Non-lactose fermenting 
*Escherichia coli*: Following in the footsteps of lactose fermenting *E. coli* high-risk clones

Razib Mazumder,† Arif Hussain,† Jody E. Phelan, Susana Campino, S. M. Arefeen Haider, Araf Mahmud, Dilruba Ahmed, Md Asadulghani, Taane G. Clark and Dinesh Mondal

Multi-resistant pathogenic strains of non-lactose fermenting *Escherichia coli* (NLF *E. coli*) are responsible for various intestinal and extraintestinal infections. Although several studies have characterised such strains using conventional methods, they have not been comprehensively studied at the genomic level. To address this gap, we used whole-genome sequencing (WGS) coupled with detailed microbiological and biochemical testing to investigate 17 NLF *E. coli* from a diagnostic centre (icddr,b) in Dhaka, Bangladesh. The prevalence of NLF *E. coli* was 10%, of which 47% (8/17) exhibited multi-drug resistant (MDR) phenotypes. All isolates (17/17) were confirmed as *E. coli* and could not ferment lactose sugar. WGS data analysis revealed international high-risk clonal lineages. The most prevalent sequence types (STs) were ST131 (23%), ST1193 (18%), ST12 (18%), ST501 (12%), ST167 (6%), ST73 (6%) and ST12 (6%). Phylogenetic analysis corroborated a striking clonal population amongst the studied NLF *E. coli* isolates. The predominant phylogroup detected was B2 (65%). The bla\textsubscript{CTX-M-15} extended-spectrum beta-lactamase gene was present in 53% of isolates (9/17), whilst 64.7% (11/17) isolates were affiliated with pathogenic pathotypes. All extraintestinal pathogenic *E. coli* pathotypes demonstrated \(\beta\)-hemolysis. Our study underscores the presence of critical pathogens and MDR clones amongst non-lactose fermenting *E. coli*. We suggest that non-lactose fermenting *E. coli* be considered equally capable as lactose fermenting forms in causing intestinal and extraintestinal infections. Further, there is a need to undertake systematic, unbiased monitoring of predominant lineages amongst non-lactose fermenting *E. coli* that would help in better treatment and prevention strategies.

**KEYWORDS**
genomic epidemiology, high risk clone, carbapenem resistance, ESBL – *Escherichia coli*, ST131 and non-ST131 lineages, non lactose fermenter
Introduction

*Escherichia coli* is a versatile gram-negative bacterium with huge genetic diversity (Nia Santos Braz et al., 2020). It is the most common organism responsible for opportunistic infections. Its primary habitat includes the lower intestinal tract of humans and animals (Nia Santos Braz et al., 2020). However, infections with variant strains of *E. coli* are responsible for various clinical manifestations ranging from diarrhoea, urinary tract infections, and life-threatening septicemia, resulting in over 2 million deaths every year globally (Hussain et al., 2012; Huang et al., 2021). Moreover, these infections are often associated with cephalosporin and carbapenem-resistant strains, impacting the mortality of patients and imparting huge healthcare costs (Logan and Weinstein, 2017).

*Escherichia coli* are non-pathogenic facultative anaerobic flora of the intestinal tract in humans. The gram-negative *E. coli* bacilli ferment lactose to produce hydrogen sulfide. However, up to 20% of *E. coli* isolates from patients are reported to be atypical, which are slow or non-lactose fermenters due to the deficiency in enzyme lactose permease encoded by the lacY gene (Nicoletti et al., 1988; Hossain, 2012; Chang et al., 2014; Yaratha et al., 2017; Johnson et al., 2019). Non-lactose fermenting (NLF) *E. coli* can be identified as colourless, transparent colonies on MacConkey agar and by negative lactose (sugar) fermentation tests (Hossain, 2012; Siqueira et al., 2021). Various strains of *E. coli* are equipped to cause different forms of enteric and extraintestinal infections in human hosts with varying propensities (Kaper et al., 2004). The pathogenic strains of *E. coli* belong to both lactose fermenting and non-lactose fermenting *E. coli* types. However, little attention is given to these atypical strains of *E. coli*, particularly the non-lactose fermenting (NLF) *E. coli* in routine diagnostic testing laboratories.

*Escherichia coli* is a well-established etiological agent of diarrhoea (Colonna et al., 1992; Nataro and Kaper, 1998; Hossain, 2012). Diarrheagenic strains of *E. coli* fall into different categories possessing distinct pathogenic mechanisms, the three important *E. coli* strains include enteropathogenic *E. coli* (EPEC), enterohemorrhagic *E. coli* (EHEC), and enteroaggregative *E. coli* (EAEC) (Nataro and Kaper, 1998). These *E. coli* categories cause a range of clinical syndromes such as watery diarrhoea in children, persistent diarrhoea, traveller’s diarrhoea and haemolytic uremic syndrome (Nataro and Kaper, 1998). A few studies have reported that NLF *E. coli* strains were identified as typical pathogenic *E. coli* isolates, having a possible diarrheagenic role (Nicoletti et al., 1988; Colonna et al., 1992; Hossain, 2012).

Urinary tract infections (UTIs) represent one of the most common bacterial infections. It is a serious public health problem affecting around 150 million people worldwide (Flores-Mireles et al., 2015). *E. coli* is the most common pathogen responsible for (un)complicated urinary tract infections (Flores-Mireles et al., 2015). Such strains harbour multiple virulence factors, which play their role in the pathophysiology of UTIs. These virulence factors affect bacterial colonisation and evasion of host defences (Hussain et al., 2014; Mazumder et al., 2020a). Recent reports suggest that clonal groups of multi-resistant, multi-virulent *E. coli* are increasingly responsible for UTIs, presenting a challenge for their treatment (Riley, 2014). In recent years, a few studies have demonstrated that NLF *E. coli* have been increasingly isolated from urine specimens in microbiology laboratories and have reported their clinical significance concerning antibiotic resistance, virulence and emerging clones (Chang et al., 2014; Chakraborty et al., 2016; Wu et al., 2017).

Rapid identification of bacterial pathogens in microbiology laboratories is critical for initiating successful infection treatment. Screening of gram-negative bacteria from urine and stool samples is routinely performed on MacConkey agar, and the colourless transparent NLF *E. coli* variants are usually missed in the screening process as little attention is given to these atypical strains of *E. coli*. Although reports of such isolates are limited worldwide, a few studies have reported the occurrence and characterisation of NLF *E. coli* from clinical specimens. They have characterised the strains using only conventional methods. This study aims to decipher the biochemical profiles, population structure, and genomic characteristics of NLF *E. coli* from stool and urine samples isolated at a referral diagnostic centre (International Center for Diarrheal Disease Research, Bangladesh; icddrb) in Dhaka, Bangladesh. Studies such as these will inform us about the evolutionary trajectories of atypical *E. coli* variants and shed light on their clinical and public health significance.

Materials and methods

Bacterial isolates and their biochemical and serological characterisation

A total of 175 *E. coli* isolates were sourced from urine (n = 98; 56%) and stool (n = 77; 44%) culture plates. The isolates were collected randomly from the Clinical Microbiology Laboratory of the icddrb, based in Dhaka, Bangladesh. The isolates were sampled between 2019 and 2020 as part of a larger 1% AMR surveillance study. From this collection of 98 urine and 77 stool *E. coli* isolates, we identified 9 (9.2%) and 8 (10.4%) NLF *E. coli* isolates, respectively. These 17 NLF *E. coli* were subjected to complete biochemical profiling and whole-genome sequencing (WGS) using an Illumina platform (NextSeq 500). NLF *E. coli* were presumptively identified by their inability to ferment lactose on the MacConkey agar. The presence of *Shigella* spp. was ruled out by serotyping using slide agglutination test employing a commercial *Shigella* grouping antisera kit that includes group A (*S. dysenteriae*), B (*S. flexineri*), C (*S. boydii*) and group D (*S. sonnei*) antisera from Denka Seiken Co. Ltd. Tokyo, Japan. Further, confirmation was made by standard biochemical methods that include motility, indole and urease tests (using MIU agar), triple sugar iron (TSI) agar test to visualise glucose and lactose fermentation reactions, catalase test (using 3% H₂O₂), oxidase test (using 1% aqueous solution of tetra-methyl-p-phenylenediamine dihydrochloride), citrate utilisation test (using Simmons citrate
agar) and acetate utilisation test (using Acetate agar) (Hawkey, 2006; Aditi et al., 2017). As mentioned in Table 1, fermentation of other sugars was confirmed using carbohydrate broth with bromothymol blue as a pH indicator, employing previously described methods (Bopp et al., 2003). Voges-Proskauer test was performed in MRVP broth and the reaction was visualised by adding 5% naphthol and 40% potassium hydroxide (KOH) (Hawkey, 2006). Gelatine liquefaction and ONPG tests were performed as described elsewhere (Hawkey, 2006; Nusrat et al., 2017). Decarboxylation reaction of amino acids was performed on decarboxylase broth supplemented by 1% of the appropriate amino acids as described earlier (Brooks and Sodeman, 1974). The Analytical profile index was evaluated using the API 20E kit (bioMérieux) as per the manufacturer's instructions (Table 1). Hemolytic properties of strains were evaluated using 5% sheep blood agar plates. Clearing zones around the colonies were suggestive of beta-haemolysis.

**Antimicrobial susceptibility testing**

The Genome Centre at icddr,b reassessed 17 NLF *E. coli* isolates with an extensive panel of 18 antibiotics (Oxoid, US) by the Kirby-Bauer disc diffusion method. The results were interpreted using the 2019 Clinical and Laboratory Standards Institute (CLSI) guidelines. Upon antibiotic susceptibility testing, the NLF *E. coli* isolates identified as intermediate and resistant were considered as resistant. They were classified as multi-drug resistant (MDR) if they were non-susceptible to at least one agent belonging to three different antimicrobial classes (Magiorakos et al., 2012).

**Whole-genome sequencing**

Total bacterial DNA was isolated and purified from the overnight grown bacterial cultures using the QIAamp DNA Mini kit (Qiagen, Germany). DNA QC and quantification were performed employing a Nanopore spectrophotometer (Thermo Fisher Scientific, United States) and Qubit 4.0 fluorometer (Life Technologies; Baddam et al., 2020). DNA libraries for the short-read paired-end sequencing were prepared using the Nextera DNA Flex library prep kit (Illumina). Size selected and pooled libraries were sequenced at the icddr,b Genome Centre in Illumina NextSeq 500 system using a NextSeq v2.5 reagent kit (2 × 150 bp) (Mazumder et al., 2020b).

**Sequence assembly and annotation**

High-quality reads were obtained by filtering the paired-end data using fastp (Chen et al., 2018). *De novo* assemblies were generated using SPAdes (v.3.11.11) (Bankevich et al., 2012). NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Tatusova et al., 2016) was used for annotating the genome assemblies. The genome statistics from the resulting file were gleaned using the QUAST software (v5.2.0) (Gurevich et al., 2013).

---

**TABLE 1** Microbiological and biochemical characteristics of 17 NLF *E. coli* isolates.

| Sl. no. | Biochemical tests                      | No. (%)          |
|--------|----------------------------------------|------------------|
| 1      | Catalase test                          | 17 (100)         |
| 2      | Oxidase test                           | 0 17 (100)       |
| 3      | TSI agar                               |                  |
| a. Acid production in slant             | 17 (100)         |
| b. Acid production in butt              | 17 (100)         |
| c. Hydrogen sulfide production(H₂S)    | 0 17 (100)       |
| d. Gas production                       | 17 (100)         |
| 4      | Motility Indole Urease Test (MIU)      |                  |
| a. Motility                             | 17 (100)         |
| b. Indole Production                    | 17 (100)         |
| c. Urea hydrolysis                      | 0 17 (100)       |
| 5      | Simmons Citrate reaction test          | 0 17 (100)       |
| 6      | Acetate                                | 17 (100)         |
| 7      | Sugar fermentation                     |                  |
| a. Glucose                              | 17 (100)         |
| b. Lactose                              | 0 17 (100)       |
| c. Sucrose                              | 11 (65)          |
| d. Mannose                              | 17 (100)         |
| e. Mannose                              | 17 (100)         |
| f. Arabinose                            | 17 (100)         |
| g. Sorbitol                             | 17 (100)         |
| h. Mannitol                             | 17 (100)         |
| i. Inositol                             | 0 17 (100)       |
| 8      | Ortho-Nitrophenyl-β-galactoside (ONPG) | 17 (100)         |
| 9      | Gelatine liquefaction                  | 0 17 (100)       |
| 10     | Voges-proskauer                        | 0 17 (100)       |
| 11     | Lysine decarboxylase                   | 17 (100)         |
| 12     | Ornithine decarboxylase                | 14(82) 3 (18)    |
| 13     | Arginine Dihydrolosa                   | 11 (65)          |
| 1      | API profile: 7144532                   | A3, A4, A5, A7, U1, U7, U8, |
| 2      | API profile: 5144572                   | A6, U4, U6       |
| 3      | API profile: 5044552                   | A1, A8, A9       |

**Bacteriological Characteristics [17 (100%)]**

| 1      | Growth on MacConkey Agar               | Colourless and transparent (NLF colony) |
| 2      | Growth on SS agar                      | Colourless and transparent (NLF colony) |
| 3      | Growth on CHROMagar™                   | Pink colour colony                      |
| 4      | Growth Temperature                     | 26–42°C                                  |
| 5      | Gram's staining                        | Gram-negative rods                      |

---
**In silico analysis**

Species identification was confirmed using Kmer Finder software (v3.2) (Hasman et al., 2014; Larsen et al., 2014; Clausen et al., 2018). Phylogenetic groups of the genomes were inferred using the Clermon Typing tool (Beghain et al., 2018). The sequence types (STs) were determined by submitting the contigs to MLST (v2.0; scheme #1) (Larsen et al., 2012). SerotypeFinder 2.0 (Joensen et al., 2015) and CHTyper 1.0 (Roer et al., 2018) software tools were used for elucidating serotypes and clonotypes, respectively.

Resistance genes were screened by BLASTn analysis of the genome assemblies against the data downloaded from the ResFinder database (Bortolaia et al., 2020). The 70% query coverage and 90% identity indicated a positive hit in the genome. Plasmid incompatibility groups were determined by PlasmidFinder (v2.1) (Carattoli et al., 2014). Mobile genetic elements associated with acquired antimicrobial resistance (AMR) genes were determined using Mobile Element finder v1.0.3 (Johansson et al., 2021). Chromosomal point mutations conferring fluoroquinolone resistance were identified by PointFinder software (Zankari et al., 2017). Contigs with blaNDM-1, and blaCTX-M-15 were analysed by BLASTn analysis against the database in NCBI to identify the chromosomal or plasmid origin of these genes.

Virulence gene content of strains was determined by BLASTn analysis of 17 genomes against the *E. coli* database selected in the virulence factor database (VFDB; Chen et al., 2005). The genes with 90% identity and 70% query coverage were considered present. Isolates were classified as extraintestinal pathogenic *E. coli* based on Johnson’s criteria (Johnson and Stell, 2000). The *in silico* analysis described above was performed on default parameters unless otherwise stated.

**Phylogenomic and pangenomic analysis**

The phylogenetic tree was inferred with the reference sequence alignment-based phylogeny builder (REALPHY) software (Bertels et al., 2014). *E. coli* K-12 MG1655 genome (accession no.U00096) was used as the single reference sequence and PhyML (Guindon et al., 2010) was used as the treebuilder. The phylogenetic tree was later visualised using the Interactive Tree of Life (iTOL; Letunic and Bork, 2016). Pan-genome analysis was conducted using the Anvi’o (v7.1) software package (Eren et al., 2015) following the microbial pan-genomics analysis workflow.

**Results**

**Bacterial and biochemical characteristics**

The colourless NLF colonies on the MacConkey agar plate appeared like *Shigella* spp. colonies. Colonies were colourless even in 72-h old MacConkey agar cultures, removing the possibility of late lactose fermenters. Serology of all colonies revealed negative reactions to the four serogroups of *Shigella*: *S. flexneri*, *S. dysenteriae*, *S. boydii*, and *S. sonnei*. Further, streaking of colonies on CHROMagar Orientation media revealed small, pink-red colonies typical of *E. coli*. The isolates were then tested by routine biochemical tests as described in Table 1 and identified as *E. coli*. Notably, all isolates were O-nitrophenyl-beta-D-galactopyranoside positive and did not ferment lactose sugar. Further confirmation was done by generating a biochemical profile for each isolate using 20 reactions of the API 20E strip. The API results gave three distinct profiles: 7144532 (n = 11), 5,044,552 (n = 3) and 5,144,572 (n = 3) and identified all isolates as *E. coli* with 99.8% probability (Table 1). Gram staining confirmed the presence of gram-negative bacilli. The optimum growth temperature of NLF *E. coli* isolates ranged from 26°C to 42°C.

**Molecular analysis and phylogenomics of NLF *Escherichia coli* isolates**

WGS of the 17 NLF *E. coli* study isolates yielded an average genome size of 5,160,336bp (range 4,890,228–5,480,336) and the average GC content was 50.6% (range 50.3–50.7%). KmerFinder software results confirmed all the isolates as *E. coli*. On average, the genome assemblies had a genome coverage of 86-fold (range: 62–124-fold; Table 2).

According to *in silico* MLST, isolates were classified into nine distinct sequence types (STs) and they represented some of the international high-risk lineages (Figure 1). ST131 was the predominant ST that accounted for 4 (23%) of the 17 NLF *E. coli* isolates. ST1193 and ST12 were present in equal proportions, 3 (18%) each, followed by two ST501 strains (12%). Other significant STs detected were ST167 (6%), ST73 (6%) and ST12 (6%). The predominant phylogroup detected was B2 (65%) which comprised all isolates affiliated with ST131, ST1193, ST12 and ST73. This was followed by group A (n = 2: ST167; 1 ST2089); other phylogroups were represented by single isolates (Figure 1). By serogroup, the 4 ST131 isolates were either O16: H4 (n = 2: fumC40: fimH30; Figure 1). All ST1193 isolates were O75: H5 and had fumC13. Other STs exhibited diverse serotypes and CH types (Figure 1).

Based on the alignment of detected core genome SNPs, a phylogenetic tree was built for 17 NLF *E. coli* from our study with the MG1655 genome as the reference genome. There was relatively low diversity of the studied NLF *E. coli* as most strains clustered together, resulting in clades and subclades. The analysis clustered the 17 isolates into two major clades, one tight cluster corresponding to phylogroup B2 and the other loosely clustered clade included strains from A, B1, D and F phylogroup. Compared to NLF *E. coli* from stool, the urine *E. coli* isolates formed a tight clade with strains belonging to ST12, ST1193 and ST131. Three strains from stool origin clustered closely with urine strains.
reinforcing the gut as a possible reservoir of UTI-causing agents. The blaCTX-M-15 MDR phenotype and β-hemolysis did not correspond with the clusters indicating their widespread presence across E. coli lineages. However, the sequence types, serogroups, and CH-types corresponded closely with the phylogenetic clusters of E. coli genomes. The ST131 strains with O16: H5 (fimC40: fimH41) formed a subcluster with ST131 strains having O25: H4 (fimC40: fimH30), indicating there are sub-lineages within ST131.

We also analysed the pan-genomes of the 17 NLF E. coli genomes (Figure 2). The number of core and accessory genes identified was 3,400 and 30,563, respectively. Contrary to the core-genome-based phylogenetic tree, no two genomes shared the same gene content with respect to the pan-genome. A closer look into the pan-genome dendrogram reflects a clustering pattern that mimics the clades of the core genome-based phylogenetic tree (Figure 2). These observations indicate that the 17 NLF E. coli genomes do not differ drastically, even regarding their accessory gene content. They all have a few unique gene families (Figure 2).

**Virulence genes**

Our analysis identified 73 virulence-associated genes out of 335 genes analysed. The distribution of 73 virulence genes amongst 17 NLF E. coli isolates is shown in Figure 3. Overall, 65% (11/17) of all NLF E. coli isolates were assigned to different pathotypes, as classified by the presence of specific virulence markers. Most urine NLF E. coli isolates were assigned to the extraintestinal pathogenic pathotype (67%, 6/9). Suspected diarrheagenic variants comprised 62% (5/8) of the stool NLF E. coli isolates. The most prevalent pathotype amongst diarrheagenic variants was enteropathogenic E. coli (38%), which harboured both aggR and aatA genes, followed by enteropathogenic E. coli (13%) that showed the presence of intimin gene (eae). One isolate from stool origin was also classified as extraintestinal pathogenic E. coli. The remaining six isolates (6/17) that did not belong to any pathotype included three strains from ST1193, two from ST131 and one from ST501. Out of four ST131 studied strains, two O25: H4 serogroup (fimH30) strains qualified as extraintestinal pathogenic and two strains with O16: H5 serogroup (fimH41) did not qualify as extraintestinal pathogenic. Amongst all the pathotypes, the seven (100%) ExpEC strains produced β-hemolysis in 5% sheep blood agar. This hemolytic phenotype was strongly associated with alphahemolysin genes (hlyA, hlyB and hlyD). Although the three ST1193 isolates did not qualify as extraintestinal pathogenic E. coli pathotypes, they all carried a similar set of virulence genes, including adhesins (fimH), toxins (sat, hlyE/clyA), siderophores (chuA, fyuA, sitABCDE, iutA) and others. The NLF E. coli sourced from urine had a higher prevalence of virulence genes, with higher aggregate virulence scores (median 90; range 72–114) compared to those of stool NLF E. coli isolates (median 73; range 62–116).

**Antimicrobial resistance**

The phenotypic and genotypic trends of AMR of the studied NLF E. coli strains were investigated (Figure 4). Overall, the NLF E. coli strains from the stool and urine samples showed high antimicrobial drug resistance rates to nalidixic acid (100%), cefazidime (94%), cephradine (94%), followed by ciprofloxacin (88%), ampicillin (82%), cefotaxime (65%) and cefepime (65%). The resistance rates to cefuroxime (59%), ceftriaxone (53%), co-trimoxazole (41%), imipenem (35%) were moderate and the
resistance rates to gentamicin (18%), amikacin (12%), doxycycline (6%), meropenem (6%) were low. In contrast, there appeared to be no resistance to chloramphenicol, fosfomycin and tigecycline.

Compared to the stool isolates, the urine NLF E. coli isolates had relatively low resistance to amikacin, cefepime, ceftriaxone and cefuroxime; consequently, there were 63% (5/8) MDR isolates in the stool category and 33% (3/9) MDR isolates in urine category.

We detected 31 unique AMR gene alleles belonging to different classes, with most genomes exhibiting at least 6 AMR genes (Figure 4). The AMR genes had a relatively similar distribution between the stool and urine NLF E. coli isolates, with strains from both the groups having a minimum of 6 AMR genes per genome, albeit with variation amongst isolates (E. coli range: stool 3–15; urine 2–13). No specific AMR profile was associated with isolates from a particular clinical source.

A variety of ESBL gene variants encoded cephalosporin resistance. Amongst the identified ESBL genes in NLF E. coli isolates, blaCTX-M-15 was predominant (9/17, 53%), but we also identified blaCTX-M-6, blaTEM-1B, blaSHV-12, and blaOXA-1a genes in several isolates (Figure 4). We observed perfect concordance (100%) between the presence of blaCTX-M-15 gene and resistance phenotypes to the following cephalosporin antibiotics, cefradine, cefotaxime, ceftriaxone, ceftazidime, and cefepime. One blaCTX-M-15 positive stool NLF E. coli harboured a blaNDM-5 gene; this strain (A4) demonstrated difficult-to-treat resistance (DTR) as it was resistant to 14 of 18 antibiotics tested, including all first-line agents. It was sensitive to chloramphenicol, doxycycline, fosfomycin and tigecycline. The probable genome locus of the blaCTX-M-15 gene was detected to be a plasmid for 4 out of 9 NLF E. coli isolates (Table 3). The genetic environment of the blaCTX-M-15 gene in most isolates (56%, 5/9) consisted of the insertion element ISeca1 preceding the locus (Table 3). The blaCTX-M-15 gene was mainly associated with the epidemiologically significant clonal E. coli lineages such as ST131, ST167, ST1193, and ST6303. Interestingly, these strains were all NLF.

We screened for mutations in the quinolone-resistance-determining regions (QRDRs) of gyrA, parC and parE. We identified amino acid substitutions at codon positions S83L (16/17 isolates), D87N (9/17 isolates) in the amino acid sequence of gyrA and S80I (9/17 isolates), S57I (one isolate), E84V (4/17 isolates), E84G (two isolates) and S80R (two isolates) in parC. We also identified mutations at codon positions S45A, I529L and L416F in the parE gene. The gyrA S83L mutation was strongly associated with resistance to ciprofloxacin. The ESBL gene blaCTX-M-15 and gyrA S83L, parC E84V were strongly linked in ST131 and ST167 lineage strains. Additionally, the isolates also harboured acquired AMR genes encoding fluoroquinolone resistance, including QnrS1 (one isolate), QnrS3 (one isolate) and aac(6’)-Ib-cr gene (3 isolates). All the three acquired fluoroquinolone resistance genes were associated with ciprofloxacin resistance.

MDR was frequently observed in isolates with multiple AMR genes identified in eight (47%) isolates. Cephalosporin resistance genes and the gyrA S83L mutation were strongly associated with
the MDR phenotype. PlasmidFinder identified seven unique plasmid replicon groups, with each isolate carrying an average of 2.5 plasmid groups (Figure 4). FIB was the most commonly encountered plasmid group ($n = 13/17$), followed by FII ($n = 9/17$), Col ($n = 8/12$) and FIA ($n = 7/17$). FIA, FIB and Col plasmids were strongly associated with ST131 _E. coli_. The New Delhi metallo-β-lactamase (NDM)-positive _E. coli_ strain (A4) concomitantly hosted three IncF-type replicons, FIA, FIB, and FII groups. We could not associate AMR genes with known plasmid replicons in two strains (A6 and U8) as no replicons were detected in them. Surprisingly, these strains belong to well-known ESBL-producing lineages (ST501 and ST131, respectively).

**Discussion**

We found a high prevalence of epidemiologically significant clonal lineages amongst NLF _E. coli_ isolates from patients with suspected UTIs and enteric infections in a referral diagnostic centre (icddr,b) in Dhaka, Bangladesh. None of the previous studies has focused on the genomic epidemiology of NLF _E. coli_. In the current study, we carried out a comprehensive characterisation of 17 NLF _E. coli_ isolates to delineate their microbiological, biochemical, and molecular characterisation using WGS to identify their population composition and detect their molecular markers of AMR and virulence. This first report forms a baseline study on the genomic epidemiology of clinical NLF _E. coli_ isolates. Knowledge of such strains’ virulence and resistance properties is important to help direct researchers, clinicians, and laboratory technologists screen, analyse and report atypical variants of _E. coli_ recovered from clinical samples.

In line with previous reports (Hossain, 2012; Yaratha et al., 2017), our study revealed that the prevalence rate of NLF _E. coli_ in patients in the community is around 10%. This makes it important to identify NLF _E. coli_ in clinical
specimens, which would help initiate appropriate antimicrobial therapy. Given the difficulty in differentiating between NLF *E. coli* and *Shigella* spp., multiple tests need to be performed as described in this study, including serology, biochemical identification (manual and API) and molecular methods like WGS. It is reported that routine MALDI-TOF MS analysis cannot reliably differentiate between *E. coli* and *Shigella* species (Ling et al., 2019). It can be noted here that the chromogenic media-CHROMagar Orientation approach cannot distinguish between lactose fermenting and NLF *E. coli* and can be used to isolate/identify both types of *E. coli* in a single step.

The predominant sequence types detected in this study were epidemiologically significant high-risk clones, including ST131, ST1193, ST12, ST73 and ST167. Our group previously reported
these clones from Bangladesh amongst ESBL-producing lactose fermenting *E. coli* (Mazumder et al., 2021). None of the studies has reported them from NLF *E. coli* except for ST1193 (Wu et al., 2017). The dominance of clonal STs amongst this collection of NLF *E. coli* is an interesting finding. It warrants close attention because the inclusion of isolates in this study was not based on any AMR phenotypes and genotypes. These STs are predominantly reported to be associated with extraintestinal infections (Pitout, 2012; Riley, 2014; Manges et al., 2019). Also, a review article on global extraintestinal pathogenic *E. coli* STs enlisted these STs to be responsible for the enormous burden of extraintestinal human infections globally (Manges et al., 2019). These highly successful clonal groups are a major mode of spreading AMR by the mechanism of global expansion (Shaik et al., 2017; Tchesnokova et al., 2019).

ST131 *E. coli* is an established pandemic extraintestinal MDR pathogen (Rogers et al., 2011; Hussain et al., 2012, 2014; Nicolas-Chanoine et al., 2014). The four ST131 isolates identified in our collection were of two O and CH types, two ST131 *E. coli* exhibited O25:H4 with fumC40:fimH30 and the other two ST131 *E. coli* had O16:H5 serogroup and fumC40:fimH41 CH type. The two strains with O25:H4 serogroup belonged to the subclade C2, known as the H30Rx sub-lineage, which is the highly drug-resistant and successful sub-lineage of the ST131 clone. ST1193 is reported to be an emerging multi-resistant clonal group (Johnson et al., 2019). The three ST1193 isolates in this study carried the same set of three nonsynonymous chromosomal mutations in gyrA (S83L), gyrA (D87N) and parC (S80I) that confer fluoroquinolone resistance and had O75:H5 serogroup with fimC14: fimH64 CH type, they were all ciprofloxacin-resistant. ST1193 prevalence is increasing rapidly worldwide and is expected to replace the most successful clone, ST131, in the near future (Pitout et al., 2022). The three ST12 isolates in this collection were qualified as ExPEC pathogens and were ciprofloxacin-resistant. ST12 is reported to be one of the predominant STs that cause bloodstream infections (Kallonen et al., 2017; Manges et al., 2019). The ST73 isolate detected in this study harboured 104 of 335 virulence genes, including several markers of extraintestinal pathogenic and sepsis-associated *E. coli*. In contrast, it carried a smaller number of AMR genes. It demonstrated β-hemolysis on 5% sheep blood agar and had the CHtype-fumC24:fimH30. The ST167 strain we identified showed a difficult-to-treat phenotype and carried bla<sub>NDM-1</sub> gene. It qualified as an EAEC pathotype and this strain has been associated with a global spread of ESBL-*E. coli* in humans (Ewers et al., 2012). ST167 associated with bla<sub>NDM-1</sub> gene was reported as an emerging carbapenem-resistant high-risk clone (Garcia-Fernandez et al., 2020).

Overall, the NLF *E. coli* isolates carried the AMR genes and were moderately resistant to antibiotics, but a higher number were assigned to different pathotypes. The isolates carried extensive virulence genes that belonged to different categories, including adhesion, invasion, colonisation, and iron uptake ability. Nevertheless, the combination of resistance and virulence in these NLF *E. coli* lineages could contribute to their increased fitness and potential to expand globally.

This study has multiple limitations. First, limited clinical and epidemiological data were available. Second, a small sample size constrained the power to generalise our findings and confirm statistical associations, particularly with the high-risk clones. The main gain of the study is that emphasis on inclusion of AMR phenotypes and genotypes. These STs are predominantly reported to be associated with extraintestinal infections (Pitout, 2012; Riley, 2014; Manges et al., 2019). Also, a review article on global extraintestinal pathogenic *E. coli* STs enlisted these STs to be responsible for the enormous burden of extraintestinal human infections globally (Manges et al., 2019). These highly successful clonal groups are a major mode of spreading AMR by the mechanism of global expansion (Shaik et al., 2017; Tchesnokova et al., 2019).

ST131 *E. coli* is an established pandemic extraintestinal MDR pathogen (Rogers et al., 2011; Hussain et al., 2012, 2014; Nicolas-Chanoine et al., 2014). The four ST131 isolates identified in our collection were of two O and CH types, two ST131 *E. coli* exhibited O25:H4 with fumC40:fimH30 and the other two ST131 *E. coli* had O16:H5 serogroup and fumC40:fimH41 CH type. The two strains with O25:H4 serogroup belonged to the subclade C2, known as the H30Rx sub-lineage, which is the highly drug-resistant and successful sub-lineage of the ST131 clone. ST1193 is reported to be an emerging multi-resistant clonal group (Johnson et al., 2019). The three ST1193 isolates in this study carried the same set of three nonsynonymous chromosomal mutations in gyrA (S83L), gyrA (D87N) and parC (S80I) that confer fluoroquinolone resistance and had O75:H5 serogroup with fimC14: fimH64 CH type, they were all ciprofloxacin-resistant. ST1193 prevalence is increasing rapidly worldwide and is expected to replace the most successful clone, ST131, in the near future (Pitout et al., 2022). The three ST12 isolates in this collection were qualified as ExPEC pathogens and were ciprofloxacin-resistant. ST12 is reported to be one of the predominant STs that cause bloodstream infections (Kallonen et al., 2017; Manges et al., 2019). The ST73 isolate detected in this study harbourled 104 of 335 virulence genes, including several markers of extraintestinal pathogenic and sepsis-associated *E. coli*. In contrast, it carried a smaller number of AMR genes. It demonstrated β-hemolysis on 5% sheep blood agar and had the CHtype-fumC24:fimH30. The ST167 strain we identified showed a difficult-to-treat phenotype and carried bla<sub>NDM-1</sub> gene. It qualified as an EAEC pathotype and this strain has been associated with a global spread of ESBL-*E. coli* in humans (Ewers et al., 2012). ST167 associated with bla<sub>NDM-1</sub> gene was reported as an emerging carbapenem-resistant high-risk clone (Garcia-Fernandez et al., 2020). Overall, the NLF *E. coli* isolates carried the AMR genes and were moderately resistant to antibiotics, but a higher number were assigned to different pathotypes. The isolates carried extensive virulence genes that belonged to different categories, including adhesion, invasion, colonisation, and iron uptake ability. Nevertheless, the combination of resistance and virulence in these NLF *E. coli* lineages could contribute to their increased fitness and potential to expand globally.

This study has multiple limitations. First, limited clinical and epidemiological data were available. Second, a small sample size constrained the power to generalise our findings and confirm statistical associations, particularly with the high-risk clones. The main gain of the study is that emphasis on inclusion of AMR phenotypes and genotypes. These STs are predominantly reported to be associated with extraintestinal infections (Pitout, 2012; Riley, 2014; Manges et al., 2019). Also, a review article on global extraintestinal pathogenic *E. coli* STs enlisted these STs to be responsible for the enormous burden of extraintestinal human infections globally (Manges et al., 2019). These highly successful clonal groups are a major mode of spreading AMR by the mechanism of global expansion (Shaik et al., 2017; Tchesnokova et al., 2019).
to accomplish successful patient treatment and to feed data to AMR surveillance programs. Further national and regional multicentre “one health” studies are required to ascertain the significance of NLF E. coli as significant pathogens impacting public health.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article.

Author contributions

RM and AH designed the study and drafted the manuscript. RM performed genome sequencing. AM, AH, and RM carried out the bioinformatics analyses and interpretation of results and prepared tables and figures. RM and SH performed microbiological tests. JP, SC, MA, DA, TC, and DM contributed to the discussions and reviewed the manuscript. DM supervised the study. All authors contributed to the article and approved the submitted version.

References

Aditi, F. Y., Rahman, S. A., and Hossain, M. M. (2017). A study on the microbiological status of mineral drinking water. *Open Microbiol. J.*, 11, 31–44. doi: 10.2174/1874285801711010031

Baddam, R., Sarkar, N., Ahmed, D., Mazumder, R., Abdullah, A., Morshed, R., et al. (2020). Genome dynamics of vibrio cholerae isolates linked to seasonal outbreaks of cholera in Dhaaka, Bangladesh. *MBio* 11:e03339-19. doi: 10.1128/mbio.03339-19

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Beghin, J., Brider-Nahmias, A., Le Nagard, H., Denamur, E., and Clermont, O. (2018). Clermon typing: an easy-to-use and accurate in silico method for Escherichia genus strain phylotyping. *Microb. Genomics* 4:192. doi: 10.1099/mgen.0.000192

Bertels, F., Silander, O. K., Pachkov, M., Rainey, P. B., and Van Nimwegen, E. (2014). Automated reconstruction of whole-genome phylogenies from short-sequence reads. *Mol. Biol. Evol.* 31, 1077–1088. doi: 10.1093/molbev/msu088

Bopp, C. A., Brenner, F. W., Fields, P. I., and Wells, J. G. (2003). *Manual of Clinical Microbiology*. 8th Edn., Washington, DC: ASM Press.

Bortolaia, V., Kaas, R. S., Ruppe, E., Roberts, M. C., Schwarz, S., Cattoir, V., et al. (2020). Res finder 4.0 for predictions of phenotypes from genotypes. *J. Antimicrob. Chemother.* 75, 3491–3500. doi: 10.1093/jac/dkaa345

Brooks, K., and Sodeman, T. (1974). A rapid method for determining deoxyribonuclease and deoxyribolase activity. *J. Clin. Pathol.* 27, 148–152. doi: 10.1136/jcp.27.2.148

Carattoli, A., Zankari, E., Garcia-Fernandez, A., Voldby Larsen, M., Lund, O., Villa, L., et al. (2014). In Silico detection and typing of plasmids using plasmid finder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* 58, 3895–3903. doi: 10.1128/AAC.02412-14

Chakraborty, A., Adhikari, P., Shenoy, S., and Saralaya, V. (2016). Virulence factor profiles, phylogenetic background, and antimicrobial resistance pattern of lactose fermenting and nonlactose fermenting *Escherichia coli* from extraintestinal sources. *Ind. J. Pathol. Microbiol.* 59, 180–184. doi: 10.4103/0377-4929.182032

Chang, J., Yu, J., Lee, H., Ryu, H., Park, K., and Park, Y. I. (2014). Prevalence and characteristics of lactose non-fermenting Escherichia coli in urinary isolates. *J. Infect. Chemother.* 20, 738–740. doi: 10.1016/J.JTAC.2014.07.005

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., et al. (2005). VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Res.* 33, D325–D328. doi: 10.1093/nar/gki008

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Faspt: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, 1884–1890. doi: 10.1093/bioinformatics/bby560

Clausen, P. T. L. C., Aarestrup, F. M., and Lund, O. (2018). Rapid and precise alignment of raw reads against redundant databases with KMA. *BMC Bioinformatics* 19, 1–8. doi: 10.1186/s12859-018-2336-6/CTABLES/2

Colonna, B., Ranucci, L., Fradiani, P. A., Casalino, M., Calconi, A., and Nicoletti, M. (1992). Organisation of aerobactin, hemolysin, and antibacterial resistance genes in lactose-negative *Escherichia coli* strains of serotype O4 isolated from children with diarrhea. *Infect. Immun.* 60, 5224–5231. doi: 10.1128/IAI.60.12.5224-5231.1992

Een, A. M., Eisen, O. C., Quince, C., Vineis, J. H., Morrison, H. G., Sogin, M. L., et al. (2015). Anvi: an advanced analysis and visualisation platform for omics data. *Peer J* 2015:1319. doi: 10.7717/PeerJ.1319/SUPP-5

Ewers, C., Bethe, A., Semmler, T., Guenther, S., and Wieler, L. H. (2012). Extended-spectrum β-lactamase-producing and AmpC-producing *Escherichia coli* from livestock and companion animals, and their putative impact on public health: a global perspective. *Clin. Microbiol. Infect.* 18, 646–655. doi: 10.1111/j.1469-0691.2012.03850.X

Flores-Mireles, A. L., Walker, J. N., Caparon, M., and Hultgren, S. J. (2015). Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat. Rev. Microbiol.* 13, 269–284. doi: 10.1038/nrmicro3432

Garcia-Fernandez, A., Villa, L., Bibolino, G., Bressan, A., Trancassini, M., Pietropaolo, V., et al. (2020). Novel insights and features of the NDM-5-producing *E. coli*. *Antimicrob. Agents Chemother.* 59, 180–184. doi: 10.1128/AAC.02412-14

Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenetic trees. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Gurevich, A., Saveliev, V., Vyakhni, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/BIOINFORMATICS/BTT086

Acknowledgments

icdbr is grateful to the governments of Bangladesh, Canada, Sweden, and the United Kingdom for providing core/unrestricted support for its operations and research. The work is funded through a Royal Society International Collaboration Award (ref. ICA1R1/191309).

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.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.
Manges, A. R., Grum, H. M., Guo, A., Edens, T. J., Fikke, C. D., and Pitout, J. D. (2019). Global extraintestinal pathogenic Escherichia coli (ExPEC) lineages. Clin. Microbiol. Rev. 32:e00135-18. doi: 10.1128/CMR.00135-18

Mazumder, R., Abdullah, A., Ahmed, D., and Hussain, A. (2020a). High prevalence of bla-cr-m-15 gene among extended-spectrum β-lactamase-producing Escherichia coli isolates causing extraintestinal infections in Bangladesh. Antibiotics. 9:796. doi: 10.3390/antibiotics9110796

Mazumder, R., Hussain, A., Husain, A., Ahmed, D., and Mondal, D. (2020b). Draft genome sequence of Chromobacterium violaceum RDN09, isolated from a patient with a wound infection in Bangladesh. Microbiol. Resour. Announc. 9.e00957-20. doi: 10.1128/mra.00957-20

Mazur, R., Hussain, A., Abdullah, A., Islam, M. N., Sadeque, M. T., Muniruzzaman, S. M., et al. (2021). International high-risk clones among extended-Spectrum β-lactamase-producing Escherichia coli in Dhaka, Bangladesh. Front. Microbiol. 12:2843. doi: 10.3389/FMICB.2021.736446/BIBTEX

Nataro, J. P. and Kaper, J. B. (1998). Diarrheagenic Escherichia coli. Clin. Microbiol. Rev. 11, 142–201. doi: 10.1128/CMR.11.1.142/FORMAT/EPUB

Nia Santos Braz, V., Melchior, K., and Gallina Moreira, C. (2020). Escherichia coli as a multifaceted pathogenic and versatile bacterium. Front. Cell. Infect. Microbiol. 10.3389/fcimb.2020.548492 doi: 10.3389/fcimb.2020.548492

Nicolaou-Cassimatis, M. H.-M., Bertrand, X., and Madec, J.-Y. (2014). Escherichia coli ST131, an intriguing clonal group. Clin. Microbiol. Rev. 27, 543–574. doi: 10.1128/CMR.00125-13

Nicoletti, M., Superti, F., Conti, C., Calconi, A., and Zagaglia, C. (1988). Virulence factors of lactose-negative Escherichia coli strains isolated from children with diarrhea in Somalia. J. Clin. Microbiol. 26, 524–529. doi: 10.1128/JCM.26.3.524-529.1988

Nusrat, N., Riddwan, B. R., Anowar, B., and Humaira, A. (2017). Isolation, identification and molecular characterisation of rhizobium species from Sesbania bispinosa cultivated in Bangladesh. Afr. J. Agric. Res. 12, 1874–1880. doi: 10.5957/AJAR201712321

Pitout, J. D. (2012). Extraintestinal pathogenic Escherichia coli: a combination of virulence with antibiotic resistance. Front. Microbiol. 3:9. 10.3389/fmicb.2012.00009

Pitout, J. D., Pirrano, G., Chen, L., Devinney, R., and Matsuura, Y. (2022). Escherichia coli ST1193: following in the footsteps of E. coli ST131. Antimicrob. Agents Chemother. 66:e0051122. doi: 10.1128/AAC.00511-22

Riley, L. W. (2014). Pandemic lineages of extraintestinal pathogenic Escherichia coli. Clin. Microbiol. Infect. 20, 380–390. doi: 10.1111/1469-0691.12646

Roer, L., Johannesen, T. B., Hansen, E. S., Stegger, M., Tchesnokova, V., Sokurenko, E., et al. (2014). CHTyper, a web tool for subtypung of Extraintestinal pathogenic Escherichia coli based on the fumC and fumH alleles. J. Clin. Microbiol. 56, 63–81. doi: 10.1128/JCM.00063-18

Roberts, B. A., Siddijah, H. E., and Paterson, D. L. (2011). Escherichia coli O25B-ST131: a pandemic, multi-resistant, community-associated strain. J. Antimicrob. Chemother. 66, 1–14. doi: 10.1093/JAC/DQX415

Shaik, S., Ranjan, A., Tiwari, S. K., Hussain, A., Nandanwar, N., Kumar, N., et al. (2017). Comparative genomic analysis of globally dominant ST131 clone with other epidemiologically successful Extraintestinal pathogenic Escherichia coli (ExPEC) lineages. Millegis. 8:e01596-17. doi: 10.1128/MBio.01596-17

Siquestra, F. M., De Carli, S., Lopes, E. C., Machado, L., Vieira, T. R., Pöppl, G., et al. (2021). Non-lactose-fermenting uropathogenic Escherichia coli from dogs: virulence profile characterization. Lett. Appl. Microbiol. 72, 596–603. doi: 10.1111/1751-7518.14354

Tatusova, T., Dicuccio, M., Badretdin, A., Chetvernin, V., Nawrocki, E. P., et al. (2016). NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 44, 6614–6624. doi: 10.1093/NAR/GKW569

Tchesnokova, V., Raday, M., Chattopadhyay, S., Larson, L., Lee Weaver, J., Kissida, D., et al. (2019). Pandemic fluoroquinolone-resistant Escherichia coli clone ST1193 emerged via simultaneous homologous recombination in 11 gene loci. Proc. Natl. Acad. Sci. U. S. A. 116, 14740–14748. doi: 10.1073/pnas.1900211116

Wu, J., Lan, F., Lu, Y., He, Q., and Li, B. (2017). Molecular characteristics of ST1193 clone among phylogenetic group B2 non-ST131 fluoroquinolone-resistant Escherichia coli. Front. Microbiol. 8:2294. doi: 10.3389/FMICB.2017.02294/BIBTEX

Yazhga, G., Perloff, S., and Changala, K. (2017). Lactose vs non-lactose fermenting E. coli: epidemiology, clinical outcomes, and resistance. Open Forum Infect. Dis. 4, S589–S590. doi: 10.1093/OFID/OFX163.1546

Zanardi, E., Allessie, R., Roensen, K. G., Cavaco, L. M., Lund, O., and Aarestrup, F. M. (2017). PointFinder: A novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. J. Antimicrob. Chemother. 72, 2764–2768. doi: 10.1093/jac/dzx217

Burgess Publishing Co.

Frontiers in Microbiology 11