008, Massot *et al.* 2016).

The co-existence of multiple strains in a single host raises an essential question about the factors that govern residency in the gut, i.e. how incoming strains overcome the colonization resistance posed by the existing *E. coli* population. The current body of evidence suggests the following.

Freter's successful competition hypothesis

Freter theorized that successful colonization occurs due to successful competition for nutrients (Freter *et al.* 1983a,b, Pereira and Berry 2017). Accordingly, the gut microbiota composition is determined by several limiting substrates, which different members of the microbiota can utilize with variable efficacy. Conway and colleagues demonstrated this principle in *E. coli* strains (*Escherichia coli* strains HS and Nissle 1917 vs. *E. coli* O157:H7) using carbohydrate metabolism in a mouse model (Maltby *et al.* 2013, Conway and Cohen 2015). This theory is in tandem with Gause's exclusion principle, which precludes two organisms' co-existence if they share the same limiting resource (Hardin 1960). Iron competition appears to influence the colonization of resident strains, as demonstrated by studies that found strains lacking siderophore genes to have a reduced ability to establish themselves in mouse models, compared to wild-type strains. Conversely, resident strains in the hu-

man gut have been found to encode siderophores (Nowrouzian *et al.* 2001), signaling their potential contribution to successful colonization. The mechanisms by which two organisms exclude each other from a particular niche might be through either direct (e.g. through bacteriocin production or phage to damage or kill competitors) or indirect (e.g. through passive resource utilization) competition (Ghoul and Mitri 2016, Pereira and Berry 2017, Granato *et al.* 2019).

Efficient utilizers

Some colonizers thrive because they utilize available nutrients much more efficiently than others who use the same nutrients. In murine studies, mice fed with *E. coli* MG1655, nonmotile *flhD* mutants were found to persist in stools collected 3 days post-feeding and dominate the population (roughly 90%) by day 15. The mutants colonized better than the wild-type parent strain and grew in the caecal mucus faster than the wild-type counterparts (Leatham *et al.* 2005, Fabich *et al.* 2011). Further analyses using high-throughput genomic approaches revealed that the *flhD* mutants possessed an enhanced ability to oxidize several carbon sources because the loss of FlhD conferred an increased expression of genes involved in carbon and energy metabolism. (The *flhDC* operon encodes the FlhD4C2 regulatory complex, which has been shown to negatively regulate the genes involved in galactose transport and the citric acid cycle while positively regulating the genes involved in ribose transport; Gauger *et al.* 2007, Fabich *et al.* 2011).

Restaurant hypothesis

Leatham-Jensen *et al.* (2012) have described how polysaccharide-degrading anaerobes break down polysaccharides into sugars, which they serve to *E. coli* cells within a shared biofilm (Ng *et al.* 2013). This is an example of 'syntrophy'—an association where one organism feeds on the nutritional products of another (Dolfing 2014)—as has been observed between *Bacteroides ovatus* and *Bacteroides thetaiotaomicron* (Tuncil *et al.* 2017). The biofilms that feed these *E. coli* strains are referred to as 'restaurants'. Many such restaurants are thought to exist, proposed to comprise a mix of different commensal strains, each serving different nutrients to the commensal *E. coli* strains residing therein (Conway and Cohen 2015).

Different nutritional requirements

This hypothesis suggests that pathogenic strains may utilize other nutrients than those of the commensal residents. For example, it has been shown that *E. coli* strain HS, *E. coli* Nissle 1917, *E. coli* MG1655, and *E. coli* EDL933 utilize unique metabolic niches in the mouse intestine (Fabich *et al.* 2008, Maltby *et al.* 2013, Conway and Cohen 2015). *Escherichia coli* HS uses six out of the 12 sugars available in the mucous layer, namely, arabinose, galactose, gluconate, lactose, ribose, and *N*-acetylglucosamine. *Escherichia coli* MG1655 uses five sugars, namely, fucose, arabinose, gluconate, *N*-acetylneuraminate, and *N*-acetylglucosamine. *Escherichia coli* Nissle 1917, on the other hand, utilizes the following seven sugars: fucose, galactose, arabinose, *N*-acetylglucosamine, gluconate, *N*-acetylneuraminate, and mannose. Lastly, *E. coli* EDL933 also uses a unique selection of carbon sources: galactose, arabinose, hexuronates, *N*-acetylglucosamine, mannose, sucrose, and ribose.

Different biogeographical niches

Pathogenic strains of *E. coli* may colonize different biogeographic regions of the gut from commensal strains. Scientists from the University of Oklahoma have shown that *E. coli* EDL933, a strain of enterohaemorrhagic *E. coli* O157:H7 colonizes a different niche in the mouse intestine to human commensal strains (Leatham *et al.* 2009, Maltby *et al.* 2013). In the bovine host (where *E. coli* O157:H7 commonly colonizes the gut innocuously), studies in a 12-month-old naturally colonized steer revealed a unique tropism of *E. coli* strain O157:H7 for the rectal mucosa adjoining the recto-anal junction (Naylor *et al.* 2003).

Hierarchical nutrient utilization

Here, closely related bacterial species may co-colonize the gut by hierarchical utilization of similar nutrients, as has been demonstrated between two generalists, *B. ovatus* and *B. thetaiotaomicron* (Tuncil *et al.* 2017). Thus, direct competition for substrates is excluded by each species utilizing the available glycans with differing priorities (i.e. order and speed of consumption).

Exploiting colonization resistance for therapy

Antibiotic treatment has significantly reduced mortality and morbidity associated with potentially life-threatening diseases, and thus saved millions of lives (Pamer 2016). However, broad-spectrum antibiotics rarely target pathogens alone and concomitantly result in deleterious effects on the commensal bacterial populations, resulting in increased susceptibility to infections due to alteration of the host microbiota (dysbiosis; Dethlefsen *et al.* 2008, Buffie *et al.* 2012, Buffie and Pamer 2013, Ravi *et al.* 2019, Ribeiro *et al.* 2020). This problem is compounded by the fact that many pathogens have become increasingly antimicrobial resistant. Consequently, the administration of live bacteria (probiotics) to restabilize the altered microbiota and restore the colonization resistance conferred by the resident microbiota has increasingly been the focus of intense research (Buffie *et al.* 2012, Buffie and Pamer 2013, Pamer 2016, Morrow and Wischmeyer 2017). This heightened interest has been fuelled by the success of faecal transplantation in patients suffering from diarrhoea associated with *Clostridiodes difficile* infection (Gough *et al.* 2011, Youngster *et al.* 2014, Manges *et al.* 2016).

Probiotics are defined as living microbes of human origin that can colonize host sites such as the gut and oropharynx when ingested in adequate quantities, and thus deliver benefits to the host (Hill *et al.* 2014, Morrow and Wischmeyer 2017). Probiotic strains are attractive as preventive measures against gut infections. Bacterial candidates that have been proposed to promote high-level colonization resistance to infection include members of the lactate-producing genera, e.g. *Lactobacillus* (Gorbach 2000) and *Bifidobacteria* (Sivan *et al.* 2015)—although evidence of their effectiveness in health in humans, reducing infections or fostering longevity is sparse (Pamer 2016).

Escherichia coli Nissle 1917 (also known as Mutaflor) is one of the most extensively researched probiotic strains globally (Jacobi and Malfertheiner 2011). Alfred Nissle observed that this strain ameliorates ulcerative colitis (Nissle 1951), following his isolation of the strain in 1917 from the stool of a soldier who did not suffer from diarrhoea like his comrades did during the First World War (Nissle 1959, Nissle 1961). The available data shows that Mutaflor possesses many traits that favor its suitability as a probiotic strain. For example, the strain forms a biofilm well and, in this way, outcompetes numerous pathogenic and nonpathogenic strains such as EPEC and ETEC (Hancock *et al.* 2010). The complete genome sequence information revealed that Mutaflor contains several fitness traits such as adhesion factors and an array of iron uptake systems that enable it to outcompete other bacteria and

block the adherence and invasion of pathogenic strains (Lodinová-Zádniková and Sonnenborn 1997, Boudeau *et al.* 2003, Altenhoefer *et al.* 2004). The strain also boosts intestinal barrier function and protects against epithelial disruption by EPEC (Zyrek *et al.* 2007). Moreover, *E. coli* Nissle 1917 induces the release of beta-defensin-2 (an antimicrobial peptide) from human epithelial cells, thus eliciting a broad antimicrobial response against both Gram-positive and Gram-negative bacteria and fungi and viruses and, thus inhibits invasion and colonization of other bacteria (Wehkamp *et al.* 2004).

Accumulating evidence suggests that probiotics comprised of a combination of strains are effective, owing to the symbiosis among strains—including those from different genera (Timmerman *et al.* 2007). An example is the VSL#3 multistrain probiotic consortium—comprising *Eubacterium faecium, Streptococcus thermophilus, Lactobacillus acidophilus*, *Bifidobacterium breve*, *Bifidobacterium infantis, Lactobacillus delbrueckii subspecies bulgaricus*, *Bifidobacterium longum*, *Lactobacillus plantarum*, and *Lactobacillus casei*—which has been shown to treat ulcerative colitis effectively (Venturi *et al.* 1999, Timmerman *et al.* 2007). Researchers working with *E. coli* EDL 933 found a combination of two strains—*E. coli* Nissle 1917 and *E. coli* HS, both promising probiotic candidates—can overlap nearly all the nutrient requirements of *E. coli* EDL 933 (Meador *et al.* 2014). However, the same combination of commensals could not prevent colonization by uropathogenic strain *E. coli* CFT073 and the enteropathogenic strain *E. coli* E2348/69. Different pathogenic strains, thus occupy distinct nutrient niches within the gut microbiome.

As interest in probiotics heightens, there is the need to understand the mechanics of how potential probiotic strains compete or co-exist with the commensal bacteria in the gut. We do not yet fully understand how probiotic strains interact with the usual residents of the gut. Previous attempts at unravelling the interactions between pathogens and commensals in the gut have been mainly carried out using well-characterized reference strains within *in vitro* environments. It will be desirable to test these hypotheses with the 'real world' strains in the natural habitat (i.e. the gut).

Within host bacterial diversity of *E. coli*

As previously discussed, continuous exposure to *E. coli via* multiple routes contributes to the turnover of strains in the vertebrate gut, thus resulting in considerable diversity among the population of *E. coli* that exists within a single individual and between different hosts. Besides immigration events, the diversity of *E. coli* is enriched by within-host evolution events (Didelot *et al.* 2016).

Members of a bacterial community undergo changes as they interact with each other and their hosts or environment. The sources of novel variation that account for within-host evolution may arise from the following.

Point mutations

These involve the substitution, insertion or deletion of a single nucleotide. Point mutations represent the smallest evolution unit (Bryant *et al.* 2012). Reeves *et al.* reported a within-host point mutation of approximately 1 per year per genome (corresponding to 1.44×10^{-7} accumulation rate per site per year) for *E. coli* (Reeves *et al.* 2011), while Lee *et al.* estimate a 2.54×10^{-10} mutation rate per site per generation (Lee *et al.* 2012).

Insertions and deletions (indels) and genomic rearrangements [466, 467]

Phylogenetic analysis of the O104:H4 strain that caused the large outbreak of gastroenteritis and HUS in Germany in 2011 revealed that the acquisition of a Shiga toxin 2-encoding prophage and an ESBL CTX-M-15-encoding plasmid contributed to its emergence [468–471]. Similarly, the acquisition of virulence gene-encoding plasmids into many ancestral *Shigella* spp. was crucial to the evolution of *Shigella* as human pathogens (Gati *et al.* 2019).

Bacterial genomes can also undergo rearrangements during DNA recombination, replication, or error-prone DNA repair (Periwal and Scaria 2015, Rodgers and McVey 2016). Genomic rearrangements alter chromosomes or large chromosomal regions. They involve the processes of deletion, duplication, insertion, inversion, or translocation. Homologous recombination leads to the reassortment of genes between chromosomal pairs; however, the genome's arrangement remains unchanged. Other forms of recombination may result in rearrangements of genomic DNA. DNA rearrangements contribute to gene expression and function and may contribute to genetic diversity in *E. coli* (Iguchi *et al.* 2006, Periwal and Scaria 2015). In long-term evolution experiments involving 12 *E. coli* populations propagated for more than 25 years in glucose-restricted environments, large-scale genomic rearrangements were found to occur in the same loci in at least two populations and contributed to genomic plasticity (Raeside *et al.* 2014). Moreover, most rearrangements were observed to confer higher fitness, highlighting the significant influence of large-scale rearrangements in genomic evolution.

Other factors that shape within-host diversity include genetic drifts: random processes by which allelic frequencies change over time due to birth and death of individuals within the population (Charlesworth 2009, Kuo *et al.* 2009), as well as by both purifying and diversifying selection (Didelot *et al.* 2016).

Recombination and horizontal gene transfer (HGT)

HGT involves the uptake of extracellular DNA (transformation), cell-to-cell transfer of genetic material *via* surface appendages (conjugation), or viral import (transduction) from genetically distant relatives. The acquisition of genetic sequences from unrelated organisms *via* HGT results in faster diversification of the genome compared to point mutation alone (Ochman *et al.* 2000)—consequently termed 'evolution in quantum leaps' (Hacker and Carniel 2001). Fragments of the chromosomal genome can be replaced with homologous sequences from another cell through homologous recombination, which plays a significant role in the evolution of *E. coli* (Didelot *et al.* 2012). In particular, the presence of mixed infection facilitates homologous recombination by providing material for import in the chromosome—as has been exemplified in *Helicobacter pylori*, where evolution is noted to accelerate up to a 100-fold in the presence of mixed infections (Kennemann *et al.* 2011, Didelot *et al.* 2013, Cao *et al.* 2015). Thus, homologous recombination is a strong driver of within-host evolution. Alternatively, nonhomologous sequences can be gained and incorporated into the genome—a phenomenon, i.e. compensated by genome degradation (Andersson and Andersson 1999).

Hypermutators

Besides spontaneous mutations, genetic variation can arise due to the breakdown of DNA repair mechanisms such as the mismatch repair system, producing bacteria with decreased replication fidelity (Mather and Harris 2013). In *E. coli*, the loss of mis-

match repair is known to increase mutation rates by up to 200-fold (Marinus 2010). Lee *et al.* (L2012) investigated the accumulation of spontaneous mutations over thousands of generations in wild-type *E. coli* compared to a mismatch repair-defective variant with the *mutL* gene encoding MutL, one of the key enzymes required for mismatch repair, deleted. In their experiment, the mutant ($MutL^-$) strain showed a 233-fold increase in mutation rate to Nal^R (resistance to nalidixic acid) compared to the wild type.

In particular, hypermutation is advantageous during infection when the ability to adapt quickly can facilitate evasion of host immune defences and antimicrobial therapy (Mather and Harris 2013). Hypermutator strains have been identified during an MRSA outbreak in a neonatal intensive care unit (Köser *et al.* 2012) and among *Staphylococcus aureus* isolates from cows suffering from benign forms of mastitis (Wang *et al.* 2013). Strikingly, in one study where cystic fibrosis patients with chronic respiratory problems were followed up over 38 years, nearly half of them harbored hypermutator *Pseudomonas aeruginosa* isolates (Marvig *et al.* 2013). Of note, the accumulation of SNPs *via* hypermutation can hamper the analysis of transmission and result in inaccurate conclusions (Mather and Harris 2013). In particular, the use of SNP thresholds to identify and delineate transmission lines could potentially exclude hypermutator strains or the inclusion of such strains may bias results (Mather and Harris 2013).

Investigating within-host bacterial diversity

To study within-host bacterial evolution in detail requires the application of whole-genome sequencing to multiple clinical samples taken from an individual host (these samples may be collected longitudinally or simultaneously from a single body site or several sites) or multiple isolates or genomes derived from a single clinical sample.

Sampling a single colony does not sufficiently capture the *E. coli* strain diversity in stools, as established by several studies (Bettelheim *et al.* 1972, Shooter *et al.* 1977, Dixit *et al.* 2018, Richter *et al.* 2018). Lidin-Janson *et al.* (1978) estimated that sampling five colonies enables detecting dominant genotypes at a > 99% probability. At the same time, Schlager *et al.* (2002) showed that sampling 28 colonies provide a > 90% chance of detecting minor genotypes in a stool sample.

In its simplest form, two genomes or more can be compared to each other by counting the number of positions where they differ. Within-host diversity is determined by comparing two or more genomes from the same host (Golubchik *et al.* 2013). Comparing two genomes from different hosts lends insight into chains of transmission. These approaches have been successfully applied to several organisms beside *E. coli* (Croxall *et al.* 2011, Alhashash *et al.* 2013, Didelot *et al.* 2013, Halachev *et al.* 2014, Worby *et al.* 2014, Kay *et al.* 2015, Ssengooba *et al.* 2016, Herranz *et al.* 2017, Trauner et al. 2017, Hall *et al.* 2019, Ley *et al.* 2019).

Genotyping multiple isolates per host can also facilitate the identification of multiply infected individuals, representing the carriage of bacterial subpopulations descended from a distinct founder strain. Multiple concurrent infections are clinically crucial for the evolution of pathogens, particularly in facilitating recombination between divergent strains as discussed above and hence, the emergence of novel genotypes (Wymant *et al.* 2018). Clinically, multiple infections are known to affect disease progression. For example, in HIV infection, the disease is reported to accelerate in the presence of dual infections (Cornelissen *et al.* 2012). A total of two plausible scenarios could explain the origin of multiple infections. In the first instance, two distinct strains could be transmitted simultaneously from an external source that harbors sufficient strain diversity. Alternatively, the two strains could have arisen from separate sources at different times.

Studies of the within-host diversity of *E. coli*

Studies of the genetic within-host diversity of *E. coli* in the human gut date back to well over a hundred years ago and have involved a wide variety of methodologies—encompassing both microbiologic and molecular techniques. The earliest studies involved the use of serotyping and subsequently multilocus enzyme electrophoresis. Later studies employed multilocus sequence typing and WGS to characterize strains and infer within-host diversity and evolution of strains. Below, we present a summary of studies investigating the within-host diversity of *E. coli* since 1890.

Pre-PCR era

The first evidence of temporal changes in the population of *E. coli* in the human gut was documented by Smith in 1899 (Smith 1899). In this pioneering study (reviewed in Wallick and Stuart 1943), Smith analysed 48 isolates obtained from four separate cultures of the stools of a single normal infant and reported a high degree of temporal antigenic stability. Subsequently, Totsuka (1902; reviewed in Wallick and Stuart (1943) and Martinson and Walk (2020)) self-collected and cultured his own stool once a week for 12 weeks and recovered up to 32 isolates per culture, yielding a total of 332 isolates. Upon agglutinating these against antisera, which he had prepared against isolates obtained from the first two cultures, he provided evidence of gradual shifts from one antigenic type to the other over the course of time in the intestine of one individual.

In a similar study, Wallick and Stuart (1943) isolated 650 *E. coli* strains from one individual over the course of 14 months and described four distinct serotypes that persisted for a month at a time—including co-occurrence of dominant and minority strains. They also inferred household transmission based on antigenic identity (serotype).

Sears, Brownlee, and Uchiyama (Sears *et al.* 1950, Sears and Brownlee 1952) investigated changes in *E. coli* serotypes over time in five adults sampled over 3 and 30 months. They were the first to coin the term 'residents' and 'transients' to describe strains of *E. coli* 'which establish themselves firmly and continue to multiply over extended periods of time' and those that persisted for only a few weeks at most respectively. For the first time, they also investigated the production of bacteriocins by resident clones to ward off the invasion of other clones.

Next, Robinet (Robinet 1962, Branche *et al.* 1963) confirmed the periodic fluctuations of antigenic types in six healthy individuals followed up for 6 months and concluded that host antibody production does not explain the turnover of resident strains of *E. coli* in the gut. Robinet was involved with further work investigating bacteriocin production by the resident clones by testing the activity of strains collected in 1 month against those recovered the previous month. These studies concluded that the turnover of strains was lower when the resident strains produced colicins.

Shooter *et al.* (1977) investigated the serotype dynamics of *E. coli* isolated from nine adults over 3 months. They combined H- and O-antigen testing for the first time and observed that several strains with an identical O antigen, in fact, displayed different H antigens. Following on from this, Selander and colleagues (Selander and Levin 1980, Caugant *et al.* 1981) sampled *E. coli* from a single individual for 11 months and, using multilocus enzyme electrophoresis, provided seminal insight on the genetic structure

of the resident and transient clones in the healthy adult gut. Importantly, they documented the loss and acquisition of plasmids in identical clones over time and the gain and loss of novel clones over the study period as essential contributors to the generation of *E. coli* diversity in the gut.

Finally, Tzabar and Pennington (1991) explored the dynamics of *E. coli* carriage in a group of individuals residing at the British Antarctic Survey research base for 26 weeks. Using MLEE, they provided evidence of strain-sharing among the members of the research station—although several isolates were lost to viability, hampering the study conclusions on the clonal turnover over time.

PCR era

Several studies (Johnson and Clabots 2006, Johnson *et al.* 2008b, Damborg *et al.* 2009) have explored the diversity of *E. coli* among adults, children and companion animals living in the same household over time. Key findings include the detection of clones that appeared to persist for 4–45 weeks, as well as documenting evidence of strain sharing among members of a household and their pets. Shared strains between humans and companion animals were found to be of phylogroups D and B2, with an ST73 strain (phylogroup B2) causing UTI in a family dog.

Martinson *et al.* (2019) studied eight healthy adults *via* biweekly sampling for 6 months to 2 years and observed that the resident clones often belonged to phylogroups A, B2, and F, including multiple resident clones persisting for more than a year in one individual.

Some studies have pointed to the possible exchange of plasmid-borne AMR genes between co-colonizing strains within a single individual. First, Stoesser *et al.* (2015) applied multiple colony sampling (16 colonies from each of eight faecal samples) followed by whole-genome sequencing to investigate the transmission of *E. coli* strains harboring ESBL genes collected in Cambodia. They reported substantial core- and accessory genome diversity, with a median of four STs recovered per individual. Remarkably, different clones from a single individual tended to share the same $bla_{CTX\text{-}M}$ variant, and identical clones were found with different $bla_{CTX\text{-}M}$ variants. Significant accessory genome diversity was also observed within and between clones—highlighting the utility of multiple colony sampling and whole-genome sampling in the analysis of *E. coli* within-host diversity.

Second, Knudsen *et al.* (2017, 2018) collected *E. coli* (and other enterobacteria) from the stools of children with cystic fibrosis or cancer along with matched healthy controls over a 9-month period, picking up to five colonies per faecal sample. Children with cystic fibrosis or cancer received antibiotic treatment over the course of the study. Of the isolated *E. coli* (90% of children with cystic fibrosis, 93% of children with cancer, and 94% of healthy controls), the prevalence of antibiotic-resistant enterobacteria did not significantly differ between the children who received antibiotics and the healthy controls at the start and at the end of the study, suggesting that the level of antibiotic resistance they observed arose from the community. As a follow-up study, seven isolates collected from three consecutive samples collected from one child with cystic fibrosis was characterized by WGS. Here, three distinct strains were detected, and the different strains were found to harbor Inc1 plasmids encoding $bla_{CTX\text{-}M\text{-}1}$. However, the plasmids from the three different strains were found to differ by only a few SNPs and varied with limited regions, suggesting recombination events. This study documented the horizontal transfer of $bla_{CTX\text{-}M\text{-}1}$-harboring plasmids within a single individual.

Stegger *et al.* (2020) analysed 20 *E. coli* isolates from each of nine urine samples collected from nine women who presented with UTI at a general practice in Zealand, Denmark, from which a single clone was detected in eight out of the nine samples. The authors then selected a total of 40 isolates belonging to the same clone: 10 each from two urine samples and two rectal swabs isolates from two healthy individuals (collected in a previous study; Nielsen *et al.* 2014) and investigated the intraclonal diversity of *E. coli* among the commensal and uropathogenic strains. A low intraclonal diversity was observed for each clone in both the commensal and pathogenic strains (0–2 nonsynonymous SNPs). The authors reached an interesting conclusion, stating that 'sampling of one colony would be enough for surveillance, outbreak investigations and clonal evolution', although there is overwhelming evidence to support the opposite. Evidence from their own study showed that among other clones, a variation in gene content of 2–15 genes was detected for all clones—which would have been unnoticeable had they isolated only a single colony from each sample.

We recently surveyed *E. coli* in healthy children (less than 5 years) from a rural Gambian community and assessed the relative contribution of independent immigration and establishment of strains vs. within-host evolution to the generation of within-host diversity in the healthy gut (Foster-Nyarko *et al.* 2021). Our results showed a dominance of independent immigration and establishment of strains, with several strains harboring putative virulence factors commonly associated with Enteropathogenic and Enteroaggregative *E. coli*—thus, adding to the body of evidence indicating a blurring of the lines of demarcation between resident strains and their pathogenic variants, while also adding significantly to the number of whole-genome sequences from commensal *E. coli* isolates, particularly from sub-Saharan Africa.

Prospective studies of travellers to low-income countries have provided vital insight into the acquisition and loss of strains, particularly, MDR clones. In a recent report (Foster-Nyarko *et al.* 2021), Kantele *et al.* investigated the dynamics of intestinal carriage of MDR bacteria during travel to a low-income country *via* daily stool sampling for the duration of their visit—unlike most previous studies, which only involved before and after samples. The study involved participants attending a medical course in Vientiane, Laos, together but staying at three different hotels, visiting several restaurants together or in small groups, and participating in daily rounds at the local hospitals. Remarkably, all the study participants acquired ESBL-producing bacteria at least at one point during the study—dominated by *E. coli*, with 95% of the participants acquiring between two and seven strains and multiple transient colonization events, with 20% of the participants sharing some common strains over the study period.

In contrast to the abundance of studies in humans, few studies have investigated the dynamics of *E. coli* residency in nonhuman mammals. In a temporal survey of *E. coli* residency in 54 mountain brushtail possums sampled on four occasions over the course of a year, Blyton *et al.* reported phylogroup B2 to dominate the resident *E. coli* population (Blyton *et al.* 2013). This study concluded that *E. coli* was rapidly gained and lost among these nonhuman mammals, with just 36% of resident strains recovered at one timepoint recovered at the subsequent timepoint. Recent data from our lab involving a range of diurnal monkey species in the Gambia also showed a dominance of *E. coli* belonging to phylogroup B2 (Foster-Nyarko *et al.* 2020a), including a prevalence of well-characterized lineages associated with human extraintestinal infections or AMR. However, the strains that occurred in these monkeys were genetically divergent from those sourced from hu-

mans, with little AMR among the simian isolates. Several monkeys were colonized by two or more strains, with the evidence of strain-sharing among monkeys sampled from the same park.

Data from poultry further support the observation that co-colonization with multiple strains is prevalent in nonhuman vertebrate hosts, reflecting the findings of previous studies in humans. Li *et al.* (2019) sampled > 100 colonies per caecum sample from nine mcr-1-positive broiler chickens from three provinces in China (a total of 962 *E. coli* isolates). A high rate of co-colonization (3–9 STs per chicken) was observed, with several birds harboring one to five Inc type plasmids encoding mrc-1—this high level of heterogeneity facilitating the transmission of mcr-1 among these chickens. The authors concluded that the 'gut is a 'melting pot' for active horizontal transfer of the mcr-1 gene'. We recently investigated the within-host diversity of *E. coli* from backyard poultry (10 backyard chickens and nine guinea fowl) drawn from rural Gambia, and found similarly high colonization with multiple strains in noncommercial birds (Foster-Nyarko *et al.* 2020b). Worryingly, a dominance of ST155—a well-characterized multidrug resistant clone—was observed in poultry *E. coli* isolates from Africa and South America. However, the isolates from these backyard birds carried few ESBL genotypes compared to what was found among other isolates from the African continent (Foster-Nyarko *et al.* 2020b).

In summary, the body of evidence shows that colonization with multiple strains is common in the vertebrate gut, with the rapid gain and loss of strains contributing to the within-host diversity of *E. coli*. At the same time, some clones can establish themselves in the gut for extended periods. As evidenced by the existing data, this level of diversity positions the gut as a hot spot for the active horizontal transfer of AMR and virulence traits within and between strains. Other vertebrate hosts may carry lineages well-established in humans—ndicating bacteria that have probably existed in the specific host species for long periods and passed down through vertical transmission. The direct contribution of nonhuman vertebrate hosts to the spread of AMR and virulence traits in humans remains to be established, despite the prevalence of isolates with similar profiles as their close relatives circulating in humans. Significant gaps remain in our understanding of the ecology of this familiar ally and foe of the vertebrate host, key examples of which are outlined below.

Conclusions and perspectives

It has been said: 'All cell biologists have at least two cells of interest: the one they are studying and *E. coli*' (Neidhardt *et al.* 1996). Yet, in the 137 years since *E. coli* was first described, most studies on the ecology of *E. coli* have been biased to pathogenic strains. Only a handful of studies have focused on the diversity of *E. coli* populations in the healthy human gut, particularly in the post-PCR era. Even fewer studies have explored the within-host diversity of non-pathogenic *E. coli* in nonhuman vertebrates in the post-PCR period, leaving much to be learned in terms of the resident populations of *E. coli* in the healthy vertebrate gut. However, such studies are vital to our understanding of the evolution of *E. coli* in the health and disease of both humans and animals.

An exploration of the commensal population of *E. coli* in nonhuman vertebrate hosts such as poultry and nonhuman primates is timely and relevant for several reasons. First, emerging infections are often linked with pathogens (such as *E. coli*) that inhabit both humans and nonhuman vertebrates and can cross the species barrier, potentially causing zoonotic as well as anthroponotic infections. Understanding the diversity within the commensal population would provide vital insights into the evolution of pathogens and AMR within this ecological context. Second, understanding the within-host evolution of *E. coli* during health may also inform new therapeutic approaches that exploit our understanding of the gut microbiome, such as faecal microbiome transplantation, probiotic development, and the control of foodborne diseases.

In particular, *E. coli* diversity studies based on whole-genome sequence data from sub-Saharan Africa are scarce. Very few studies have examined the population of *E. coli* among healthy individuals in this setting, particularly in healthy children. However, given the increased exposure to the environment, unsanitary conditions, and proximity to animals especially in rural areas, such studies are likely to yield crucial insights into the dynamics of strain turnover and residency of *E. coli* in the gut, ultimately facilitating our understanding of infection by this organism.

Furthermore, the bulk of existing surveys of *E. coli* in poultry have been biased towards the surveillance of AMR in commercial birds; thus, we know very little of the dynamics of colonization and the diversity of *E. coli* in this niche. Also, backyard birds, which are kept in proximity to humans, have been largely neglected in ecological studies. However, given the frequent contact with humans, studies involving backyard poultry might yield critical insights into the potential for the dissemination of strains and AMR between the two host species.

The direct contribution of human-habituated animals to *E. coli* transmission remains unknown. To address this would require future studies incorporating samples from whole households and animals living in proximity. The inclusion of nonhuman primate species in future studies of resident *E. coli* diversity would elucidate the full scale of the impact of human encroachment on natural habitats. The recent advances in Illumina sequencing platforms twinned with the potential of long-read sequencing (such as Pacific Biosciences sequencing and Oxford Nanopore technology) provide a timely opportunity to apply these approaches to these lines of study.

Outstanding questions that need addressing in future studies include: what factors govern co-existence or competition between the different *E. coli* strains co-colonizing the vertebrate gut? Do they exploit different microgeographical or nutrient niches in the gut? What is the importance, if any, of flagellar motility? How far do within-species antibacterial factors (e.g. colicins and type VI secretions systems) play a role in competitive exclusion? Can differences in the distribution of *E. coli* strains *in vivo* be accounted for by behaviors *in vitro* (e.g. competitive growth under different nutrient conditions)?

Biolog assays (Biolog Inc., Hayward, CA) are one way to screen for the key metabolic differences between different strains that may account for their co-existence or competitive exclusion within the gut. Candidate strains could then be selected for onward *in vitro* competition assays via individual and competitive fitness experiments and indirect antagonism assays as have been used by Durso *et al.* to assess fitness differences and competition between commensal strains and O157:H7 strains (Durso *et al.* 2004). Applikon Biotechnology (https://www.applikon-biotechnology.com/en/products/cultivation-systems/micro-matrix/) have designed mini-bioreactors that can be used for multiple competition assays in parallel, using 24-well microtiter plates with built-in controls for pH, temperature, and dissolved oxygen. Such experiments may shed light on the physiological and biochemical properties underlying the co-colonization of different strains in the vertebrate gut.

Acknowledgments

The authors are grateful to Professor Alan McNally and Dr Mark Webber for critically reviewing the thesis chapter that served as the starting material for this manuscript during E.F.-N's PhD viva and encouraging publishing it as a review article.

Conflicts of interest. None declared.

References

Abia ALK, Ubomba-Jaswa E. Dirty money on holy ground: isolation of potentially pathogenic bacteria and fungi on money collected from church offerings. *Iran J Public Health* 2019;**48**:849–57.

Achtman M, Heuzenroeder M, Kusecek B *et al.* Clonal analysis of *Escherichia coli* O2:K1 isolated from diseased humans and animals. *Infect Immun* 1986;**51**:268–76.

Aijuka M, Santiago AE, Girón JA *et al.* Enteroaggregative *Escherichia coli* is the predominant diarrheagenic *E . coli* pathotype among irrigation water and food sources in South Africa. *Int J Food Microbiol* 2018;**278**:44–51.

Akoachere JF, Gaelle N, Dilonga HM *et al.* Public health implications of contamination of Franc CFA (XAF) circulating in Buea (Cameroon) with drug resistant pathogens. *BMC Res Notes* 2014;**7**:16.

Albert MJ, Alam K, Islam M *et al.* *Hafnia alvei*, a probable cause of diarrhea in humans. *Infect Immun* 1991;**59**:1507–13.

Alhashash F, Weston V, Diggle M *et al.* Multidrug-resistant *Escherichia coli* bacteremia. *Emerg Infect Dis* 2013;**19**:1699–701.

Alm EW, Walk ST, Gordon DM.. In:Walk,ST,Feng,PCH (eds). *Population Genetics of Bacteria: A tribute to Thomas S. Whittam*. Washington, DC: ASM Press, 2011; Chapter 6, 67–89.

Altenhoefer A, Oswald S, Sonnenborn U *et al.* The probiotic *Escherichia coli* strain Nissle 1917 interferes with invasion of human intestinal epithelial cells by different enteroinvasive bacterial pathogens. *FEMS Immunol Med Microbiol* 2004;**40**:223–9.

Andersson JO, Andersson SG. Insights into the evolutionary process of genome degradation. *Curr Opin Genet Dev* 1999;**9**:664–71.

Bäckhed F, Ley RE, Sonnenburg JL *et al.* Host-bacterial mutualism in the human intestine. *Science* 2005;**307**:1915–20.

Baeshen MN, Al-Hejin AM, Bora RS *et al.* Production of biopharmaceuticals in E. coli: Current scenario and future perspectives. *J Microbiol Biotechnol* 2015;**25**:953–62.

Banerjee R, Johnson JR A new clone sweeps clean: the enigmatic emergence of *Escherichia coli* sequence type 131. *Antimicrob Agents Chemother* 2014;**58**:4997–5004.

Beghain J, Bridier-Nahmias A, Le Nagard H *et al.* ClermonTyping: an easy-to-use and accurate in silico method for Escherichia genus strain phylotyping. *Microb Genomics* 2018;**4**:e000192.

Ben Zakour NL, Alsheikh-Hussain AS, Ashcroft MM *et al.* Sequential acquisition of virulence and fluoroquinolone resistance has shaped the evolution of *Escherichia coli* ST131. *Mbio* 2016;**7**:e00347–00316.

Bennani H, Mateus A, Mays N *et al.* Overview of evidence of antimicrobial use and antimicrobial resistance in the food chain. *Antibiotics* 2020;**9**:49.

Benz F, Huisman JS, Bakkeren E *et al.* Plasmid- and strain-specific factors drive variation in ESBL-plasmid spread in vitro and in vivo. *ISME J* 2021;**15**:862–78.

Berg RD. The indigenous gastrointestinal microflora. *Trends Microbiol* 1996;**4**:430–5.

Bertin Y, Girardeau JP, Chaucheyras-Durand F *et al.* Enterohaemorrhagic *Escherichia coli* gains a competitive advantage by using ethanolamine as a nitrogen source in the bovine intestinal content. *Environ Microbiol* 2011;**13**:365–77.

Bettelheim KA, Faiers M, Shooter RA. Serotypes of *Escherichia coli* in normal stools. *Lancet* 1972;**2**:1223–4.

Bettelheim KA. Commemoration of the publication 100 years ago of the papers by Dr. Th. Escherich in which are described for the first time the organisms that bear his name. *Zentralbl Bakteriol Mikrobiol Hyg A* 1986;**261**:255–65.

Bielaszewska M, Schiller R, Lammers L *et al.* Heteropathogenic virulence and phylogeny reveal phased pathogenic metamorphosis in *Escherichia coli* O2:H6. *EMBO Mol Med* 2014;**6**:347–57.

Blattner FR, Plunkett G, Bloch CA *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* 1997;**277**:1453–62.

Blattner FR. Biological frontiers. *Science* 1983;**222**:719–20.

Blount ZD, Barrick JE, Davidson CJ *et al.* Genomic analysis of a key innovation in an experimental *Escherichia coli* population. *Nature* 2012;**489**:513–8.

Blount ZD. The unexhausted potential of *E. coli*. *Elife* 2015;**4**:e05826.

Blum SE, Heller ED, Sela S *et al.* Genomic and phenomic study of mammary pathogenic *Escherichia coli*. *PLoS One* 2015;**10**:e0136387.

Blyton MD, Banks SC, Peakall R *et al.* High temporal variability in commensal *Escherichia coli* strain communities of a herbivorous marsupial. *Environ Microbiol* 2013;**15**:2162–72.

Bongers JH, Franssen F, Elbers AR *et al.* Antimicrobial resistance of *Escherichia coli* isolates from the faecal flora of veterinarians with different professional specialties. *Vet Q* 1995;**17**:146–9.

Borges CA, Tarlton NJ, Riley LW. *Escherichia coli* from commercial broiler and backyard chickens share sequence types, antimicrobial resistance profiles, and resistance genes with human extraintestinal pathogenic *Escherichia coli*. *Foodborne Pathog Dis* 2019;**16**:813–22.

Boudeau J, Glasser AL, Julien S *et al.* Inhibitory effect of probiotic *Escherichia coli* strain Nissle 1917 on adhesion to and invasion of intestinal epithelial cells by adherent-invasive *E. coli* strains isolated from patients with Crohn's disease. *Aliment Pharmacol Ther* 2003;**18**:45–56.

Bourque R, Lallier R, Larivière S. Influence of oral antibiotics on resistance and enterotoxigenicity of *Escherichia coli*. *Can J Comp Med* 1980;**44**:101–8.

Branche WC, Young VM, Robinet HG *et al.* Effect of colicine production on *Escherichia coli* in the normal human intestine. *Exp Biol Med* 1963;**114**:198–201.

Brenner DJ, Davis BR, Steigerwalt AG *et al.* Atypical biogroups of *Escherichia coli* found in clinical specimens and description of *Escherichia hermannii* sp. nov. *J Clin Microbiol* 1982a;**15**:703–13.

Brenner DJ, McWhorter AC, Knutson JK *et al.* *Escherichia vulneris*: a new species of *Enterobacteriaceae* associated with human wounds. *J Clin Microbiol* 1982b;**15**:1133–40.

Brown AC, Valiere A. Probiotics and medical nutrition therapy. *Nutr Clin Care* 2004;**7**:56–68.

Brown L, Langelier C, Reid MJ *et al.* Antimicrobial resistance: a call to action! *Clin Infect Dis* 2017;**64**:106–7.

Bryant J, Chewapreecha C, Bentley SD. Developing insights into the mechanisms of evolution of bacterial pathogens from whole-genome sequences. *Fut Microbiol* 2012;**7**:1283–96.

Buffie CG, Jarchum I, Equinda M *et al.* Profound alterations of intestinal microbiota following a single dose of clindamycin results in sustained susceptibility to *Clostridium difficile*-induced colitis. *Infect Immun* 2012;**80**:62–73.

Buffie CG, Pamer EG. Microbiota-mediated colonization resistance against intestinal pathogens. *Nat Rev Immunol* 2013;**13**:790–801.

Burgess NR, McDermott SN, Whiting J. Aerobic bacteria occurring in the hind-gut of the cockroach, *Blatta orientalis*. *J Hyg* 1973;**71**:1–7.

Burmeister AR. Horizontal gene transfer. *Evol Med Public Health* 2015;**2015**:193–4.

Cao Q, Didelot X, Wu Z *et al.* Progressive genomic convergence of two *Helicobacter pylori* strains during mixed infection of a patient with chronic gastritis. *Gut* 2015;**64**:554–61.

Capaldi RA, Schulenberg B, Murray J *et al.* Cross-linking and electron microscopy studies of the structure and functioning of the *Escherichia coli* ATP synthase. *J Exp Biol* 2000;**203**:29–33.

Castanon JIR. History of the use of antibiotic as growth promoters in European poultry feeds. *Poult Sci* 2007;**86**:2466–71.

Castellani A, Chalmers A. *J. Manual of Tropical Medicine*. 3rd edn. London: Baillière, Tindall and Cox; 1919, 941.

Caugant DA, Levin BR, Selander RK. Genetic diversity and temporal variation in the *E. coli* population of a human host. *Genetics* 1981;**98**:467–90.

Chang DE, Smalley DJ, Tucker DL *et al.* Carbon nutrition of *Escherichia coli* in the mouse intestine. *Proc Natl Acad Sci USA* 2004;**101**:7427–32.

Charlesworth B. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. *Nat Rev Genet* 2009;**10**:195–205.

Chaudhuri RR, Henderson IR. The evolution of the *Escherichia coli* phylogeny. *Infect Genet Evol* 2012;**12**:214–26.

Chaudhuri RR, Sebaihia M, Hobman JL *et al.* Complete genome sequence and comparative metabolic profiling of the prototypical enteroaggregative *Escherichia coli* strain 042. *PLoS ONE* 2010;**5**:e8801.

Chaumeil PA, Mussig AJ, Hugenholtz P *et al.* GTDB-Tk: a toolkit to classify genomes with the genome taxonomy database. *Bioinformatics* 2019;**36**:1925–27.

Chen SL, Wu M, Henderson JP *et al.* Genomic diversity and fitness of *E. coli* strains recovered from the intestinal and urinary tracts of women with recurrent urinary tract infection. *Sci Transl Med* 2013;**5**:184ra160.

Cheng G, Dai M, Ahmed S *et al.* Antimicrobial drugs in fghting against antimicrobial resistance. *Front Microbiol* 2016;**7**:470.

Claesson MJ, O'Sullivan O, Wang Q *et al.* Comparative analysis of pyrosequencing and a phylogenetic microarray for exploring microbial community structures in the human distal intestine. *PLoS ONE* 2009;**4**:e6669.

Claus SP, Guillou H, Ellero-Simatos S. The gut microbiota: a major player in the toxicity of environmental pollutants? *NPJ Biofilms Microbiomes* 2016;**2**:16003.

Clermont O, Christenson JK, Denamur E *et al.* The clermont *Escherichia coli* phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. *Environ Microbiol Rep* 2013;**5**:58–65.

Clermont O, Gordon DM, Brisse S *et al.* Characterization of the cryptic *Escherichia* lineages: rapid identification and prevalence. *Environ Microbiol* 2011;**13**:2468–77.

Codling C, O'Mahony L, Shanahan F *et al.* A molecular analysis of fecal and mucosal bacterial communities in irritable bowel syndrome. *Dig Dis Sci* 2010;**55**:392–7.

Cointe A, Birgy A, Bridier-Nahmias A *et al. Escherichia coli* O80 hybrid pathotype strains producing Shiga toxin and ESBL: molecular characterization and potential therapeutic options. *J Antimicrob Chemother* 2020;**75**:537–42.

Conway T, Cohen PS. Commensal and pathogenic *Escherichia coli* metabolism in the gut. *Microbiol Spectr* 2015;**3**:10.

Cooper TF. Recombination speeds adaptation by reducing competition between beneficial mutations in populations of *Escherichia coli*. *PLoS Biol* 2007;**5**:e225.

Cornelissen M, Pasternak AO, Grijsen ML *et al.* HIV-1 dual infection is associated with faster CD4+ T-cell decline in a cohort of men with primary HIV infection. *Clin Infect Dis* 2012;**54**:539–47.

Crick FH, Barnett L, Brenner S *et al.* General nature of the genetic code for proteins. *Nature* 1961;**192**:1227–32.

Croxall G, Hale J, Weston V *et al.* Molecular epidemiology of extraintestinal pathogenic *Escherichia coli* isolates from a regional cohort of elderly patients highlights the prevalence of ST131 strains with increased antimicrobial resistance in both community and hospital care settings. *J Antimicrob Chemother* 2011;**66**:2501–8.

Croxen MA, Finlay BB. Molecular mechanisms of *Escherichia coli* pathogenicity. *Nat Rev Microbiol* 2010;**8**:26–38.

Damborg P, Nielsen SS, Guardabassi L. *Escherichia coli* shedding patterns in humans and dogs: insights into within-household transmission of phylotypes associated with urinary tract infections. *Epidemiol Infect* 2009;**137**:1457–64.

Datta N. Drug resistance and R factors in the bowel bacteria of London patients before and after admission to hospital. *BMJ* 1969;**2**:407–11.

David LA, Maurice CF, Carmody RN *et al.* Diet rapidly and reproducibly alters the human gut microbiome. *Nature* 2014;**505**:559–63.

Davies M, Stewart PR. Transferable drug resistance in man and animals: genetic relationship between R-plasmids in enteric bacteria from man and domestic pets. *Aust Vet J* 1978;**54**:507–12.

de Kraker ME, Jarlier V, Monen JC *et al.* The changing epidemiology of bacteraemias in Europe: trends from the European Antimicrobial Resistance Surveillance System. *Clin Microbiol Infect* 2013;**19**:860–8.

de la Cruz F, Davies J. Horizontal gene transfer and the origin of species: lessons from bacteria. *Trends Microbiol* 2000;**8**:128–33.

De Rauw K, Thiry D, Caljon B *et al.* Characteristics of Shiga toxin producing- and enteropathogenic *Escherichia coli* of the emerging serotype O80:H2 isolated from humans and diarrhoeic calves in Belgium. *Clin Microbiol Infect* 2019;**25**:111.

Decano AG, Downing T. An *Escherichia coli* ST131 pangenome atlas reveals population structure and evolution across 4,071 isolates. *Sci Rep* 2019;**9**:17394.

Denamur E, Clermont O, Bonacorsi S *et al.* The population genetics of pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2020; **19**:37–54.

Derakhshandeh A, Eraghi V, Boroojeni AM *et al.* Virulence factors, antibiotic resistance genes and genetic relatedness of commensal *Escherichia coli* isolates from dogs and their owners. *Microb Pathog* 2018;**116**:241–5.

Deriu E, Liu JZ, Pezeshki M *et al.* Probiotic bacteria reduce *Salmonella typhimurium* intestinal colonization by competing for iron. *Cell Host Microbe* 2013;**14**:26–37.

Dethlefsen L, Huse S, Sogin ML *et al.* The pervasive effects of an antibiotic on the human gut microbiota, as revealed by deep 16S rRNA sequencing. *PLoS Biol* 2008;**6**:e280.

Díaz-Jiménez D, García-Meniño I, Herrera A *et al.* Genomic characterization of *Escherichia coli* isolates belonging to a new hybrid aEPEC/ExPEC pathotype O153:H10-A-ST10 eae-beta1 occurred in meat, poultry, wildlife and human diarrheagenic samples. *Antibiotics* 2020;**9**:192.

Didelot X, Méric G, Falush D *et al.* Impact of homologous and non-homologous recombination in the genomic evolution of *Escherichia coli*. *BMC Genomics* 2012;**13**:256.

Didelot X, Nell S, Yang I *et al.* Genomic evolution and transmission of *Helicobacter pylori* in two South African families. *Proc Natl Acad Sci* 2013;**110**:13880–5.

Didelot X, Walker AS, Peto TE *et al.* Within-host evolution of bacterial pathogens. *Nat Rev Microbiol* 2016;**14**:150–62.

Dixit OVA, O'Brien CL, Pavli P *et al.* Within-host evolution versus immigration as a determinant of *Escherichia coli* diversity in the human gastrointestinal tract. *Environ Microbiol* 2018;**20**:993–1001.

Dogan B, Rishniw M, Bruant G *et al.* Phylogroup and lpfA influence epithelial invasion by mastitis associated *Escherichia coli*. *Vet Microbiol* 2012;**159**:163–70.

Dolfing J. Syntrophy in microbial fuel cells. *ISME J* 2014;**8**:4–5.

Dunn SJ, Connor C, McNally A. The evolution and transmission of multi-drug resistant *Escherichia coli* and *Klebsiella pneumoniae:* the complexity of clones and plasmids. *Curr Opin Microbiol* 2019;**51**:51–56.

Duriez P, Clermont O, Bonacorsi S *et al.* Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. *Microbiology* 2001;**147**: 1671–6.

Durso LM, Smith D, Hutkins RW. Measurements of fitness and competition in commensal *Escherichia coli* and *E . coli* O157:H7 strains. *Appl Environ Microbiol* 2004;**70**:6466–72.

Dutta S, Pazhani GP, Nataro JP *et al.* Heterogenic virulence in a diarrheagenic *Escherichia coli*: evidence for an EPEC expressing heat-labile toxin of ETEC. *Int J Med Microbiol* 2015;**305**:47–54.

Ellis EL, Delbrück M. The growth of bacteriophage. *J Gen Physiol* 1939;**22**:365–84.

Ercumen A, Pickering AJ, Kwong LH *et al.* Animal feces contribute to domestic fecal contamination: evidence from *E . coli* measured in water, hands, food, flies, and soil in Bangladesh. *Environ Sci Technol* 2017;**51**:8725–34.

Escherich T, Pfaundler M. *Bacterium coli commune*. In: Kolle W, Wassermann A (eds). *Handbuch der pathogenen mikroorganismen*, Vol **2**. Jena: Verlag von Gustav Fischer, 1903, 334–474.

Escherich T. *Die darmbakterien des däuglings und ihre beziehungen zur physiologie der verdauung*. Stuttgart: Verlag von Ferdinand Enke, 1886.

Escherich T. The intestinal bacteria of the neonate and breast-fed infant. 1884. *Clin Infect Dis* 1988;**10**:1220–5.

Escherich T. Ueber die bacterien des milchkotes. *Ærztliches Intelligenz-Blatt, Münchner Medicinische Wochenschrift* 1888;**32**:243.

Escobar-Páramo P, Clermont O, Blanc-Potard AB *et al.* A specific genetic background is required for acquisition and expression of virulence factors in *Escherichia coli*. *Mol Biol Evol* 2004a;**21**: 1085–94.

Escobar-Páramo P, Giudicelli C, Parsot C *et al.* The evolutionary history of *Shigella* and enteroinvasive *Escherichia coli* revised. *J Mol Evol* 2003;**57**:140–8.

Escobar-Páramo P, Grenet K, Le Menac'h A *et al.* Large-scale population structure of human commensal *Escherichia coli* isolates. *Appl Environ Microbiol* 2004b;**70**:5698–700.

Ewers C, Grobbel M, Stamm I *et al.* Emergence of human pandemic O25:H4-ST131 CTX-M-15 extended-spectrum-β-lactamase-producing *Escherichia coli* among companion animals. *J Antimicrob Chemother* 2010;**65**:651–60.

Fabich AJ, Jones SA, Chowdhury FZ *et al.* Comparison of carbon nutrition for pathogenic and commensal *Escherichia coli* strains in the mouse intestine. *Infect Immun* 2008;**76**:1143–52.

Fabich AJ, Leatham MP, Grissom JE *et al.* Genotype and phenotypes of an intestine-adapted *Escherichia coli* K-12 mutant selected by animal passage for superior colonization. *Infect Immun* 2011;**79**: 2430–9.

Farmer JJ, Fanning GR, Davis BR *et al. Escherichia fergusonii* and *Enterobacter taylorae*, two new species of *Enterobacteriaceae* isolated from clinical specimens. *J Clin Microbiol* 1985;**21**:77–81.

Flint HJ. The impact of nutrition on the human microbiome. *Nutr Rev* 2012;**70 Suppl 1**:S10–13.

Foster-Nyarko E, Alikhan NF, Ikumapayi UN *et al.* Genomic diversity of *Escherichia coli* from healthy children in rural Gambia. *PeerJ* 2021;**9**:e10572.

Foster-Nyarko E, Alikhan NF, Ravi A *et al.* Genomic diversity of Escherichia coli from non-human primates in the Gambia. *Microb Genom* 2020a;**7**:mgen000484.

Foster-Nyarko E, Alikhan NF, Ravi A *et al.* Genomic diversity of *Escherichia coli* from non-human primates in the Gambia. *Microb Genom* 2020b;**6**:mgen000428.

Foxman B. The epidemiology of urinary tract infection. *Nat Rev Urol* 2010;**7**:653–60.

Frank C, Werber D, Cramer JP *et al.* Epidemic profile of Shiga-toxin-producing *Escherichia coli* O104:H4 outbreak in Germany. *N Engl J Med* 2011;**365**:1771–80.

Freter R, Brickner H, Botney M *et al.* Mechanisms that control bacterial populations in continuous-flow culture models of mouse large intestinal flora. *Infect Immun* 1983a;**39**:676–85.

Freter R, Brickner H, Fekete J *et al.* Survival and implantation of *Escherichia coli* in the intestinal tract. *Infect Immun* 1983b;**39**:686–703.

Friedmann H. Escherich and *Escherichi a*. *EcoSal Plus* 2014;**6**. DOI: 10.1128/ecosalplus.ESP-0025-2013. Reprinted from *Adv Appl Microbiol* 2014;60:133-196.

Friedmann HC. From "butyribacterium" to "*E. coli*": an essay on unity in biochemistry. *Perspect Biol Med* 2004;**47**:47–66.

Fukiya S, Mizoguchi H, Tobe T *et al.* Extensive genomic diversity in pathogenic *Escherichia coli* and *Shigella* strains revealed by comparative genomic hybridization microarray. *J Bacteriol* 2004;**186**: 3911–21.

Gati NS, Middendorf-Bauchart B, Bletz S *et al.* Origin and evolution of hybrid Shiga toxin-producing and uropathogenic *Escherichia coli* strains of sequence type 141. *J Clin Microbiol* 2019;**58**:e01309–19.

Gauger EJ, Leatham MP, Mercado-Lubo R *et al.* Role of motility and the flhDC Operon in *Escherichia coli* MG1655 colonization of the mouse intestine. *Infect Immun* 2007;**75**:3315–24.

Ghoul M, Mitri S. The ecology and evolution of microbial competition. *Trends Microbiol* 2016;**24**:833–45.

Gilroy R, Ravi A, Getino M *et al.* Extensive microbial diversity within the chicken gut microbiome revealed by metagenomics and culture. *PeerJ* 2021;**9**:e10941.

Gioia-Di Chiacchio RM, Cunha MPV, de Sá LRM *et al.* Novel hybrid of typical enteropathogenic *Escherichia coli* and Shiga-toxin-producing *E . coli* (tEPEC/STEC) emerging from pet birds. *Front Microbiol* 2018;**9**:2975.

Golubchik T, Batty EM, Miller RR *et al.* Within-host evolution of *Staphylococcus aureus* during asymptomatic aarriage. *PLoS ONE* 2013;**8**:e61319.

Good BH, McDonald MJ, Barrick JE *et al.* The dynamics of molecular evolution over 60,000 generations. *Nature* 2017;**551**:45–50.

Gorbach SL. Probiotics and gastrointestinal health. *Am J Gastroenterol* 2000;**95**:S2–4.

Gordon DM, Cowling A. The distribution and genetic structure of *Escherichia coli* in Australian vertebrates: host and geographic effects. *Microbiology (Reading)* 2003;**149**:3575–86.

Gordon DM, O'Brien CL, Pavli P. *Escherichia coli* diversity in the lower intestinal tract of humans. *Environ Microbiol Rep* 2015;**7**:642–8.

Gough E, Shaikh H, Manges AR. Systematic review of intestinal microbiota transplantation (fecal bacteriotherapy) for recurrent *Clostridium difficile* infection. *Clin Infect Dis* 2011;**53**:994–1002.

Granato ET, Meiller-Legrand TA, Foster KR. The evolution and ecology of bacterial warfare. *Curr Biol* 2019;**29**:R521–37.

Guinée P, Ugueto N, van Leeuwen N. Escherichia coli with resistance factors in vegetarians, babies, and nonvegetarians. *Appl Microbiol* 1970;**20**:531–5.

Gyles CL, Palchaudhuri S, Maas WK. Naturally occurring plasmid carrying genes for enterotoxin production and drug resistance. *Science* 1977;**198**:198–9.

Hacker J, Carniel E. Ecological fitness, genomic islands and bacterial pathogenicity. A Darwinian view of the evolution of microbes. *EMBO Rep* 2001;**2**:376–81.

Halachev MR, Chan JZM, Constantinidou CI *et al.* Genomic epidemiology of a protracted hospital outbreak caused by multidrug-resistant *Acinetobacter baumannii* in Birmingham, England. *Genom Med* 2014;**6**:70.

Hall MD, Holden MT, Srisomang P *et al.* Improved characterisation of MRSA transmission using within-host bacterial sequence diversity. *Elife* 2019;**8**:e46402.

Hancock V, Dahl M, Klemm P. Probiotic *Escherichia coli* strain Nissle 1917 outcompetes intestinal pathogens during biofilm formation. *J Med Microbiol* 2010;**59**:392–9.

Hardin G. The competitive exclusion principle. *Science* 1960;**131**:1292–7.

Harshey RM, Matsuyama T. Dimorphic transition in *Escherichia coli* and Salmonella typhimurium: surface-induced differentiation into hyperflagellate swarmer cells. *Proc Natl Acad Sci USA* 1994;**91**:8631–5.

Hata H, Natori T, Mizuno T *et al.* Phylogenetics of family *Enterobacteriaceae* and proposal to reclassify *Escherichia hermannii* and *Salmonella subterranea* as *Atlantibacter hermannii* and *Atlantibacter subterranea* gen. nov., comb. nov. *Microbiol Immunol* 2016;**60**:303–11.

Hayashi K, Morooka N, Yamamoto Y *et al.* Highly accurate genome sequences of *Escherichia coli* K-12 strains MG1655 and W3110. *Mol Syst Biol* 2006;**2**:2006.0007.

Hayashi T, Makino K, Ohnishi M *et al.* Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and genomic comparison with a laboratory strain K-12. *DNA Res* 2001;**8**:11–22.

Herranz M, Pole I, Ozere I *et al.* *Mycobacterium tuberculosis* acquires limited genetic diversity in prolonged infections, reactivations and transmissions involving multiple hosts. *Front Microbiol* 2017;**8**:2661.

Hildebrand A, Schlacta T, Warmack R *et al.* Engineering *Escherichia coli* for improved ethanol production from gluconate. *J Biotechnol* 2013;**168**:101–6.

Hill C, Guarner F, Reid G *et al.* Expert consensus document. The International Scientific Association for probiotics and prebiotics consensus statement on the scope and appropriate use of the term probiotic. *Nat Rev Gastroenterol Hepatol* 2014;**11**:506–14.

Hooper LV, Macpherson AJ. Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat Rev Immunol* 2010;**10**:159–69.

Hu Y, Coates AR. Transposon mutagenesis identifies genes which control antimicrobial drug tolerance in stationary-phase *Escherichia coli*. *FEMS Microbiol Lett* 2005;**243**:117–24.

Huang CJ, Lin H, Yang X. Industrial production of recombinant therapeutics in *Escherichia coli* and its recent advancements. *J Ind Microbiol Biotechnol* 2012;**39**:383–99.

Huys G, Cnockaert M, Janda JM *et al.* *Escherichia albertii* sp. nov., a diarrhoeagenic species isolated from stool specimens of Bangladeshi children. *Int J Syst Evol Microbiol* 2003;**53**:807–10.

Hyma KE, Lacher DW, Nelson AM *et al.* Evolutionary genetics of a new pathogenic *Escherichia* species: *Escherichia albertii* and related *Shigella boydii* strains. *J Bacteriol* 2005;**187**:619–28.

Iguchi A, Iyoda S, Terajima J *et al.* Spontaneous recombination between homologous prophage regions causes large-scale inversions within the *Escherichia coli* O157:H7 chromosome. *Gene* 2006;**372**:199–207.

Ingle DJ, Clermont O, Skurnik D *et al.* Biofilm formation by and thermal niche and virulence characteristics of *Escherichia* spp. *Appl Environ Microbiol* 2011;**77**:2695–700.

Inoue T, Shingaki R, Hirose S *et al.* Genome-wide screening of genes required for swarming motility in *Escherichia coli* K-12. *J Bacteriol* 2007;**189**:950–7.

Jacob F, Perrin D,Sanchez C *et al.* [Operon: a group of genes with the expression coordinated by an operator]. *C R Hebd Seances Acad Sci* 1960;**250**:1727–9.

Jacobi CA, Malfertheiner P. *Escherichia coli* Nissle 1917 (Mutaflor): new insights into an old probiotic bacterium. *Dig Dis* 2011;**29**:600–7.

Jain C, Rodriguez-R LM, Phillippy AM *et al.* High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 2018;**9**:5114.

Johnson JR, Clabots C, Kuskowski MA. Multiple-host sharing, long-term persistence, and virulence of *Escherichia coli* clones from human and animal household members. *J Clin Microbiol* 2008a;**46**:4078–82.

Johnson JR, Clabots C. Sharing of virulent *Escherichia coli* clones among household members of a woman with acute cystitis. *Clin Infect Dis* 2006;**43**:e101–108.

Johnson JR, Miller S, Johnston B *et al.* Sharing of *Escherichia coli* sequence type ST131 and other multidrug-resistant and urovirulent *E . coli* strains among dogs and cats within a household. *J Clin Microbiol* 2009;**47**:3721–5.

Johnson JR, Owens K, Gajewski A *et al.* *Escherichia coli* colonization patterns among human household members and pets, with attention to acute urinary tract infection. *J Infect Dis* 2008b;**197**:218–24.

Judicial Commission of the International Committee on Bacterial Nomenclature. Conservation of the family name *Enterobacteriaceae*, of the name of the type genus, and designation of the type species. Opinion No. 15. *Int Bull Bacteriol Nomencl Taxon* 1958;**8**:73–4.

Kadooka Y, Sato M, Imaizumi K *et al.* Regulation of abdominal adiposity by probiotics (*Lactobacillus gasseri* SBT2055) in adults with obese tendencies in a randomized controlled trial. *Eur J Clin Nutr* 2010;**64**:636–43.

Kamada N, Chen GY, Inohara N *et al.* Control of pathogens and pathobionts by the gut microbiota. *Nat Immunol* 2013;**14**:685–90.

Kang JH, Yun SI, Park MH *et al.* Anti-obesity effect of *Lactobacillus gasseri* BNR17 in high-sucrose diet-induced obese mice. *PLoS ONE* 2013;**8**:e54617.

Kaper JB, Nataro JP, Mobley HL. Pathogenic *Escherichia coli*. *Nat Rev Microbiol* 2004;**2**:123–40.

Karch H, Denamur E, Dobrindt U *et al.* The enemy within us: lessons from the 2011 European *Escherichia coli* O104:H4 outbreak. *EMBO Mol Med* 2012;**4**:841–8.

Kaup B. Bringer-Meyer S Sahm H Metabolic engineering of *Escherichia coli*: construction of an efficient biocatalyst for D-mannitol formation.

Kay GL, Sergeant MJ, Zhou Z *et al.* Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. *Nat Commun* 2015;**6**:6717.

Kennemann L, Didelot X, Aebischer T *et al.* *Helicobacter pylori* genome evolution during human infection. *Proc Natl Acad Sci USA* 2011;**108**:5033–8.

Kessler R, Nisa S, Hazen TH *et al.* Diarrhea, bacteremia and multiorgan dysfunction due to an extraintestinal pathogenic *Escherichia coli* strain with enteropathogenic *E. coli* genes. *Pathog Dis* 2015;**73**:ftv076.

Kim B, Park H, Na D *et al.* Metabolic engineering of *Escherichia coli* for the production of phenol from glucose. *Biotechnol J* 2014;**9**:621–9.

Knudsen PK, Brandtzaeg P, Høiby EA *et al.* Impact of extensive antibiotic treatment on faecal carriage of antibiotic-resistant enterobacteria in children in a low resistance prevalence setting. *PLoS ONE* 2017;**12**:e0187618.

Knudsen PK, Gammelsrud KW, Alfsnes K *et al.* Transfer of a *blaCTX-M-1*-carrying plasmid between different *Escherichi coli* strains within the human gut explored by whole genome sequencing analyses. *Sci Rep* 2018;**8**:280.

Kosek M, Bern C, Guerrant RL. The global burden of diarrhoeal disease, as estimated from studies published between 1992 and 2000. *Bull World Health Organ* 2003;**81**:197–204.

Köser CU, Holden MTG, Ellington MJ *et al.* Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med* 2012;**366**:2267–75.

Kuo CH, Moran NA, Ochman H. The consequences of genetic drift for bacterial genome complexity. *Genome Res* 2009;**19**:1450–4.

Lan R, Alles MC, Donohoe K *et al.* Molecular evolutionary relationships of enteroinvasive *Escherichia coli* and *Shigella* spp. *Infect Immun* 2004;**72**:5080–8.

Lawrence JG, Ochman H, Hartl DL. Molecular and evolutionary relationships among enteric bacteria. *J Gen Microbiol* 1991;**137**:1911–21.

Lawrence JG, Ochman H. Amelioration of bacterial genomes: rates of change and exchange. *J Mol Evol* 1997;**44**:383–97.

Lawrence JG, Ochman H. Molecular archaeology of the *Escherichia coli* genome. *Proc Natl Acad Sci USA* 1998;**95**:9413–7.

Lawrence JG, Ochman H. Reconciling the many faces of lateral gene transfer. *Trends Microbiol* 2002;**10**:1–4.

Lazarus B, Paterson DL, Mollinger JL *et al.* Do human extraintestinal *Escherichia coli* infections resistant to expanded-spectrum cephalosporins originate from food-producing animals? A systematic review. *Clin Infect Dis* 2015;**60**:439–52.

Le Bouguénec C, Schouler C. Sugar metabolism, an additional virulence factor in enterobacteria. *Int J Med Microbiol* 2011;**301**:1–6.

Leatham MP, Banerjee S, Autieri SM *et al.* Precolonized human commensal *Escherichia coli* strains serve as a barrier to *E . coli* O157:H7 growth in the streptomycin-treated mouse intestine. *Infect Immun* 2009;**77**:2876–86.

Leatham MP, Stevenson SJ, Gauger EJ *et al.* Mouse intestine selects nonmotile flhDC mutants of *Escherichia coli* MG1655 with increased colonizing ability and better utilization of carbon sources. *Infect Immun* 2005;**73**:8039–49.

Leatham-Jensen MP, Frimodt-Moller J, Adediran J *et al.* The streptomycin-treated mouse intestine selects *Escherichia coli* envZ missense mutants that interact with dense and diverse intestinal microbiota. *Infect Immun* 2012;**80**:1716–27.

Lederberg J, Lederberg EM. Replica plating and indirect selection of bacterial mutants. *J Bacteriol* 1952;**63**:399–406.

Lee H, Popodi E, Tang H *et al.* Rate and molecular spectrum of spontaneous mutations in the bacterium *Escherichia coli* as determined by whole-genome sequencing. *Proc Natl Acad Sci USA* 2012;**109**:E2774–2783.

Lee JW. Protocol measuring horizontal gene transfer from algae to non-photosynthetic organisms. *MethodsX* 2019;**6**:1564–74.

Lehman IR, Bessman MJ, Simms ES *et al.* Enzymatic synthesis of deoxyribonucleic acid. I. Preparation of substrates and partial purification of an enzyme from *Escherichia coli*. *J Biol Chem* 1958;**233**:163–70.

Leimbach A, Hacker J, Dobrindt U. *E. coli* as an all-rounder: the thin line between commensalism and pathogenicity. In: Dobrindt U, Hacker JH, Svanborg C (eds). *Between Pathogenicity and Commensalism*. Berlin Heidelberg: Springer, 2013, 3–32.

Lenski RE. Experimental studies of pleiotropy and epistasis in *Escherichia coli*. I. Variation in competitive fitness among mutants resistant to virus T4. *Evolution* 1988;**42**:425–32.

Lescat M, Clermont O, Woerther PL *et al.* Commensal *Escherichia coli* strains in Guiana reveal a high genetic diversity with host-dependant population structure. *Environ Microbiol Rep* 2013;**5**:49–57.

Ley RE, Hamady M, Lozupone C *et al.* Evolution of mammals and their gut microbes. *Science* 2008;**320**:1647–51.

Ley SD, de Vos M, Van Rie A *et al.* Deciphering within-host microevolution of *Mycobacterium tuberculosis* through whole-genome sequencing: the phenotypic impact and way forward. *Microbiol Mol Biol Rev* 2019;**83**. DOI: 10.1128/MMBR.00062-18.

Li SP, Tan J, Yang X *et al.* Niche and fitness differences determine invasion success and impact in laboratory bacterial communities. *ISME J* 2019;**13**:402–12.

Lidin-Janson G, Kaijser B, Lincoln K *et al.* The homogeneity of the faecal coliform flora of normal school-girls, characterized by serological and biochemical properties. *Med Microbiol Immunol (Berl)* 1978;**164**:247–53.

Linn S, Arber W. Host specificity of DNA produced by *Escherichia coli*, X. In vitro restriction of phage fd replicative form. *Proc Natl Acad Sci USA* 1968;**59**:1300–6.

Linton KB, Lee PA, Richmond MH *et al.* Antibiotic resistance and transmissible R-factors in the intestinal coliform flora of healthy adults and children in an urban and a rural community. *J Hyg* 1972;**70**:99–104.

Liu S, Feng J, Pu J *et al.* Genomic and molecular characterisation of *Escherichia marmotae* from wild rodents in Qinghai-Tibetplateau as a potential pathogen. *Sci Rep* 2019;**9**:10619.

Liu S, Jin D, Lan R *et al. Escherichia marmotae* sp. nov., isolated from faeces of *Marmota himalayana*. *Int J Syst Evol Microbiol* 2015;**65**:2130–4.

Liu T, Khosla C. Genetic engineering of *Escherichia coli* for biofuel production. *Annu Rev Genet* 2010;**44**:53–69.

Lodinová-Zádniková R, Sonnenborn U. Effect of preventive administration of a nonpathogenic *Escherichia coli* strain on the colonization of the intestine with microbial pathogens in newborn infants. *Biol Neonate* 1997;**71**:224–32.

Lu S, Jin D, Wu S *et al.* 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. *Emerg Microbes Infect* 2016;**5**:e122.

Lukjancenko O, Wassenaar TM, Ussery DW. Comparison of 61 sequenced *Escherichia coli* genomes. *Microb Ecol* 2010;**60**:708–20.

Luria SE, Delbrück M. Mutations of bacteria from virus sensitivity to virus resistance. *Genetics* 1943;**28**:491–511.

Lwoff A. Lysogeny. *Bacteriol Rev* 1953;**17**:269–337.

Lynch SV, Boushey HA. The microbiome and development of allergic disease. *Curr Opin Allergy Clin Immunol* 2016;**16**:165–71.

Macfarlane GT, Macfarlane S. Human colonic microbiota: ecology, physiology and metabolic potential of intestinal bacteria. *Scand J Gastroenterol* 1997;**222**:3–9.

Mainil J. *Escherichia coli* virulence factors. *Vet Immunol Immunopathol* 2013;**152**:2–12..

Maltby R, Leatham-Jensen MP, Gibson T *et al.* Nutritional basis for colonization resistance by human commensal *Escherichia coli* strains HS and Nissle 1917 against *E . coli* O157:H7 in the mouse intestine. *PLoS ONE* 2013;**8**:e53957.

Manges AR, Steiner TS, Wright AJ. Fecal microbiota transplantation for the intestinal decolonization of extensively antimicrobial-resistant opportunistic pathogens: a review. *Infect Dis* 2016;**48**:587–92.

Mariani-Kurkdjian P, Lemaître C, Bidet P *et al.* Haemolytic-uraemic syndrome with bacteraemia caused by a new hybrid *Escherichia coli* pathotype. *New Microbes New Infect* 2014;**2**:127–31.

Marinus MG. DNA methylation and mutator genes in *Escherichia coli* K-12. *Mutat Res* 2010;**705**:71–6.

Martinson JNV, Pinkham NV, Peters GW *et al.* Rethinking gut microbiome residency and the Enterobacteriaceae in healthy human adults. *ISME J* 2019;**13**.2306–18

Martinson JNV, Walk ST. Residency in the gut of healthy human adults. *EcoSal Plus* 2020;**9**:10.

Marvig RL, Johansen HK, Molin S *et al.* Genome analysis of a transmissible lineage of *Pseudomonas aeruginosa* reveals pathoadaptive mutations and distinct evolutionary paths of hypermutators. *PLos Genet* 2013;**9**:e1003741.

Massot M, Daubié AS, Clermont O *et al.* Phylogenetic, virulence and antibiotic resistance characteristics of commensal strain populations of *Escherichia coli* from community subjects in the Paris area in 2010 and evolution over 30 years. *Microbiology* 2016;**162**:642–50.

Mather AE, Harris SR Playing fast and loose with mutation. *Nat Rev Microbiol* 2013;**11**:822.

Mau B, Glasner JD, Darling AE *et al.* Genome-wide detection and analysis of homologous recombination among sequenced strains of *Escherichia coli*. *Genome Biol* 2006;**7**:R44.

McNally A, Kallonen T, Connor C *et al.* Diversification of colonization factors in a multidrug-resistant *Escherichia coli* lineage evolving under negative frequency-dependent selection. *Mbio* 2019;**10**. DOI: 10.1128/mBio.00644-19.

McNally A, Oren Y, Kelly D *et al.* Combined analysis of variation in cre, accessory and regulatory genome regions provides a super-resolution view into the evolution of bacterial populations. *PLos Genet* 2016;**12**:e1006280.

Meador JP, Caldwell ME, Cohen PS *et al. Escherichia coli* pathotypes occupy distinct niches in the mouse intestine. *Infect Immun* 2014;**82**:1931–8.

Meselson M, Yuan R. DNA restriction enzyme from E. coli. *Nature* 1968;**217**:1110–4.

Mitsuoka T, Ohno K, Benno Y *et al.* [The fecal flora of man. IV. Communication: comparison of the newly developed method with the old conventional method for the analysis of intestinal flora (author's transl)]. *Zentralbl Bakteriol Orig A* 1976;**234**:219–33.

Miyoshi M, Ogawa A, Higurashi S *et al.* Anti-obesity effect of *Lactobacillus gasseri* SBT2055 accompanied by inhibition of pro-inflammatory gene expression in the visceral adipose tissue in diet-induced obese mice. *Eur J Nutr* 2014;**53**:599–606.

Morrow LE, Wischmeyer P. Blurred lines: dysbiosis and probiotics in the ICU. *Chest* 2017;**151**:492–9.

Muegge BD, Kuczynski J, Knights D *et al.* Diet drives convergence in gut microbiome functions across mammalian phylogeny and within humans. *Science* 2011;**332**:970–4.

Mueller NT, Bakacs E, Combellick J *et al.* The infant microbiome development: mom matters. *Trends Mol Med* 2015;**21**:109–17.

Muto A, Osawa S. The guanine and cytosine content of genomic DNA and bacterial evolution. *Proc Natl Acad Sci USA* 1987;**84**:166–9.

Naylor SW, Low JC, Besser TE *et al.* Lymphoid follicle-dense mucosa at the terminal rectum is the principal site of colonization of enterohemorrhagic *Escherichia coli* O157:H7 in the bovine host. *Infect Immun* 2003;**71**:1505–12.

Neidhardt FC, Eingraham JL, Low KB (eds). *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. Washington, DC: ASM Press; 1996.

Ng KM, Ferreyra JA, Higginbottom SK *et al.* Microbiota-liberated host sugars facilitate post-antibiotic expansion of enteric pathogens. *Nature* 2013;**502**:96–9.

Nicholson BA, West AC, Mangiamele P *et al.* Genetic characterization of expec-like virulence plasmids among a subset of NMEC. *PLoS ONE* 2016;**11**:e0147757.

Nicholson JK, Holmes E, Kinross J *et al.* Host-gut microbiota metabolic interactions. *Science* 2012;**336**:1262–7.

Nielsen KL, Dynesen P, Larsen P *et al.* Faecal *Escherichia coli* from patients with *E . coli* urinary tract infection and healthy controls who have never had a urinary tract infection. *J Med Microbiol* 2014;**63**:582–9.

Nijsten R, London N, van den Bogaard A *et al.* In-vitro transfer of antibiotic resistance between faecal *Escherichia coli* strains isolated from pig farmers and pigs. *J Antimicrob Chemother* 1996;**37**:1141–54.

Nissle A. [Explanations of the significance of colonic dysbacteria & the mechanism of action of *E . coli* therapy (mutaflor)]. *Medizinische* 1959;**4**:1017–22.

Nissle A. [Mutaflor and its medical significance]. *Z Klin Med* 1951;**2**:68.

Nissle A. [Old and new experiences on therapeutic successes by restoration of the colonic flora with mutaflor in gastrointestinal diseases]. *Med Welt* 1961;**29-30**:1519–23.

Nowrouzian F, Adlerberth I, Wold AE. P fimbriae, capsule and aerobactin characterize colonic resident *Escherichia coli*. *Epidemiol Infect* 2001;**126**:11–8.

Nüesch-Inderbinen M, Cernela N, Wüthrich D *et al.* Genetic characterization of Shiga toxin producing *Escherichia coli* belonging to the emerging hybrid pathotype O80:H2 isolated from humans 2010-2017 in Switzerland. *Int J Med Microbiol* 2018;**308**:534–8.

Nwankwo EO, Ekwunife N, Mofolorunsho KC. Nosocomial pathogens associated with the mobile phones of healthcare workers in a hospital in Anyigba, Kogi state, Nigeria. *J Epidemiol Glob Health* 2014;**4**:135–40.

Ocejo M, Tello M, Oporto B *et al.* Draft genome sequence of *Escherichia marmotae* E690, isolated from beef cattle. *Microbiol Resour Announc* 2020;**9**:e00739–20.

Ochman H, Jones IB. Evolutionary dynamics of full genome content in *Escherichia coli*. *EMBO J* 2000;**19**:6637–43.

Ochman H, Lawrence JG, Groisman EA. Lateral gene transfer and the nature of bacterial innovation. *Nature* 2000;**405**:299–304.

Orskov F, Orskov I. *Escherichia coli* serotyping and disease in man and animals. *Can J Microbiol* 1992;**38**:699–704.

Pal S, Juyal D, Adekhandi S *et al.* Mobile phones: reservoirs for the transmission of nosocomial pathogens. *Adv Biomed Res* 2015;**4**:144.

Pallen MJ, Telatin A, Oren A. The next million names for Archaea and Bacteria. *Trends Microbiol* 2020; **29**:289–298.

Pamer EG. Resurrecting the intestinal microbiota to combat antibiotic-resistant pathogens. *Science* 2016;**352**:535–8.

Peigne C, Bidet P, Mahjoub-Messai F *et al.* The plasmid of *Escherichia coli* strain S88 (O45:K1:H7) that causes neonatal meningitis is closely related to avian pathogenic *E . coli* plasmids and is associated with high-level bacteremia in a neonatal rat meningitis model. *Infect Immun* 2009;**77**:2272–84.

Penders J, Thijs C, Vink C *et al.* Factors influencing the composition of the intestinal microbiota in early infancy. *Pediatrics* 2006;**118**:511–21.

Pennisi E. Laboratory workhorse decoded. *Science* 1997;**277**:1432–4.

Pereira FC, Berry D. Microbial nutrient niches in the gut. *Environ Microbiol* 2017;**19**:1366–78.

Periwal V, Scaria V. Insights into structural variations and genome rearrangements in prokaryotic genomes. *Bioinformatics* 2015;**31**:1–9.

Perna NT, Plunkett G, Burland V *et al.* Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* 2001;**409**:529–33.

Picard B, Garcia JS, Gouriou S *et al.* The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. *Infect Immun* 1999;**67**:546–53.

Pitout JD. Extraintestinal pathogenic *Escherichia coli*: a combination of virulence with antibiotic resistance. *Front Microbiol* 2012;**3**:9.

Poulsen LK, Lan F, Kristensen CS *et al.* Spatial distribution of *Escherichia coli* in the mouse large intestine inferred from rRNA in situ hybridization. *Infect Immun* 1994;**62**:5191–4.

Poulsen LK, Licht TR, Rang C *et al.* Physiological state of *Escherichia coli* BJ4 growing in the large intestines of streptomycin-treated mice. *J Bacteriol* 1995;**177**:5840–5.

Priest FG, Barker M. Gram-negative bacteria associated with brewery yeasts: reclassification of *Obesumbacterium proteus* biogroup 2 as *Shimwellia pseudoproteus* gen. nov., sp. nov., and transfer of *Escherichia blattae* to *Shimwellia blattae* comb. nov. *Int J Syst Evol Microbiol* 2010;**60**:828–33.

Pupo GM, Lan R, Reeves PR. Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. *Proc Natl Acad Sci USA* 2000;**97**:10567–72.

Quince C, Ijaz UZ, Loman N *et al.* Extensive modulation of the fecal metagenome in children with Crohn's disease during exclusive enteral nutrition. *Am J Gastroenterol* 2015;**110**:1718–29; quiz 1730.

Raeside C, Gaffé J, Deatherage DE *et al.* Large chromosomal rearrangements during a long-term evolution experiment with *Escherichia coli*. *Mbio* 2014;**5**:e01377–01314.

Rasko DA, Rosovitz MJ, Myers GS *et al.* The pangenome structure of *Escherichia coli*: comparative genomic analysis of *E . coli* commensal and pathogenic isolates. *J Bacteriol* 2008;**190**:6881–93.

Ravi A, Halstead FD, Bamford A *et al.* Loss of microbial diversity and pathogen domination of the gut microbiota in critically ill patients. *Microb Genom* 2019;**5**:e000293.

Reeves PR, Liu B, Zhou Z *et al.* Rates of mutation and host transmission for an *Escherichia coli* clone over 3 years. *PLoS ONE* 2011;**6**:e26907.

Ribeiro CFA, Silveira GGOS, Cândido ES *et al.* Effects of antibiotic treatment on gut microbiota and how to overcome its negative impacts on human health. *ACS Infect Dis* 2020;**6**:2544–59.

Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. *Proc Natl Acad Sci USA* 2009;**106**:19126–31.

Richter TKS, Hazen TH, Lam D *et al.*. Temporal variability of *Escherichia coli* diversity in the gastrointestinal tracts of Tanzanian children with and without exposure to antibiotics. *mSphere* 2018;**3**:e00558–18.

Richter TKS, Michalski JM, Zanetti L *et al.* Responses of the human gut *Escherichia coli* population to pathogen and antibiotic disturbances. *mSystems* 2018;**3**. DOI: 10.1128/mSystems.00047-18.

Riley LW. Distinguishing pathovars from nonpathovars: *Escherichia coli*. *Microbiol Spectr* 2020;**8**:33385193.

Robinet HG. Relationship of host antibody to fluctuations of *Escherichia coli* serotypes in the human intestine. *J Bacteriol* 1962;**84**:896–901.

Robins-Browne RM, Holt KE, Ingle DJ *et al.* Are *Escherichia coli* pathotypes still relevant in the era of whole-genome sequencing? *Front Cell Infect Microbiol* 2016;**6**:141.

Rodgers K, McVey M. Error-prone repair of DNA double-strand breaks. *J Cell Physiol* 2016;**231**:15–24.

Round JL, Mazmanian SK. The gut microbiota shapes intestinal immune responses during health and disease. *Nat Rev Immunol* 2009;**9**:313–23.

Schlager TA, Hendley JO, Bell AL *et al.* Clonal diversity of *Escherichia coli* colonizing stools and urinary tracts of young girls. *Infect Immun* 2002;**70**:1225–9.

Sears HJ, Brownlee I, Uchiyama JK. Persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *J Bacteriol* 1950;**59**:293–301.

Sears HJ, Brownlee I. Further observations on the persistence of individual strains of *Escherichia coli* in the intestinal tract of man. *J Bacteriol* 1952;**63**:47–57.

Selander RK, Levin BR. Genetic diversity and structure in *Escherichia coli* populations. *Science* 1980;**210**:545–7.

Sheldon IM, Rycroft AN, Dogan B *et al.* Specific strains of *Escherichia coli* are pathogenic for the endometrium of cattle and cause pelvic inflammatory disease in cattle and mice. *PLoS One* 2010;**5**:e9192.

Shooter RA, Bettleheim KA, Lennox-King SM *et al.* Escherichia coli serotypes in the faeces of healthy adults over a period of several months. *J Hyg* 1977;**78**:95–8.

Shulman ST, Friedmann HC, Sims RH. Theodor Escherich: the first pediatric infectious diseases physician? *Clin Infect Dis* 2007;**45**:1025–9.

Sivan A, Corrales L, Hubert N *et al.* Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science* 2015;**350**:1084–9.

Skerman V, McGowan V, Sneath P. Approved lists of bacterial names. *Int J Syst Evol Microbiol* 1980;**30**:225–420.

Skurnik D, Bonnet D, Bernède-Bauduin C *et al.* Characteristics of human intestinal *Escherichia coli* with changing environments. *Environ Microbiol* 2008;**10**:2132–7.

Slanetz LW, Bartley CH. Numbers of enterococci in water, sewage, and feces determined by the membrane filter technique with an improved medium. *J Bacteriol* 1957;**74**:591–5.

Smith HL. Zur kenntnis der colibacillen des sa¨uglingsstuhles. *Zentr Bakt Parasitenk I Orig* 1899;**25**:689–93.

Sorbara MT, Pamer EG. Interbacterial mechanisms of colonization resistance and the strategies pathogens use to overcome them. *Mucosal Immunol* 2019;**12**:1–9.

Soysal N, Mariani-Kurkdjian P, Smail Y *et al.* Enterohemorrhagic *Escherichia coli* hybrid pathotype O80:H2 as a new therapeutic challenge. *Emerg Infect Dis* 2016;**22**:1604–12.

Ssengooba W, de Jong BC, Joloba ML *et al.* Whole genome sequencing reveals mycobacterial microevolution among concurrent isolates from sputum and blood in HIV infected TB patients. *BMC Infect Dis* 2016;**16**:371.

Stegger M, Leihof RF, Baig S *et al.* A snapshot of diversity: intraclonal variation of *Escherichia coli* clones as commensals and pathogens. *Int J Med Microbiol* 2020;**310**:151401.

Steinsland H, Lacher DW, Sommerfelt H *et al.* Ancestral lineages of human enterotoxigenic *Escherichia coli*. *J Clin Microbiol* 2010;**48**:2916–24.

Stevens A. Incorporation of the adenine ribonucleotide into RNA by cell fractions from E. coli. *Biochem Biophys Res Commun* 1960;**3**:92–6.

Stoesser N, Sheppard AE, Moore CE *et al.* Extensive within-host diversity in fecally carried extended-spectrum-beta-lactamase-producing *Escherichia coli* isolates: implications for transmission analyses. *J Clin Microbiol* 2015;**53**:2122–31.

Stoesser N, Sheppard AE, Pankhurst L *et al.* Evolutionary history of the global emergence of the *Escherichia coli* epidemic clone ST131. *Mbio* 2016;**7**:e02162.

Suvarna K, Stevenson D, Meganathan R *et al.* Menaquinone (vitamin K2) biosynthesis: localization and characterization of the *menA* gene from *Escherichia coli*. *J Bacteriol* 1998;**180**:2782–7.

Swartz JR. Advances in *Escherichia coli* production of therapeutic proteins. *Curr Opin Biotechnol* 2001;**12**:195–201.

Syed SA, Abrams GD, Freter R. Efficiency of various intestinal bacteria in assuming normal functions of enteric flora after association with germ-free mice. *Infect Immun* 1970;**2**:376–86.

Tanaka M, Nakayama J. Development of the gut microbiota in infancy and its impact on health in later life. *Allergol Int* 2017;**66**:515–22.

Tannock GW, Tiong IS, Priest P *et al.* Testing probiotic strain *Escherichia coli* Nissle 1917 (Mutaflor) for its ability to reduce carriage of multidrug-resistant *E. coli* by elderly residents in long-term care facilities. *J Med Microbiol* 2011;**60**:366–70.

Tenaillon O, Skurnik D, Picard B *et al.* The population genetics of commensal *Escherichia coli*. *Nat Rev Microbiol* 2010;**8**:207–17.

Tettelin H, Masignani V, Cieslewicz MJ *et al.* Genome analysis of multiple pathogenic isolates of *Streptococcus agalactiae*: implications for the microbial "pan-genome". *Proc Natl Acad Sci USA* 2005;**102**:13950–5.

Tettelin H, Medini D. *The Pangenome: Diversity, Dynamics and Evolution of Genomes*. Cham: Springer, 2020.

Teuber M. Spread of antibiotic resistance with food-borne pathogens. *Cell Mol Life Sci* 1999;**56**:755–63.

The HC, Thanh DP, Holt KE *et al.* The genomic signatures of *Shigella* evolution, adaptation and geographical spread. *Nat Rev Microbiol* 2016;**14**:235–50.

Thiry D, Saulmont M, Takaki S *et al.* Enteropathogenic *Escherichia coli* O80:H2 in young calves with diarrhea, Belgium. *Emerg Infect Dis* 2017;**23**:2093–5.

Timmerman HM, Niers LE, Ridwan BU *et al.* Design of a multispecies probiotic mixture to prevent infectious complications in critically ill patients. *Clin Nutr* 2007;**26**:450–9.

Toombs-Ruane LJ, Benschop J, French NP *et al.* Carriage of extended-spectrum-beta-lactamase- and AmpC beta-lactamase-producing *Escherichia coli* strains from humans and pets in the same households. *Appl Environ Microbiol* 2020;**86**. DOI: 10.1128/AEM.01613-20.

Totsuka K. Studien über *Bacterium coli*. *Z Hyg Infektionskrankh* 1902;**45**:115–24.

Touchon M, Hoede C, Tenaillon O *et al.* Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. *PLos Genet* 2009;**5**:e1000344.

Trauner A, Liu Q, Via LE *et al.* The within-host population dynamics of *Mycobacterium tuberculosis* vary with treatment efficacy. *Genome Biol* 2017;**18**:71.

Tuncil YE, Xiao Y, Porter NT *et al.* Reciprocal prioritization to dietary glycans by gut bacteria in a competitive environment promotes stable coexistence. *Mbio* 2017;**8**:e01068–17.

Turnbaugh PJ, Ridaura VK, Faith JJ *et al.* The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci Transl Med* 2009;**1**:6ra14.

Tzabar Y, Pennington TH. The population structure and transmission of *Escherichia coli* in an isolated human community; studies on an Antarctic base. *Epidemiol Infect* 1991;**107**:537–42.

Van Boeckel TP, Brower C, Gilbert M *et al.* Global trends in antimicrobial use in food animals. *Proc Natl Acad Sci* 2015;**112**:5649.

van der Putten BCL, Matamoros S, Mende DR *et al. Escherichia ruysiae* sp. nov., a novel Gram-stain-negative bacterium, isolated from a faecal sample of an international traveller. *Int J Syst Evol Microbiol* 2021;**71**:004609.

van Elsas JD, Semenov AV, Costa R *et al.* Survival of *Escherichia coli* in the environment: fundamental and public health aspects. *ISME J* 2011;**5**:173–83.

Ventola CL. The antibiotic resistance crisis: part 1: causes and threats. *P T* 2015;**40**:277–83.

Venturi A, Gionchetti P, Rizzello F *et al.* Impact on the composition of the faecal flora by a new probiotic preparation: preliminary data on maintenance treatment of patients with ulcerative colitis. *Aliment Pharmacol Ther* 1999;**13**:1103–8.

Walk ST, Alm EW, Gordon DM *et al.* Cryptic lineages of the genus *Escherichi a*. *Appl Environ Microbiol* 2009;**75**:6534–44.

Wallick H, Stuart CA. Antigenic relationships of *Escherichia coli* isolated from one individual. *J Bacteriol* 1943;**45**:121–6.

Wang S, Wu C, Shen J *et al.* Hypermutable *Staphylococcus aureus* strains present at high frequency in subclinical bovine mastitis isolates are associated with the development of antibiotic resistance. *Vet Microbiol* 2013;**165**:410–5.

Watanabe H, Wada A, Inagaki Y *et al.* Outbreaks of enterohaemorrhagic *Escherichia coli* O157:H7 infection by two different genotype strains in Japan, 1996. *Lancet* 1996;**348**:831–2.

Webber MA, Piddock LJV. . In: *Frontiers in Antimicrobial Resistance: a Tribute to Stuart B Levy*. Washington, DC: ASM Press, 2005, 374–86.

Wehkamp J, Harder J, Wehkamp K *et al.* NF-kappaB- and AP-1-mediated induction of human beta defensin-2 in intestinal epithelial cells by *Escherichia coli* Nissle 1917: a novel effect of a probiotic bacterium. *Infect Immun* 2004;**72**:5750–8.

Welch RA, Burland V, Plunkett G *et al.* Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. *Proc Natl Acad Sci USA* 2002;**99**:17020–4.

Wiedemann B, Habermann R, Knothe H *et al.* [Stablility of *Escherichia coli*-flora in healthy men. 3. Occurrence of permanent and transitory strains in infants]. *Arch Hyg Bakteriol* 1971;**154**:581–9.

Wiser MJ, Ribeck N, Lenski RE. Long-term dynamics of adaptation in asexual populations. *Science* 2013;**342**:1364–7.

Worby CJ, Lipsitch M, Hanage WP. Within-host bacterial diversity hinders accurate reconstruction of transmission networks from genomic distance data. *PLoS Comput Biol* 2014;**10**:e1003549.

Wu GD, Chen J, Hoffmann C *et al.* Linking long-term dietary patterns with gut microbial enterotypes. *Science* 2011;**334**:105–8.

Wymant C, Hall M, Ratmann O *et al.* PHYLOSCANNER: inferring transmission from within- and between-host pathogen genetic diversity. *Mol Biol Evol* 2018;**35**:719–33.

Xu J. Microbial ecology in the age of genomics and metagenomics: concepts, tools, and recent advances. *Mol Ecol* 2006;**15**:1713–31.

Yar DD. Bacterial contaminants and antibiogram of Ghana paper currency notes in circulation and their associated health risks in asante-mampong, Ghana. *Int J Microbiol* 2020;**2020**:8833757.

Youngster I, Russell GH, Pindar C *et al.* Oral, capsulized, frozen fecal microbiota transplantation for relapsing *Clostridium difficile* infection. *JAMA* 2014;**312**:1772–8.

Zhou Z, Alikhan NF, Mohamed K *et al.* The enterobase user's guide, with case studies on *Salmonella* transmission, *Yersinia pestis* phylogeny, and *Escherichia* core genomic diversity. *Genome Res* 2020;**30**:138–52.

Zogg AL, Zurfluh K, Schmitt S *et al.* Antimicrobial resistance, multilocus sequence types and virulence profiles of ESBL producing and non-ESBL producing uropathogenic *Escherichia coli* isolated from cats and dogs in Switzerland. *Vet Microbiol* 2018;**216**:79–84.

Zyrek AA, Cichon C, Helms S *et al.* Molecular mechanisms underlying the probiotic effects of *Escherichia coli* Nissle 1917 involve ZO-2 and PKCzeta redistribution resulting in tight junction and epithelial barrier repair. *Cell Microbiol* 2007;**9**:804–16.