frontiers | Frontiers in Bacteriology

TYPE Review
PUBLISHED 09 October 2025
DOI 10.3389/fbrio.2025.1677775

OPEN ACCESS

EDITED BY
Zhuo Ma,
Albany College of Pharmacy and Health Sciences, United States

REVIEWED BY
Biswanath Jana,
Washington University in St. Louis, United States
Janani Prahlad,
University of Minnesota, United States

*CORRESPONDENCE
Khuthadzo L. Mudau
50173065@mylife.unisa.ac.za

RECEIVED 01 August 2025
ACCEPTED 19 September 2025
PUBLISHED 09 October 2025

CITATION
Mudau KL, Ntobeng LR, Kalu CM and Tekere M (2025) Pathogenicity and virulence factors of *Escherichia coli* discovered using next generation sequencing technologies and proteomics. *Front. Bacteriol.* 4:1677775. doi: 10.3389/fbrio.2025.1677775

Pathogenicity and virulence factors of *Escherichia coli* discovered using next generation sequencing technologies and proteomics

Khuthadzo L. Mudau*, Lesoka R. Ntobeng, Chimdi M. Kalu and Memory Tekere

Department of Environmental Science, College of Agriculture and Environmental Sciences, University of South Africa, Florida, South Africa

Escherichia coli is a gastrointestinal bacterium previously known for its commensal activities in the human digestive systems. Their occurrence in drinking water and natural water sources has been used as a faecal pollution footprint or marker to determine the extent of pollution. However, their ability to cause diseases as an opportunistic bacterium is a global concern. Hence, unveiling their diverse virulence factors and pathogenicity through diverse technologies becomes pertinent. The advent of next-generation sequencing technologies and proteomics have significantly propelled these studies forward. Utilizing next-generation sequencing and proteomics, scientists have unveiled a multitude of pathogenicity and virulence factors linked to *E. coli*. This review underscores the advancements made in uncovering *E. coli's* pathogenicity, virulence factors, and specific attributes through next-generation sequencing and selected proteomics investigations. The review presents and describes discovered pathogenicity and virulence factors. It concludes that while significant progress has been made, there is still much work to be done that can utilize next-generation sequencing and proteomics in this area of research fully. The in-depth study of *E. coli's* virulence factors and pathogenicity could provide preventive/curative insight into a pattern or technologies that could be adopted to minimize the outbreak of disease associated with the bacterium even at their opportunistic level.

KEYWORDS

Escherichia coli, pathogenicity, virulence, next-generation sequencing, proteomics, water quality

1 Introduction

Theodor Escherich first described *Escherichia coli* in 1885 (Escherich, 1885; Mueller and Tainter, 2023). There are hundreds of strains of the bacterium *E. coli* (Berthe et al., 2013; Mueller and Tainter, 2023). Many of the strains are typically benign and proliferate in human the digestive tracts. However, due to its virulence, *E. coli* can elude the host's immune system and become resistant to commonly prescribed antibiotics (CDC, 2016; Jang et al., 2017). Attributed to extensive characterization of its phenotypes and genotypes, *E. coli* has served as a valuable model organism for research and education since the 1940s.

Nine identified pathovars of *E. coli* strains isolated from humans can cause diarrheagenic and extraintestinal diseases (Donnenberg, 2013; Yang et al., 2017). Seven of these pathotypes are enteric pathogenic *E. coli*, including Enteropathogenic *E. coli* (EPEC), Enterohaemorrhagic *E. coli* (EHEC), Enterotoxigenic *E. coli* (ETEC), Enteroinvasive *E. coli* (EIEC), Enteroaggregative *E. coli* (EAEC), Diffusely adherent *E. coli* (DAEC), and a recently discovered pathotype, Adherent-Invasive *E. coli* (AIEC) (Herzog et al., 2019; Mandomando et al., 2020). These particular pathotypes are mainly responsible for triggering diarrhea and various intestinal disorders. For instance, Enterohemorrhagic *E. coli* (EHEC) pathotypes pose significant public health concerns as they are known foodborne pathogens and have been linked to fatal outbreaks in both developed and developing countries (Kaper et al., 2004; Donnenberg, 2013; Foster et al. 2015). These pathotypes cause diseases by expressing genes that encode virulence factors, and recent studies have emphasized their potential impact on a range of disorders (Jangid et al., 2024; Palmela et al., 2018; Desvaux et al., 2020).

Pathogenic microorganisms produce specialized molecules, predominantly proteins, known as virulence factorscontrolled by specific genes (Lewis et al., 2015; Tenaillon et al., 2023). Each type of pathogenic *E. coli* has its distinct pathogenicity mechanisms and virulence factor profile, which is determined by specific gene clusters. It is interesting to note that various enteric and extraintestinal pathotypes of *E. coli* share common virulence factors and strategies (Tenaillon et al., 2010; Habibi et al., 2017; Peng et al., 2024). The most common virulence factors and strategies shared by enteric and extraintestinal pathogenic *E. coli* includes iron acquisition systems (siderophores), adhesins, fimbriae (including P fimbriae), lipopolysaccharides (LPS), capsules, secretory IgA proteases and various toxins (such as hemolysin and cytotoxic necrotizing factor) (Clements et al., 2012; Lindstedt et al., 2018; Potgieter et al., 2018).

In recent years, research has identified several virulence factors associated with enteric *E. coli* pathotypes, which are implicated in intestinal and extraintestinal disorders (Lo et al., 2015; Pakbin et al., 2021). The use of NGS has significantly aided in the identification of *E. coli* virulence factors, allowing for improved diagnosis of clinical specimens, the distinction between different infections causing diseases, identification of the sources of illness, and the implementation of disease control measures to prevent further spread of diseases.

2 Pathogenicity and virulence factors of *E. coli*

E. coli naturally resides in the gastrointestinal tract (GIT) of humans and animals. Pathogenic *E. coli* can be classified into two main groups: diarrheagenic *E. coli* (DEC) and extra-intestinal pathogenic *E. coli* (ExPEC) (Desvaux et al., 2020). ExPEC is responsible for a variety of infections such as sepsis, neonatal meningitis, and urinary tract infections (UTI) due to the release of toxins (Daga et al., 2019; Paramita et al., 2020). According to Garcia et al. (2013), ExPEC secretes toxins such as hemolysin A (HlyA), cytotoxic necrotizing factor 1 (CNF 1), and cytolethal distending toxin (CDT). HlyA forms membrane pores to lyse erythrocytes and effector immune cells (Lerm et al., 1999; Hofman et al., 2000) and can also induce Ca^{2+} oscillations in renal epithelial cells, resulting in increased production of IL-6 and IL-8 (Uhlen et al., 2000).

There are six well-recognized pathotypes of DEC, including enterotoxigenic *E. coli* (ETEC), enteropathogenic *E. coli* (EPEC), enteroaggregative *E. coli* (EAEC), Shiga toxin (Stx)-producing *E. coli* (STEC), enteroinvasive *E. coli* (EIEC), and diffusely adherent *E. coli* (DAEC) (Martins et al., 2015; Tourret and Denamur, 2016; Clermont et al., 2019). These pathotypes are known as the primary causative agents of childhood and traveler's diarrhea (Martins et al., 2015; Tourret and Denamur, 2016; Clermont et al., 2019). EAEC has been traditionally recognized as an intestinal pathogen and is unlikely to cause disease in individuals outside of the intestinal environment (Mandomando et al., 2020; Meza-Segura et al., 2020). Additionally, EAEC produces various toxins, including the enteroaggregative heat-stable toxin 1 (EAST1) encoded by the heat-stable enterotoxins (ast) A gene, which is located adjacent to the plasmid-encoded toxin gene (Eichhorn et al., 2015). This toxin has been compared to the heat-stable *E. coli* STa enterotoxin, suggesting its role in causing secretory diarrhea. According to Dadonaite et al. (2018) and Roser and Ritchie (2021), diarrhea led to the deaths of at least 370,000 children in 2019, with 800,000fatalities per year based on 2013 data. While diarrhea is a global issue, its prevalence and impact are most pronounced in low-income countries. Many outbreaks are attributed to intestinal pathogenic *E. coli* strains, which exhibit distinct virulence traits, different O: H serotypes, and characteristic clinical syndromes (LeJeune et al., 2008; Rojas-Lopez et al., 2021). Nonetheless, these *E. coli* strains also share certain steps in the pathogenesis mechanism, including attachment to the intestinal mucosa and the possession of plasmids that encode virulence factors (Potgieter

and Pinto, 2019; Rojas-Lopez et al., 2021). A table summarising the virulence factors, the discovery methods, associated *E. coli* strains, the diseases they case, and their modes of action is presented in Supplementary Table 1.

3 Comparative genomics and molecular basis of pathogenicity and virulence in *E. coli*

The comparative genomics approach aims to systematically organize the major virulence factors in *E. coli*, providing a comprehensive and comparative understanding across different pathotypes, phylogroups, and virulence factor categories (Desvaux et al., 2020; Clark and Maresso, 2021). This organization offers valuable insights into vaccine targets categorized by phylogroup and also identifies alternative strategies to effectively counteract this pathogen from multiple perspectives (Kaper et al., 2004; Rojas-Lopez et al., 2018; Paramita et al., 2020). Currently, this is made possible through genome sequencing technologies, which generate extensive datasets revealing the genome structure and diverse genes present in *E. coli* (Clermont et al., 2019; Tenaillon et al., 2023). The information derived from genome sequencing and RNA sequencing projects enables a deeper understanding of the mechanisms utilized by *E. coli* to cause diseases in humans and facilitates the classification of pathogenic variants (pathovars) based on their virulence gene content (Hazen et al., 2017; Geurtsen et al., 2022). These projects have significantly advanced our understanding of the pathogenicity and virulence of *E. coli* in recent years. In a study by Pakbin et al. (2021), the virulence factors of enteric pathogenic *E. coli* were reviewed to gain insight into the diverse virulence factors associated with encoding genes used by different pathotypes of enteric pathogenic *E. coli* to cause intestinal and extraintestinal diseases in humans. The most notable outbreak involving a hybrid isolate of *E. coli* O104:H4 was discovered in 2011 in Europe (Beutin and Martin, 2012). For example, an *E. coli* outbreak in Germany in 2011, which affected over 3,400 people and led to 39 deaths, was attributed to a hybrid pathogenic *E. coli* (Bielaszewska et al., 2011). Through a combination of comparative genomics, transcriptomics, and functional characterization of virulence mechanisms, it was highlighted that four EPEC/ETEC isolates are likely from EPEC isolates that have acquired ETEC virulence genes via mobile genetic elements, most likely divergent plasmids.

For instance, human and cattle studies have identified *E. coli* isolates containing combinations of EPEC/ETEC and STEC/ETEC virulence genes (Dutta et al., 2015; Li et al., 2018). Previous reports have characterized hybrid *E. coli* isolates with a combination of canonical virulence genes from different *E. coli* pathovars. Among these isolates, three contained genes encoding the Shiga toxin of STEC and the heat-stable enterotoxin (ST) of ETEC (Franz et al., 2014; Li et al., 2018). Another study found an isolate containing the LEE region of EPEC and the LT genes of ETEC (Dutta et al., 2015). These findings underscore the limitations of simplistic pathovar definitions and indicate the presence of numerous circulating hybrid isolates. Phylogenomic comparisons have shown that E. coli isolates with identical virulence gene content can be found in different locations within the species' phylogenomic framework. While these phylogroups often consist of isolates from a single pathovar, they may also contain isolates from different pathovars, such as EPEC and ETEC (Franzin and Sircili, 2015; Iqbal et al., 2016). As a result, there is a growing need for greater utilization of next-generation sequencing (NGS) approaches and proteomic techniques to gain deeper insights and address the limitations associated with identifying diverse *E. coli* pathovars and understanding the connection between genes and associated pathogenicity and virulence.

4 Pathogenicity and virulence factors of *E. coli* discovered Using NGS technologies

Understanding and defining pathogenic microorganisms is crucial for effective disease treatment, recovery, and ensuring patient safety. Next-Generation Sequencing (NGS) has significantly contributed to the generation of extensive datasets of genomes and transcriptomes. With NGS, the complete DNA sequence of a bacterial genome can be determined in a single sequence run, providing valuable information on resistance, virulence, and typing, which is especially useful for outbreak investigations (Deurenberg et al., 2017). The availability of whole genome sequencing (WGS) from bacterial pathogens was hastened by notable infectious disease events, such as the cholera epidemic in Haiti following the 2010 earthquake (Barzilay et al., 2013) and the international *E. coli* 0104:H4 outbreak linked to fenugreek sprout consumption (Mellmann et al., 2011; King et al., 2012). NGS technologies allow for the sequencing of entire genomes of multiple pathogens in a single sequence run, whether from bacterial isolates of different patients or various species in patient material from a single individual (metagenomics). Although *E. coli* strains are typically non-harmful and part of the normal intestinal microflora, certain pathogenic groups like DEC and ExPEC can cause illness in humans and animals beyond the GI tract (Wang et al., 2022).

The different pathotypes of *E. coli* each have unique pathogenicity mechanisms and a specific set of virulence factors encoded by specific gene clusters. These genes may be involved in activities such as adhesion, invasion, attachment, iron acquisition, motility, and toxin activity (Sarowska et al., 2019). *E. coli* pathotypes can be categorized into four main virulence classes: colonization, fitness, toxins, and effectors, each containing specific virulence factors with distinct functions (Pakbin et al., 2021). It's worth noting that various enteric and extraintestinal pathotypes of *E. coli* share common virulence factors and strategies (Mainil, 2013). In recent years, many virulence factors associated with *E. coli* pathotypes linked to intestinal and extraintestinal disorders have been identified (Gomes et al., 2016; Clermont et al., 2019). Studying these virulence factors and their associated genes can provide valuable insights into the interactions between these factors in *E. coli* pathotypes and host proteins at the molecular level, shedding light on how they cause

diseases and enabling the development of preventive strategies (Songe et al., 2016). The extensive use of NGS holds promise in overcoming the limitations in identifying these virulence factors and associated genes.

Virulence genes have the potential to enhance the pathogenicity of *E. coli* by inhibiting or evading the host's immune system and deriving nutrients from the host, consequently depriving the host of essential nutrients (Wijetunge et al., 2015; Clermont et al., 2019). In recent years, much attention has been given to the analysis of the phylogenetic affiliation of pathogenic *E. coli* strains, to understand the sources of such ExPEC and to limit the spread of multidrug resistance among such strains. *E. coli* strains mainly fall into four phylogenetic groups (A, B1, B2, and D) and those virulent extra-intestinal strains mainly belong to groups B2 and D (Carlos et al., 2010; Alfinete et al., 2021). It has been found that pathogenic *E. coli* strains causing extraintestinal infections mainly belong to group B2 and, to a lesser extent, group D. In contrast, commensal strains belong to groups A and B1 (Chakraborty et al., 2015; Cordoni et al., 2016). In a study by Lyhs et al. (2012), 207 *E. coli* isolates from poultry meat products were characterized using the polymerase chain reaction (PCR) method. The findings revealed the presence of a virulent gene in each isolate. The majority of the isolates were categorized into phylogenetic group D, followed by groups A and B2 (Lyhs et al., 2012). Based on virulence factor gene PCR, 23.2 percent of the strains were classified as ExPEC strains, containing avian pathogenic, uropathogenic, and neonatal meningitis-causing *E. coli* (Johnson et al., 2008; Zhao et al., 2018). Furthermore, all *E. coli* strains obtained from raw meat and shellfish contained at least one of the 16 virulence genes detected (Van et al., 2008; Bradshaw, 2024). It is imperative to identify such virulence genes in the food industry using NGS as their presence in the identified *E. coli* strains raises significant concerns for human health.

Over the past two decades, molecular diagnostic methods have made significant advances and have played an increasingly critical role in medical microbiology laboratories. These methods have significantly reduced the time from sample collection to obtaining results and have enabled the detection of non-cultivable pathogens (Molechan et al., 2019). However, whole-genome sequencing (WGS) has emerged as the gold standard due to its accessibility and affordability, revolutionizing outbreak investigations (Nutman and Marchaim, 2019). In a study on genome-based characterization of *E. coli* causing bloodstream infections through NGS by Paramita et al. (2020), 22 patients were found to have *E. coli* isolates exhibiting high diversity in serotypes, sequence types, virulence genes, and antimicrobial resistance (AMR) genes. Of the 22 *E. coli* isolates, 12 different sequence types (STs) were identified. Notably, five (22.7 percent) of the *E. coli* samples belonged to ST131, all of which had serotype O25:H4. Additionally, each of ST38, ST405, and ST69 were observed in three (13.6 percent) out of the 22 *E. coli* isolates. Paramita et al. (2020) provided further explanations regarding the observed stereotypes, particularly that ST38 had two stereotypes (086:H18 and an undefined stereotype belonging to H30). Additionally, ST405 was associated with a particular O102 stereotype, while ST69 comprised numerous diverse stereotypes. It was noted that strains with the same stereotype and sequence types (STs) harbored the same set of virulence genes (Bien et al., 2012; Lemaitre et al., 2014; Paramita et al., 2020).

Recent studies have provided strong support for the significant advancements in genomics facilitated by NGS, which has greatly expanded our understanding of genome structure, function, and dynamics. NGS technology has empowered extensive research, allowing scientists to delve into the complexities of genetic information in unprecedented ways. For instance, in the context of understanding the pathogenicity and virulence factors of *E. coli*, Nafea et al. (2024), emphasized that NGS accurately identifies and detects a diverse range of pathogens, including viruses, bacteria, fungi, and parasites, across various sample types, such as clinical specimens, environmental samples, and vectors. However, the use of NGS is not without its limitations, particularly in pathogen detection, which include the need for robust bioinformatics pipelines, standardized methods, and optimized workflows for different sample types (Nafea et al., 2024). It's important to note that while NGS shows immense potential in pathogen identification, metagenomic NGS procedures are not yet approved by the Food and Drug Administration (Kim et al., 2016; Kong et al., 2024), and its application in clinical pathogen recognition is still in its early stages. Despite these challenges, ongoing advancements in sequencing technologies, data analysis tools, and collaborative efforts among researchers, clinicians, and public health agencies are expected to further enhance the application of NGS in identifying various pathogens. These developments are poised to significantly improve the diagnostics, surveillance, and control of infectious diseases.

4.1 Notable studies which paved NGS in the study of *E. coli* and pathogenicity

Before NGS technologies, researchers relied on real-time PCR, microarrays, and culture methods to uncover the pathogenicity and virulence factors of *E. coli* (Li et al., 2024). These early studies laid the foundation for subsequent research based on NGS. Some of these studies involved genetic and genomic characterization using DNA-DNA hybridization (DDH) procedures (Marmur et al., 1963). This method classified strains as members of *E. coli* if they exhibited at least 70 percent DNA similarity to the reference strains (Brenner et al., 1972). It's important to note that DDH percentages do not precisely reflect the actual DNA identity between strains (Rosselló-Mora, 2006), but this approach was pivotal in establishing a threshold-based method for defining bacterial species.

Report from World Health Organization (2022) stipulated thatalmost 600 million people suffer from illnesses caused by consuming contaminated food each year, leading to an estimated 420,000 deaths. Among the various pathogenic bacteria associated with foodborne illnesses, *E. coli* holds a unique position. This bacterium exhibits a dual nature, serving as a beneficial resident of the gut and a harmful pathogen, making it a focal point of scientific research (Kornacki and Marth, 1982; Liu et al., 2017). Most of the strains of *E. coli* possess beneficial attributes essential to bodily processes such as vitamin K productionwhich aid in the

prevention of harmful bacterial colonization (Jangid et al., 2024). Particularly, the O157 strain has garnered attention due to its involvement in serious foodborne outbreaks. Several research efforts are currently underway to thoroughly understand the epidemiology, pathogenesis, transmission dynamics, and preventive measures associated with *E. coli* O157 (Carlos et al., 2010; Schwaiger et al., 2024).

In their 2017 study, Messerer et al. utilized NGS to explore the horizontal gene transfer of pathogenicity islands in *E. coli*. The research findings indicated that pathogenic *E. coli* (ExPEC) harbors pathogenicity islands (PAIs), and the mechanisms by which these PAIs are acquired remain poorly understood. These PAIs equip ExPEC with the ability to efficiently colonize and invade hosts. The study also demonstrated that tetracycline-resistant transconjugants had a higher efficiency in spreading the acquired PAIs through conjugation (Messerer et al., 2017). This phenomenon potentially contributes to the widespread dissemination of the *E. coli* HPI despite its lack of self-transferability. Notably, the Integrative and Conjugative Elements (ICEs) type of the HPI, while still mobile, is less prevalent within the *E. coli* species (Martin et al., 2013; Jangid et al., 2024).

4.2 *E. coli* pathogenesis-related genes discovered using RNA-seq transcriptomics

NGS is already in use in numerous medical microbiology laboratories for activities such as outbreak management, molecular case finding, pathogen characterization, and surveillance. With the ability to rapidly identify bacteria using the 16S-23S rRNA region, perform taxonomy and metagenomics approaches on clinical samples, and determine zoonotic microorganism transmission from animals to humans, NGS is proving to be highly beneficial (Deurenberg et al., 2017; Cui et al., 2024).

Another cutting-edge method in the field is RNA sequencing (RNA-Seq), which offers an unbiased high-throughput sequencing approach to capture the global transcriptional response of an organism under specific conditions. This approach enables the simultaneous analysis of all regions of the genome, which sets it apart from methods like microarray and quantitative reverse-transcription PCR (qRT-PCR), which are limited to known genomic regions as targets. Moreover, RNA-Seq can be used to analyze isolates with diverse or unknown genomic content, unlike microarray analysis, which relies on samples exhibiting sequence similarity to known targets used to develop the microarray probes (Lyhs et al., 2012; Deurenberg et al., 2017; Clermont et al., 2019). This passage provides a comprehensive overview of the application of RNA-Seq in characterizing the transcriptomes of various human disease-associated bacteria (Servin, 2014; Zhao et al., 2022). The studies cited showcase the use of RNA-Seq in elucidating differences in quorum sensing regulons of distinct isolates of *Pseudomonas aeruginosa*, as well as in identifying global regulators in enterotoxigenic *E. coli* (ETEC). However, the text also highlights the absence of prior studies utilizing RNA-Seq to delineate the global transcriptional response of enteropathogenic *E. coli* (EPEC) and the variability in the global virulence regulons of EPEC across divergent phylogenomic lineages of *E. coli*.

4.3 *E. coli* pathogenetic proteins discovered using proteomics approaches

Proteins serve as macromolecular machines that carry out a wide range of biochemical functions, such as acting as building blocks, transporters, and enzymes. These protein functions are intricately linked with other components of organisms, including genes, RNA, and metabolites. Proteomics involves the comprehensive study of the sets of proteins produced by organisms (Geiger et al., 2010; Aebersold and Mann, 2023). The entire set of proteins produced by an organism is known as its proteome, which can vary across cells and is influenced to some degree by the underlying transcriptome. In the past, proteins were primarily studied using low-throughput methods that focused on a relatively small set of proteins, providing qualitative data on their structure, function, and interactions with other cellular components. However, these traditional techniques offered only limited insight into the overall proteome of a cell (Cho, 2007; Tenaillon et al., 2016). In contrast to gel-based and antibody-based methods, mass spectrometry (MS) has been employed to generate extensive datasets on the proteome (Chandramouli and Qian, 2009). The typical workflow in proteomics begins with the extraction of total proteins from the tissue, followed by trypsin digestion, chromatographic separation of short peptides, and MS analysis. This is followed by the identification of proteins in the sample and the compilation of a protein list. Similar to NGS, proteomics studies have been advanced by the development of various instruments that enable peptide separation, mass analysis, and other downstream applications. These instruments must meet specific criteria, including high throughput and confidence in peptide identification, with notable examples being Orbitrap and time-of-flight mass analyzers (Geiger et al., 2010). A common approach to improving MS performance has been the development of hybrid systems, which combine different ion analyzers in a triple quadrupole instrument, enhancing MS capacity over a single quadrupole. In the triple quadrupole, data on m/z values are combined with data on molecule fragmentation patterns to improve accuracy, made possible by the inclusion of a second quadrupole acting as a collision cell (Aebersold and Mann, 2023; Geiger et al., 2010). Adhesins of pathogenic *E. coli* can also include outer-membrane proteins, such as intimin of uropathogenic *E. coli* (UPEC) and enterohemorrhagic *E. coli* (EHEC), or other non-fimbrial proteins. Certain surface structures can trigger signal transduction pathways or cytoskeletal rearrangements leading to disease. The identification of such proteins has been facilitated by advanced methods of protein identification.

In a study conducted by Bose et al. (2017), an in silico investigation was reported on the protein-protein interactions (PPIs) between human cells and four EHEC strains, namely EDL933, Sakai, EC4115, and TW14359. The purpose of the study was to gain insights into the virulence and host-colonization

strategies of these strains. The research encompassed the intra-species PPI data from humans and different strains of *E. coli*, aiming to infer potential Host Pathogenic Interactions (HPIs) between the host and the pathogen proteins. The findings revealed that the invading pathogens have developed mechanisms to evade the host's immune defenses, particularly by counteracting the redox stresses induced by the host. Given the significance of PPIs in bacterial colonization and survival within the host, it was conjectured that the abundance of host and bacterial proteins involved in these interactions would reflect the bacteria's colonization potential (Bose et al., 2017). Higher HCI- values of the 0157:H7 strains indicate their superior ability to colonize the human gut. The identified HPI is anticipated to contribute to understanding the biological underpinnings of diseases caused by various bacterial pathogens, including E. coli (Liang et al., 2005; Bose et al., 2017).

5 Summary and conclusions

A wide variety of pathogenicity mechanisms and virulence factor profiles are present across enteric *E. coli* pathotypes, raising significant concerns for public health and food safety (Han et al., 2020). The evolution of enteric *E. coli* pathotypes has led to the emergence of distinct pathotypes capable of toxin secretion, aggregative colonization, and survival in the gastrointestinal tract. These adaptations, driven by key genetic elements, have given rise to new pathotypes (Beutin and Martin, 2012; Lang et al., 2018). The emergence of novel *E. coli* pathotypes, like Shiga toxin-producing *E. coli* (STEC), has highlighted the need for robust surveillance systems, as evidenced by the contamination of raw and pasteurized bulk milk in South Africa resulting in fatalities (Ntuli, 2017). To address this, studies and monitoring systems have been integrated into one-health approaches and networks, recognizing that enteric pathogenic *E. coli* can be transmitted through food, water, animals, and humans (Yang et al., 2017; Kaczvinsky et al., 2024). It is important to note that enteric *E. coli* pathotypes have a significant impact on various functions within host cells, such as protein synthesis, gene transcription, secretion of molecules and ions, cytoskeleton rearrangement, apoptosis, and signal transduction. These pathotypes possess a wide array of virulence factors which are encoded by specific gene clusters on the chromosome or mobile genetic elements (Ntuli et al., 2018). To effectively monitor these pathotypes, it is essential to utilize NGS and proteomics to obtain accurate information about their virulence profiles. While numerous pathogenicity and virulence factors have been identified, there is still a need for further research leveraging NGS and proteomics technologies. The discovery of additional pathogenicity and virulence factors could lead to the development of new control methods, novel treatments, and improved immunization strategies against *E. coli*.

Author contributions

KM: Writing – review & editing, Writing – original draft. LN: Visualization, Software, Formal analysis, Writing – original draft, Resources, Funding acquisition, Project administration, Methodology, Data curation, Investigation, Supervision, Validation, Writing – review & editing, Conceptualization. CK: Writing – review & editing, Conceptualization, Funding acquisition, Validation, Software, Methodology, Supervision, Resources, Writing – original draft, Project administration, Formal analysis, Investigation, Data curation, Visualization. MT: Supervision, Investigation, Software, Methodology, Writing – review & editing, Conceptualization, Funding acquisition, Writing – original draft, Formal analysis, Visualization, Project administration, Resources, Data curation, Validation.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. The authors wish to acknowledge the Water Research Commission (South Africa) for funding under the research grant 2022/2023-00801.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbrio.2025.1677775/full#supplementary-material

References

Aebersold, R., and Mann, M. (2023). Mass spectrometry-based proteomics. *Nature* 422, 198–207. doi: 10.1038/nature01511

Alfinete, N. W., Bolukaoto, J. Y., Heine, L., Potgieter, N., and Barnard, T. G. (2021). Virulence and phylogenetic analysis of enteric pathogenic Escherichia coli isolated from children with diarrhoea in South Africa. *Int. J. Infect. Dis.* 114, 226–232. doi: 10.1016/j.ijid.2021.11.017

Barzilay, E. J., Schaad, N., Magloire, R., Mung, K. S., Boncy, J., Dahourou, G. A., et al. (2013). Cholera surveillance during the Haiti epidemic–the first 2 years. *N Engl. J. Med.* 368, 599–609. doi: 10.1056/NEJMoa1204927

Bien, J., Sokolova, O., and Bozko, P. (2012) Role of Uropathogenic Escherichia coli Virulence Factors in Development of Urinary Tract Infection and Kidney Damage. *Int J Nephrol.* 681473. doi: 10.1155/2012/681473

Berthe, T., Ratajczak, M., Clermont, O., Denamur, E., and Petit, F. (2013). Evidence for the coexistence of distinct *Escherichia coli* populations in various aquatic environments and their survival in estuary water. *Appl. Environ. Microbiol.* 79, 4684–4693. doi: 10.1128/AEM.00698-13

Beutin, L., and Martin, A. (2012). Outbreak of Shiga toxin-producing *Escherichia coli* (STEC) O104:H4 infection in Germany causes a paradigm shift with regard to human pathogenicity of STEC strains. *J. Food Prot* 75, 408–418. doi: 10.4315/0362-028X.JFP-11-452

Bielaszewska, M., Mellmann, A., Zhang, W., Köck, R., Fruth, A., Bauwens, A., et al. (2011). Characterisation of the Escherichia coli strain associated with an outbreak of haemolytic uraemic syndrome in Germany 2011: a microbiological study. *Lancet Infect. Dis.* 11, 671–676. doi: 10.1016/S1473-3099(11)70165-7

Bose, T., Venkatesh, K. V., and Mande, S. S. (2017). Computational analysis of host-pathogen protein interactions between humans and different strains of enterohemorrhagic escherichia coli. *Front. Cell Infect. Microbiol.* 7. doi: 10.3389/fcimb.2017.00128

Bradshaw, C. (2024). Phosphorus distribution with respect to particle size and microbial activity in dairy manure and implications for manure management. Canada: Doctoral dissertation, University of Northern British Columbia. doi: 10.24124/2024/59478

Brenner, D. J., Fanning, G. R., Skerman, F. J., and Falkow, S. (1972) Polynucleotide sequence divergence among strains of Escherichia coli and closely related organisms. *J Bacteriol.* 109(3), 953–65. doi: 10.1128/jb.109.3.953-965.1972

Carlos, C., Pires, M. M., Stoppe, N. C., Hachich, E. M., Sato, M. I., Gomes, T. A., et al. (2010). *Escherichia coli* phylogenetic group determination and its application in the identification of the major animal source of fecal contamination. *BMC Microbiol.* 10, 161. doi: 10.1186/1471-2180-10-161

CDC (2016). Escherichia coli *(*E. coli*)* (Atlanta, GA: US: Department of Health and Huma Services, CDC).

Chakraborty, A., Saralaya, V., Adhikari, P., Shenoy, S., Baliga, S., and Hegde, A. (2015). Characterization of escherichia coli phylogenetic groups associated with extraintestinal infections in South Indian population. *Ann. Med. Health Sci. Res.* 5, 241–246. doi: 10.4103/2141-9248.160192

Chandramouli, K., and Qian, P. Y. (2009). Proteomics: challenges, techniques, and possibilities to overcome biological sample complexity. *Hum. Gen. Proteom* 4, 23920. doi: 10.4061/2009/239204

Cho, W. C. S. (2007). Proteomics technologies and challenges. *Geno Prot Bioinfo* 5 (2), 77–85. Available online at: https://creativecommons.org/licenses/by/4.0/.

Clark, J. R., and Maresso, A. M. (2021). Comparative pathogenomics of *Escherichia coli*: polyvalent vaccine target identification through virulome analysis. *Infect. Immun.* 89, e0011521. doi: 10.11128/IAI.00115-21

Clements, A., Young, J. C., Constantinou, N., and Frankel, G. (2012) Infection strategies of enteric pathogenic Escherichia coli. *Gut Microbes*. 3(2), 71–87. doi: 10.4161/gmic.19182

Clermont, O., Dixit, O. V. A., Vangchhia, B., Condamine, B., Dion, S., Bridier-Nahmias, A., et al. (2019). Characterization and rapid identification of phylogroup G in *Escherichia coli*, a lineage with high virulence and antibiotic resistance potential. *Environ. Microbio* 21, 3107–3117. doi: 10.1111/1462-2920.14713

Cordoni, G., Woodward, M. J., Wu, H., Alanazi, M., Wallis, T., and La Ragione, R. M. (2016). Comparative genomics of European avian pathogenic E. *coli (APEC). BMC Genomics* 17, 960. doi: 10.1186/s12864-016-3289-7

Cui, L., Zheng, J., Lin, Y., Lin, P., Lu, Y., Zheng, Y., et al. (2024). Decoding the ribosome's hidden language: rRNA modifications as key players in cancer dynamics and targeted therapies. *Clin. Transl. Med.* 14, e1705. doi: 10.1002/ctm2.1705

Dadonaite, B., Ritchie, H., and Roser, M. (2018). *Diarrheal Diseases* (Our World Data). Available online at: www.ourworldindata.org/diarrheal-diseases (Accessed June 18, 2024).

Daga, A. P., Koga, V. L., Soncini, J. G. M., de Matos, C. M., Perugini, M. R. E., Pelisson, M., et al. (2019). *Escherichia coli* Bloodstream Infections in Patients at a University Hospital: Virulence factors and clinical characteristics. *Front. Cell Infect. Microbiol.* 9. doi: 10.3389/fcimb.2019.00191

Desvaux, M., Dalmasso, G., Beyrouthy, R., Barnich, N., Delmas, J., Bonnet, R., et al. (2020). Pathogenicity factors of genomic islands in intestinal and extraintestinal escherichia coli. *Front. Microbiol.* 11. doi: 10.3389/fmicb.2020.02065

Deurenberg, R. H., Bathoorn, E., Chlebowicz, M. A., Couto, N., Ferdous, M., García-Cobos, S., et al. (2017). Application of next-generation sequencing in clinical microbiology and infection prevention. *J. Biotechnol.* 243, 16–24. doi: 10.1016/j.jbiotec.2016.12.022

Donnenberg, M. (2013). Escherichia coli: *Pathotypes and Principles of Pathogenesis* (New York, NY, USA: Academic Press). Available online at: http://lib.ugent.be/catalog/ebk01:2550000001064180 (Acessed May 20, 2024).

Dutta, D., Dobson, A. J., Houtz, P. L., Gläßer, C., Revah, J., Korzelius, J., et al. (2015). Regional cell-specific transcriptome mapping reveals regulatory complexity in the adult drosophila midgut. *Cell Rep.* 12, 346–358. doi: 10.1016/j.celrep.2015.06.009

Eichhorn, I., Heidemanns, K., Semmler, T., Kinnemann, B., Mellmann, A., Harmsen, D., et al. (2015). Highly virulent Non-O157 enterohemorrhagic Escherichia coli (EHEC) serotypes reflect similar phylogenetic lineages, providing new insights into the evolution of EHEC. *Appl. Environ. Microbiol.* 81, 7041–7047. doi: 10.1128/AEM.01921-15

Escherich, T. (1885). Die Darmbakterien des Neugeborenen und Säuglings. *Fortschr Med.* 3, 515–522.

Foster, M. A., Iqbal, J., Zhang, C., McHenry, R., Cleveland, B. E., Romero-Herazo, Y., et al. (2015). Enteropathogenic and enteroaggregative E. coli in stools of children with acute gastroenteritis in Davidson County, Tennessee. *Diagn. Microbiol. Infect. disease.* 83, 319–324. doi: 10.1016/j.diagmicrobio.2015.07.016

Franz, E., Delaquis, P., Morabito, S., Beutin, L., Gobius, K., Rasko, D. A., et al. (2014). Exploiting the explosion of information associated with whole genome sequencing to tackle Shiga toxin-producing *Escherichia coli* (STEC) in global food production systems. *Int. J. F Microbiol.* 187, 57–72. doi: 10.1016/j.ijfoodmicro.2014.07.002

Franzin, F. M., and Sircili, M. P. (2015). Locus of enterocyte effacement: a pathogenicity island involved in the virulence of enteropathogenic and enterohemorrhagic Escherichia coli subjected to a complex network of gene regulation. *BioMed. Res. Int.* 2015, 534738. doi: 10.1155/2015/534738

Garcia, T. A., Ventura, C. L., Smith, M. A., Merrell, D. S., and O'Brien, A. D. (2013). Cytotoxic necrotizing factor 1 and hemolysin from uropathogenic *Escherichia coli* elicit different host responses in the murine bladder. *Inf Immune* 81, 99–109. doi: 10.1128/IAI.00605-12

Geiger, T., Cox, J., and Mann, M. (2010). Proteomics on an Orbitrap benchtop mass spectrometer using all-ion fragmentation. *Mol. Cell Proteom* 9, 2252–2261. doi: 10.1074/mcp.M110.001537

Geurtsen, J., de Been, M., Weerdenburg, E., Zomer, A., McNally, A., Poolman, J., et al. (2022). Genomics and pathotypes of the many faces of Escherichia coli. *FEMS Microbiol. Rev.* 46, fuac031. doi: 10.1093/femsre/fuac031

Gomes, I., Ayoub, M. A., Fujita, W., Jaeger, W. C., Pfleger, K. D., Devi, L. A., et al. (2016). G protein-coupled receptor heteromers. *Annu. Rev. Pharmacol. Toxicol.* 56, 403–425. doi: 10.1146/annurev-pharmtox-011613-135952

Habibi, M., Asadi Karam, M. R., and Bouzari, S. (2017). Evaluation of prevalence, immunogenicity, and efficacy of FyuA iron receptor in uropathogenic Escherichia coli isolates as a vaccine target against urinary tract infection. *Microb. Pathog.* 110, 477–483. doi: 10.1016/j.micpath.2017.07.037

Han, Z., An, W., Yang, M., and Zhang, Y. (2020). Assessing the impact of source water on tap water bacterial communities in 46 drinking water supply systems in China. *Water Res.* 172, 115469. doi: 10.1016/j.watres.2020.115469

Hazen, T. H., Michalski, J., Luo, Q., Shetty, A. C., Daugherty, S. C., Fleckenstein, J. M., et al. (2017). Comparative genomics and transcriptomics of *Escherichia coli* isolates carrying virulence factors of both enteropathogenic and enterotoxigenic E. coli. *Sci. Rep.* 7, 3513. doi: 10.1038/s41598-017-03489-z

Herzog, K., Debertolis, L., Kastelic, J. P., Schmicke, M., Ulbrich, S. E., Bollwein, H., et al. (2019) Effects of intravenous infusion of E. coli lipopolysaccharide in early pregnant cows. *Reproduction*. 157(1), 65–76. doi: 10.1530/REP-18-0174

Hofman, P., Le Negrate, G., Mograbi, B., Hofman, V., Brest, P., Alliana-Schmid, A., et al. (2000). *Escherichia coli* cytotoxic necrotizing factor-1 (CNF-1) increases the adherence to epithelia and the oxidative burst of human polymorphonuclear leukocytes but decreases bacteria phagocytosis. *J. Leukoc. Biol.* 68, 522–528. doi: 10.1189/jlb.68.4.522

Iqbal, J., Dufendach, K. R., Wellons, J. C., Kuba, M. G., Nickols, H. H., Gómez-Duarte, O. G., et al. (2016). Lethal neonatal meningoencephalitis is caused by multidrug-resistant, highly virulent Escherichia coli. *Infect. Dis. Rep.* 48, 461–466. doi: 10.3109/23744235.2016.1144142

Jang, J., Hur, H. G., Sadowsky, M. J., Byappanahalli, M. N., Yan, T., Ishii, S., et al. (2017) Environmental Escherichia coli: ecology and public health implications-a review. *J Appl Microbiol.* 123(3), 570–581. doi: 10.1111/jam.13468

Jangid, H., Kumar, D., Kumar, G., Kumar, R., and Mamidi, N. (2024) An Emerging Foodborne Pathogen Spotlight: A Bibliometric Analysis and Scholarly Review of Escherichia coli O157 Research. *Antibiotics (Basel)*. 13(1), 60. doi: 10.3390/antibiotics13010060

Johnson, T. J., Wannemuehler, Y., Johnson, S. J., Stell, A. L., Doetkott, C., Johnson, J. R., et al. (2008). Comparison of extraintestinal pathogenic Escherichia coli strains from

human and avian sources reveals a mixed subset representing potential zoonotic pathogens. *Appl. Environ. Microbiol.* 74, 7043–7050. doi: 10.1128/AEM.01395-08

Kaczvinsky, C., Levy, H., Preston, S., Youngflesh, C., Clucas, G., Lynch, H. J., et al. (2024). The influence of biotic and abiotic factors on the bacterial microbiome of gentoo penguins (Pygoscelis papua) in their natural environment. *Sci. Rep.* 14, 17933. doi: 10.1038/s41598-024-66460-9

Kaper, J., Nataro, J., and Mobley, H. (2004). Pathogenic escherichia coli. *Nat. Rev. Microbiol.* 2, 123–140. doi: 10.1038/nrmicro818

Kim, E. J., Chang, H. J., Kwak, S., and Park, J. H. (2016) Virulence Factors and Stability of Coliphages Specific to Escherichia coli O157:H7 and to Various E. coli Infection. *J Microbiol Biotechnol*. 26(12), 2060–2065. doi: 10.4014/jmb.1609.09039

King, L. A., Nogareda, F., Weill, F. X., Mariani-Kurkdjian, P., Loukiadis, E., Gault, G., et al. (2012). Outbreak of Shiga toxin-producing Escherichia coli O104:H4 associated with organic fenugreek sprouts, France, June 2011. *Clin. Infect. Dis.* 54, 1588–1594. doi: 10.1093/cid/cis255

Kong, H., Hu, Z., Zhang, L., Chen, Q., Yang, L., Li, J., et al. (2024) Clinical risk factors and outcomes of carbapenem-resistant Escherichia coli nosocomial infections in a Chinese teaching hospital: a retrospective study from 2013 to 2020. *Microbiology Spectrum*. 12(7), e04228–23. doi: 10.1016/j.lwt.2022.113913

Kornacki, J. L., and Marth, E. H. (1982). Foodborne illness caused by escherichia coli. *J. Food Prot* 45, 1051–1067. doi: 10.4315/0362-028X-45.11.1051

Lang, C., Fruth, A., Holland, G., Laue, M., Mühlen, S., Dersch, P., et al. (2018). Novel type of pilus associated with a Shiga-toxigenic E. coli hybrid pathovar conveys aggregative adherence and bacterial virulence. *Emerg. Microbes Infect.* 7, 1–16. doi: 10.1038/s41426-018-0209-8

LeJeune, J., Homan, J., Linz, G., and Pearl, D. L. (2008). Role of the european starling in the transmission of E. coli 0157 on dairy farms. *Univ. California Agric. Natural Resources Proc. Vertebr Pest Conf* 23, 31–34. doi: 10.5070/V423110392

Lemaître, C., Bidet, P., Benoist, J. F., Schlemmer, D., Sobral, E., d'Humières, C., et al. (2014). The ssbL gene harboured by the ColV plasmid of an *Escherichia coli* neonatal meningitis strain is an auxiliary virulence factor boosting the production of siderophores through the shikimate pathway. *J. Bacteriol* 196, 1343–1349. doi: 10.1128/JB.01153-13

Lerm, M., Selzer, J., Hoffmeyer, A., Rapp, U. R., Aktories, K., Schmidt, G., et al. (1999). Deamidation of Cdc42 and Rac by Escherichia coli cytotoxic necrotizing factor 1: activation of c-Jun N-terminal kinase in HeLa cells. *Infect. Immun.* 67, 496–503. doi: 10.1128/IAI.67.2.496-503.1999

Lewis, S. B., Cook, V., and Tighe, R. (2015). Enterohemorrhagic *Escherichia coli* colonization of human colonic epithelium *in vitro* and ex vivo. *Infect. Immun.* 83, 942–949. doi: 10.1128/IAI.02928-14

Li, J., Ma, L., Liao, X., Liu, D., Lu, X., Chen, S., et al. (2018) Ultrasound-Induced Escherichia coli O157:H7 Cell Death Exhibits Physical Disruption and Biochemical Apoptosis. *Front Microbiol*. 9:2486. doi: 10.3389/fmicb.2018.02486

Li, C., Ouyang, Z., and Liu, J. (2024). Bacterial growth and cultivation. *Mol. Med. Microbiol.*, 155–175. doi: 10.1016/B978-0-12-818619-0.00070.8

Liang, B., Goodman, L., Tummala-Narra, P., and Weintraub, S. (2005). A theoretical framework for understanding help-seeking processes among survivors of intimate partner violence. *Am. J. Community Psychol.* 36, 71–84. doi: 10.1007/s10464-005-6233-6

Lindstedt, B. A., Finton, M. D., Porcellato, D., and Brandal, L. T. (2018). High frequency of hybrid *Escherichia coli*strains with combined Intestinal Pathogenic *Escherichia coli* (IPEC) and Extraintestinal Pathogenic *Escherichia coli* (ExPEC) virulence factors isolated from human faecal samples. *BMC Infect. Dis.* 18, 544. doi: 10.1186/s12879-018-3449-2

Liu, G., Zhang, Y., Knibbe, W. J., Feng, C., Liu, W., Medema, G., et al. (2017). Potential impacts of changing supply-water quality on drinking water distribution: a review. *Water Res.* 116:135–48. doi: 10.1016/j.watres.2017.03.031

Lo, Y., Zhang, L., and Foxman, B. (2015). Whole-genome sequencing of uropathogenic *Escherichia coli* reveals long evolutionary history of diversity and virulence. *Infect. Genet. Evol.* 34, 244–250. doi: 10.1016/j.meegid.2015.06.023

Lyhs, U., Ikonen, I., Pohjanvirta, T., Raninen, K., Perko-Mäkelä, P., Pelkonen, S., et al. (2012). Extraintestinal pathogenic Escherichia coli in poultry meat products on the Finnish retail market. *Acta Vet. Scand.* 54, 64. doi: 10.1186/1751-0147-54-64

Mainil, J. (2013). Escherichia coli virulence factors. *Vet. Immunol. Immunopathol.* 152, 2–12. doi: 10.1016/j.vetimm.2012.09.032

Mandomando, I., Vubil, D., Boisen, N., Quintó, L., Ruiz, J., Sigaúque, B., et al. (2020) Escherichia coli ST131 clones harbouring AggR and AAF/V fimbriae causing bacteremia in Mozambican children: Emergence of new variant of fimH27 subclone. *PLoS Negl Trop Dis*. 14(5), e0008274. doi: 10.1371/journal.pntd.0008274

Martin, P., Marcq, I., Magistro, G., Penary, M., Garcie, C., Payros, D., et al. (2013). Interplay between siderophores and colibactin genotoxin biosynthetic pathways in Escherichia coli. *PLoS Pathog.* 9, e1003437. doi: 10.1371/journal.ppat.1003437

Marmur, J., Falkow, S., and Mandel, M. (1963) New approaches to bacterial taxonomy. *Annu Rev Microbiol.* 17, 329–72. doi: 10.1146/annurev.mi.17.100163.001553

Mellmann, A., Harmsen, D., Cummings, C. A., Zentz, E. B., Leopold, S. R., Rico, A., et al. (2011) Prospective genomic characterization of the German enterohemorrhagic Escherichia coli O104:H4 outbreak by rapid next generation sequencing technology. *PLoS One*. 6(7), e22751. doi: 10.1371/journal.pone.0022751

Martins, F. H., Guth, B. E., Piazza, R. M., Leão, S. C., Ludovico, A., Ludovico, M. S., et al. (2015) Diversity of Shiga toxin-producing Escherichia coli in sheep flocks of Paraná State, southern Brazil. *Vet Microbiol*. 175(1), 150–6. doi: 10.1016/j.vetmic.2014.11.003

Messerer, M., Fischer, W., and Schubert, S. (2017). Investigation of horizontal gene transfer of pathogenicity islands in *Escherichia coli* using next-generation sequencing. *PLoS One* 12, e0179880. doi: 10.1371/journal.pone.0179880

Meza-Segura, M., Zaidi, M. B., Vera-Ponce de León, A., Moran-Garcia, N., Martinez-Romero, E., Nataro, J. P., et al. (2020). New insights into DAEC and EAEC pathogenesis and phylogeny. *Front. Cell. Infection Microbiol.* 10, 572951. doi: 10.3389/fcimb.2020.572951

Molechan, C., Amoako, D. G., Abia, A. L. K., Somboro, A. M., Bester, L. A., Essack, S. Y., et al. (2019). Molecular epidemiology of antibiotic-resistant Enterococcus spp. from the farm-to-fork continuum in intensive poultry production in KwaZulu-Natal, South Africa. *Sci. Total Environ.* 692, 868–878. doi: 10.1016/j.scitotenv.2019.07.324

Mueller, M., and Tainter, C. R. (2023). "*Escherichia coli* Infection," in *StatPearls* (StatPearls Publishing, Treasure Island (FL). Available online at: https://www.ncbi.nlm.nih.gov/books/NBK564298/.

Nafea, A. M., Wang, Y., Wang, D., Salama, A. M., Aziz, M. A., Xu, S., et al. (2024) Application of next-generation sequencing to identify different pathogens. *Front Microbiol*. 14, 1329330. doi: 10.3389/fmicb.2023.1329330

Ntuli, V. (2017) Shigatoxin producing Escherichia coli O157 and non-O157 serotypes in producer-distributor bulk milk. University of Pretoria (South Africa).

Ntuli, V., Njage, P. M. K., Bonilauri, P., Serraino, A., and Buys, E. M. (2018). Qualitative risk assessment of Hemolytic Uremic Syndrome associated with consumption of bulk milk sold directly from producer to consumer in South Africa. *J. Food Prot* 81, 472–481. doi: 10.4315/0362-028X.JFP-17-199

Nutman, A., and Marchaim, D. (2019). How to: Molecular investigation of a hospital outbreak. *Clin. Microbiol. Infect.* 25, 688–695. doi: 10.1016/j.cmi.2018.09.017

Pakbin, B., Bruck, W. M., and Rossen, J. W. A. (2021). Virulence factors of enteric pathogenic *escherichia coli*: A review. *Int. J. Mol. Sci.* 22, 9922. doi: 10.3390/ijms22189922

Palmela, C., Chevarin, C., Xu, Z., Torres, J., Sevrin, G., Hirten, R., et al. (2018). Adherent-invasive Escherichia coli in inflammatory bowel disease. *Gut* 67, 574–587. doi: 10.1136/gutjnl-2017-314903

Paramita, R. I., Nelwan, E. J., Fadilah, F., Renesteen, E., Puspandari, N., Erlina, L., et al. (2020). Genome-based characterization of Escherichia coli causing bloodstream infection through next-generation sequencing. *PLoS One* 15(12), e0244358. doi: 10.1371/journal.pone.0244358

Peng, Z., Wang, X., Huang, J., and Li, B. (2024). Pathogenic escherichia coli. *Mol. Med. Microbiol.*, 1065–1096. doi: 10.1016/B978-0-12-818619-0.00069-1

Potgieter, S. C., and Pinto, A. J. (2019). Reproducible microbial community dynamics of two drinking water systems treating similar source waters. *Microbiol* 678920. doi: 10.11101/678920

Potgieter, S., Pinto, A., Sigudu, M., du Preez, H., Ncube, E., Venter, S., et al. (2018). Long-term spatial and temporal microbial community dynamics in a large-scale drinking water distribution system with multiple disinfectant regimes. *Water Res.* 139, 406–419. doi: 10.1016/j.watres.2018.03.077

Rojas-Lopez, M., Monteiro, R., Pizza, M., Desvaux, M., and Rosini, R. (2018). Intestinal pathogenic Escherichia coli: insights for vaccine development. *Front. Microbiol.* 9. doi: 10.3389/fmicb.2018.00440

Rojas López, A., Monzón, P., and Acerenza, L. (2021) A model for the regulation of apoptosis intrinsic pathway: The potential role of the transcriptional regulator E2F in the point of no return. *J Theor Biol*. 525:110765. doi: 10.1016/j.jtbi.2021.110765

Roser, M., and Ritchie, H. (2021). *Burden of Disease* (Our World Data). Available online at: https://ouroworldindata.org/burden-of-disease. (Accessed May 31, 2024).

Rosselló-Mora, R. (2006). "DNA-DNA reassociation methods applied to microbial taxonomy and their critical evaluation," in *Molecular Identification, Systematics, and Population Structure of Prokaryotes*. Ed. E. Stackebrandt (Springer, Berlin), 23–50. doi: 10.1007/978-3-540-31292-5_2

Sarowska, J., Futoma-Koloch, B., Jama-Kmiecik, A., Frej-Madrzak, M., Ksiazczyk, M., Bugla-Ploskonska, G., et al. (2019). Virulence factors, prevalence and potential transmission of extraintestinal pathogenic Escherichia coli isolated from different sources: recent reports. *Gut pathogens*. 11, 10. doi: 10.1186/s13099-019-0290-0

Schwaiger, G., Matt, M., Streich, P., Bromann, S., Clauß, M., Elsner, M., et al. (2024). Standard addition method for rapid, cultivation-independent quantification of Legionella pneumophila cells by qPCR in biotrickling filters. *Analyst* 149, 2978–2987. doi: 10.1039/d3an02207b

Servin, A. L. (2014). Pathogenesis of human diffusely adhering Escherichia coli expressing Afa/Dr adhesins (Afa/Dr DAEC): current insights and future challenges. *Clin. Microbiol. Rev.* 27, 823–869. doi: 10.1128/CMR.00036-14

Songe, M. M., Hang'ombe, B. M., Knight-Jones, T. J., and Grace, D. (2016) Antimicrobial Resistant Enteropathogenic Escherichia coli and Salmonella spp. in Houseflies Infesting Fish in Food Markets in Zambia. *Int J Environ Res Public Health*. 14 (1), 21. doi: 10.3390/ijerph14010021

Tenaillon, O., Barrick, J. E., Ribeck, N., Deatherage, D. E., Blanchard, J. L., Dasgupta, A., et al. (2016). Tempo mode of genome evolution in a 50,000-generation experiment. *Nature* 536, 165–170. doi: 10.1038/nature18959

Tenaillon, O., Skurnik, D., Picard, B., and Denamur, E. (2010). The population genetics of commensal Escherichia coli. *Nat. Rev. Microbiol.* 8, 207–217. doi: 10.1038/nrmicro2298

Tenaillon, M. I., Burban, E., Huynh, S., Wojcik, A., Thuillet, A. C., Manicacci, D., et al. (2023) Crop domestication as a step toward reproductive isolation. *Am J Bot*. 110 (7), e16173. doi: 10.1002/ajb2.16173

Tourret, J., and Denamur, E. (2016) Population Phylogenomics of Extraintestinal Pathogenic Escherichia coli. *Microbiol Spectr*. 4(1). doi: 10.1128/microbiolspec.UTI-0010-2012

Uhlén, P., Laestadius, A., Jahnukainen, T., Söderblom, T., Bäckhed, F., Celsi, G., et al. (2000). Alpha-haemolysin of uropathogenic E. coli induces Ca2+ oscillations in renal epithelial cells. *Nature* 405, 694–697. doi: 10.1038/35015091

Van, T. T., Chin, J., Chapman, T., Tran, L. T., and Coloe, P. J. (2008). Safety of raw meat and shellfish in Vietnam: An analysis of Escherichia coli isolations for antibiotic resistance and virulence genes. *Int. J. Food Microbiol.* 124, 217–223. doi: 10.1016/j.ijfoodmicro.2008.03.029

Wang, Z., Zheng, X., Guo, G., Hu, Z., Miao, J., Dong, Y., et al. (2022). O145 may be emerging as a predominant serogroup of Avian pathogenic Escherichia coli (APEC) in China. *Veterinary Microbiol.* 266, 109358. doi: 10.1016/j.vetmic.2022.109358

WHO [World Health Organization]. (2022). Guidelines for drinking-water quality: Fourth edition incorporating the first and second addenda. Environment, Climate Change and Health (ECH), Water, Sanitation, Hygiene and Health (WSH). Available: https://www.who.int/publications/i/item/9789240045064 (Accessed May 30, 2025).

Wijetunge, D. S., Gongati, S., DebRoy, C., Kim, K. S., Couraud, P. O., Romero, I. A., et al. (2015). Characterizing the pathotype of neonatal meningitis causing Escherichia coli (NMEC). *BMC Microbiol.* 15, 211. doi: 10.1186/s12866-015-0547-9

Yang, S. C., Lin, C. H., Aljuffali, I. A., and Fang, J. Y. (2017). Current pathogenic Escherichia coli foodborne outbreak cases and therapy development. *Arch. Microbiol.* 199, 811–825. doi: 10.1007/s00203-017-1393-y

Zhao, W. D., Liu, D. X., and Chen, Y. H. (2018). *Escherichia coli* hijack Caspr1 receptor to invade cerebral vascular and neuronal hosts. *Cell Microbiol.* 5, 418–420. doi: 10.15698/mic2018.09.647

Zhao, H. X., Zhang, T. Y., Wang, H., Hu, C. Y., Tang, Y. L., Xu, B., et al. (2022). Occurrence of fungal spores in drinking water: A review of pathogenicity, odor, chlorine resistance, and control strategies. *Sci. Total Environ.* 853, 158626. doi: 10.1016/j.scitotenv.2022.158626