Microbes and Infectious Diseases 2025; 6(3): 4035-4046

Microbes and Infectious Diseases

Journal homepage: https://mid.journals.ekb.eg/

Original article

E. coli clinical isolates: Phylogrouping and virulence factors in Alexandria, Egypt

Shahinda Rezk [1,2 *], Ahmed Jamal [1], Abeer Ghazal [1], Aliaa Aboulela [1]

1- Department of Microbiology, Medical Research Institute, Alexandria University, Egypt.
2- Institute of Virology, Medical Center-University of Freiburg, Freiburg 79104, Germany

ARTICLE INFO

Article history:
Received 12 January 2025
Received in revised form 1 February 2025
Accepted 4 February 2025

Keywords:
Phylogroups
Virulence genes
Antibiotic resistance

ABSTRACT

Background: *Escherichia coli (E. coli*) can be categorized into many pathotypes which represent groups of strains with specific pathogenic characteristics. **Aim:** This study aimed to identify phylogroups and different virulence factors in *E coli* isolated from clinical specimens. **Methods:** Fifty clinical *E. coli* isolates were included in this study. Identification and antibiotic susceptibility testing were performed by Vitek® 2 compact system. Virulence genes detection and phylogroup gene alleles were performed using SYBR Green Real-time PCR. **Results:** The results revealed high antibiotic resistance; 62% of the isolates were multidrug resistant (MDR) and 8% were extensively drug resistant (XDR). The most prevalent virulence gene was *papG* (78%) while the least was *hlyA* (22%). Phylogroup B2 was the most predominant (52%), followed by (10%) group D, (8%) group C and group E each, (6%) group B1, (4%) clade1 and (2%) group F. **Conclusion:** Phylogroup B2 had all the virulence genes investigated in this study and it was more associated with *sat* and *hlyA* virulence genes.

Introduction

Escherichia coli (E. coli) has the potential to cause a vast variety of infections including diarrheal infections, urinary tracts infections, septicemia, and meningitis. It has been classified into different phylotypes which diverge in metabolic features, virulence traits, resistance to antimicrobials, and tendency to induce infections. Accordingly, allocating *E. coli* strains to previously described phylogroups would disclose a wealth of information [1, 2].

In 2000, a triplex PCR method was developed by Clermont and colleagues to detect (*chuA*, *yjaA*, and *TspE4.C2*) genes in *E. coli* strains. Concerning the presence or absence of these three genes, an *E. coli* strain can be assigned to one of the major phylogroups, A, B1, B2, or D [3]. Few years later a quadruplex PCR was conducted in 2013 by Clermont in order to categorize an *E. coli* strain into one of the following phylogroups: clade I, B1, B2, C, D, E, F, and A by adding a fourth gene target, arpA, to the three candidate markers [4].

It has been declared that *E. coli* phylotyping ensured that extraintestinal infections are primarily related to phylotype B2 and, to a lower degree to phylotype D, on the other hand the majority of the commensals seemed commonly fitted to phylotypes B1 and A. Pathogenic *E. coli* typically exhibit differences in both their

DOI: 10.21608/MID.2025.352346.2458
* *Corresponding author:* Shahinda Rezk
E-mail address: shahinda.rezk@alexu.edu.eg

phylogenetic backgrounds and virulence characteristics [5].

E. coli produce many virulence factors which contribute in attachment, invading host tissues, biofilm development, competing with other bacteria and evade host immunogenic responses. These factors are generally divided as superficial factors and secreted factors [6].

Adhesion is mainly arbitrated to surface factors such as fimbriae, afimbrial adhesives and outer membrane proteins (e.g., type 1 fimbriae, pyelonephritis associated pili, Afa,…etc.) which are coded by clusters of genes termed as pathogenicity islands and other mobile genetic elements [7].

Moreover, it possesses and produces many toxigenic proteins such as (hemolysins, shiga like toxin, and heat stable or heat liable toxin) and iron acquisition systems like (enterobactin and salmochelin) which improve their pathogenicity and surviveness [8].

This study aimed to identify phylogroups and detect specific virulence factors in *E. coli* isolates from clinical specimens. We sought to examine the relationship between phylogenetic groups and virulent gene profiles, with the goal of enhancing our understanding of *E. coli* pathogenicity.

Materials and methods

This study was carried out on 50 *E. coli* strains isolated from clinical specimens of hospitalized patients referred to the Microbiology department of the Medical Research Institute, Alexandria University from June 2021 to September 2022.

Bacterial identification and antibiotic susceptibility testing (AST) were done by Vitek® 2 compact system (BioMerieux, France). Antibiotic susceptibility testing was done using AST-N222 card including (ampicillin, ampicillin/sulbactam, piperacillin/tazobactam, cefazolin, cefoxitin, ceftazidime, ceftriaxone, cefepime, meropenem, impenem, amikacin, gentamicin, tobramycin, ciprofloxacin, levofloxacin, trimethoprim/sulfamethoxazole, colistin, minocycline, and nitrofurantion).

This was followed by DNA extraction from the bacterial isolates by boiling method. The extracted DNA was stored at -20°C.

Detection of *E. coli* phylogroups and virulence genes

Both Phylotypes and the virulence genes detection were using SYBR Green Real Time PCR (Thermo Fisher Scientific, United States). The reaction was accomplished on a thermal cycler with a programmed cycling condition: Initial denaturation for 10 mins at 95 ° C for one cycle followed by 40 cycles of (Denaturation at 95 ° C for 15 sec, annealing for 30 sec and extension at 72° C for 30 secs. The annealing temperatures for each pair of primers used is described in **tables (1, 2).** A meting curve analysis was done to determine the specificity of the products. Phylotyping was done according to Clermont classification. The isolates were assigned to different phylogroups: A, B1, B2, C, D, E, F, and clade I.

Statistical analysis

Descriptive and statistical analyses were performed in RStudio by R platform (version R.4.2.3). Finalfit, ggplot2, FactoMineR, factoextra and Rstatix are the main used packages. Categorical data is presented in count and percent while numerical data are summarized in median and interquartile range (IQR). Chi-square or Fisher-Exact tests were used for categorical data as appropriate. Wilcoxon rank sum test was used to test scored data. Also, dendrogram of detected phylotypes according to Clermont classification was done by hierarchical clustering. Moreover, we performed correspondence analysis (CA) for deeper understanding of relationships between phylotypes and detected genes. *P*-value was considered statistically significant at level ≤ 0.05. Illustrative figures including pie chart, bar plots and biplot were generated through either Microsoft Excel or R Studio.

Table 1. Primers used for the detection of *E. coli* phylogroups and their annealing temperature.

Annealing temperature °C	Primers	Nucleotide sequences	References
54	*chuA*	F 5′-ATGGTACCGGACGAACCAAC-3′	[4,9]
		R 5′-TGCCGCCAGTACCAAAGACA-3′	
50	*yjaA*	F 5′-CAAACGTGAAGTGTCAGGAG-3′	
		R 5′-AATGCGTTCCTCAACCTGTG-3′	
56	*TspE4.C2*	F 5′-CACTATTCGTAAGGTCATCC-3′	
		R 5′-AGTTTATCGCTGCGGGTCGC-3′	
54	*arpA*	F 5′-AACGCTATTCGCCAGCTTGC-3′	
		R 5′-TCTCCCCATACCGTACGCTA-3′	
57	*arpA group E*	F 5′-GATTCCATCTTGTCAAAATATGCC-3′	[4,9]
		R 5′-GAAAAGAAAAAGAATTCCCAAGAG-3′	
57	*trpA group C*	F 5′-AGTTTTATGCCCAGTGCGAG-3′	[4,9]
		R 5′-TCTGCGCCGGTCACGCCC-3′	
50	*trpA internal control*	F 5′-CGGCGATAAAGACATCTTCAC-3′	[4,9]
		R 5′-GCAACGCGGCCTGGCGGAAG-3′	

Table 2. Primers used for the detection of chosen *E. coli* virulence genes.

Annealing temperature °C	Primers	Nucleotide sequences	References
55	*pap G II*	F 5’-GGGATGAGCGGGCCTTTGAT-3’	[10]
		R 5’- CGGGCCCCCAAGTAACTCG-3’	
47	*cnf1*	F 5’-AAGATGGAGTTICCTA’IGCAGGAG-3’	[11]
		R 5’-CATTCAGAGTCCTGCCCTCATTATT-3’	
55	*sfa*	F 5’-GTGGATACGACGATTACTGTG-3’	[12]
		R 5’-CCGCCAGCATTCCCTGTATTC-3’	
57	*hlyA*	F 5’-AGATTCTTGGGCATGTATCCT-3’	[13]
		R 5’-TTGCTTTGCAGACTGTAGTGT-3’	
55	*sat*	F 5’-GGTATTGATATCTCCGGTGAAC-3’	[14]
		R 5’-ATAGCCGCCTGACATCAGTAAT-3’	

Results

This study was carried out on 50 *E. coli* isolates obtained from multiple types of specimens from hospitalized patients. Most of the isolates (70%) were obtained from pediatric patients, with median age of 8.5 years. The patient cohort consisted of 60% females and 40% males. The majority of the clinical isolates were from urine samples (76.0%), followed by wound swabs (16.0%), blood (6.0%), and sputum (2.0%).

In the analysis of antibiotic resistance among 50 *E. coli* isolates, varying levels of resistance were observed across different antibiotic classes. Penicillin resistance was highest for ampicillin (94%) and lowest for piperacillin/tazobactam (24%). Cephalosporin resistance ranged from 74% for cefazolin to 26% for cefoxitin. Notably, 18% of isolates exhibited resistance to both carbapenems tested (imipenem and meropenem). Quinolone resistance was observed in 66% of isolates. Aminoglycoside resistance varies, with 26% resistant to tobramycin and 6% to amikacin. Additionally, 60% of isolates were resistant to trimethoprim/sulfamethoxazole, while only 4% showed resistance to minocycline. Interestingly, all 50 *E. coli* isolates remained susceptible to polymyxin-B (colistin). In a subgroup analysis of 38 urinary *E. coli* isolates, 89% were sensitive to nitrofurantoin, with the remaining 11% showing intermediate resistance. The majority 31 (62%) of 50 *E. coli* isolates were grouped as multidrug resistant (MDR) and 4 (8%) isolates were grouped as extensively drug resistant (XDR). Only 15 (30%) isolates were non-multidrug resistant isolates.

Escherichia coli (E. coli) phylotyping showed a great variety among *E. coli* clinical isolates. Seven phylogroups were detected among *E.*

coli clinical isolates, the majority (52%) were classified as group B2 followed by (10%) group D, (8%) for group C and group E each, (6%) group B1, (4%) clade1, and (2%) group F. However, (10%) of the isolates were unclassified. Genes combinations assigned to different phylogroups are shown in **table (3).** The comparison of the phylotypes according to the sample types is shown in **table (4).**

Assessment of the 31 multidrug-resistant (MDR) isolates revealed a predominance of phylogroup B2, accounting for 61.3% (19 isolates). Phylogroups D, C, and E each represented 9.7% (3 isolates) of the MDR isolates, while phylogroup B1 constituted 3.2% (1 isolate). The remaining 6.5% (2 isolates) could not be classified into known phylogroups. Among the extensively drug-resistant (XDR) isolates, there was an equal distribution across phylogroups B2, D, E, and unknown, with each group representing 25% (1 isolate) of the XDR population. In contrast, the non-MDR isolates were predominantly of phylogroup B2, comprising 40% (6 isolates) of this category.

Regarding the prevalence of the virulence genes, Pap G was the most prevalent virulence gene detected among *E. coli* isolates (78%), followed by *sfa* in 50%, *sat* in 32%, *cnf1* in 28%, and *hlyA* in 22%.

This study revealed a diverse distribution of virulent genes across *E. coli* phylotypes shown in **table (5).** In phylogroup B1, *cnf1* and *sfa* were most prevalent, each detected in 66.7% of isolates, while *papG* was found in 33%. Phylogroup B2 exhibited all virulent genes in varying frequencies. For phylogroup C, *papG* and *sfa* were identified in 75% and 50% of isolates, respectively. Phylogroup D demonstrated high prevalence of *papG* (100%), *sfa* (80%), and *sat* (40%). In phylogroup E, *papG* and *sfa* were equally common (50% each), with *sat* present in 25% of isolates. Phylogroup F showed 100% prevalence for both *papG* and *cnf1*. Clade I exhibited *papG* in all isolates (100%), with *cnf1* and *sfa* each detected in 50%. Among unclassified isolates, *papG* was universally present (100%), while *cnf1, sfa,* and *sat* were equally distributed (40% each).

The comparison of virulence genes presented in B2 phylotype and non-B2 phylotypes shows no statistically significant differences except for *hlyA* virulence gene which was more presented in B2 phylogroup than other phylogroups ($p = 0.01$), also there was no correlation detected among phylogroups and number of virulence genes in the 50 *E. coli* isolates in the current study.

This study investigated the associations between *E. coli* phylotypes and virulence genes among 50 isolates. The analysis revealed distinct patterns of gene-phylogroup relationships: Phylogroup B2 showed a stronger association with sat compared to other phylogroups, the *papG* gene demonstrated a ubiquitous presence across all phylogroups, as evidenced by its central position in the graphical representation, the *sfa* gene exhibited a closer relationship with phylogroup D and *cnf1* displayed a stronger association with phylogroups B1 and F. These findings, illustrated in **figure (1)**, suggest that certain virulence genes may have preferential distribution among specific *E. coli* phylogroups, potentially influencing their pathogenic capabilities. This distribution pattern could provide insights into the evolution and adaptation of different *E. coli* lineages.

In the current study there was no association between sample type and virulence genes except for *hylA* gene which was more presented in blood samples (66.7%) than in urine samples (23.7%) and was not detected in sputum or wound samples ($p = 0.055$).

Certain *E. coli* phylogroups showed notable associations with specific antibiotics. Phylotypes C and E demonstrated significant resistance to cefazolin ($p= 0.002$). In contrast, phylotypes B1, clade 1, and F exhibited significant sensitivity to several antibiotics: ceftazidime ($p= 0.003$), ceftriaxone ($p = 0.009$), and cefepime ($p= 0.009$). Furthermore, clade 1 and phylotype F were found to be significantly sensitive to carbapenems (meropenem and imipenem) ($p=0.033$) and quinolones (ciprofloxacin and levofloxacin) ($p=0.057$).

Table 3. Phylogenetic typing by Clermont classification of *E. coli* isolates (N= 50).

Phylotype	n	*arpA*	*chuA*	*yjaA*	*tspE4.C2*	Total No. (%)
B2	23	-	+	+	+	26 (52)
B2	2	-	+	-	+	
B2	1	-	+	+	-	
E or Clade I	6	+	+	+	-	6 (12)
D or E	3	+	+	-	-	5 (10)
D or E	2	+	+	-	+	
Unknown	1	+	-	+	+	5 (10)
Unknown	4	+	+	+	+	
A or C	4	+	-	+	-	4 (8)
B1	3	+	-	-	+	3 (6)
F	1	-	+	-	-	1 (2)
Total						50

n: number of isolates

Table 4. Comparison of phylotypes of *E. coli* isolates according to sample types (N= 50).

Phylotype	Uroisolates n= 38	Non Uroisolates* n= 12	^{MC}p-value
B2	20 (52.6)	6 (50.0)	1.000
D	4 (10.5)	1 (8.3)	1.000
Unknown	4 (10.5)	1 (8.3)	1.00
C	3 (7.9)	1 (8.3)	1.000
E	1 (2.6)	3 (25.0)	0.044
B1	3 (7.9)	0 (0.0)	0.6
Clade I	2 (5.3)	0 (0.0)	1.000
F	1 (2.6)	0 (0.0)	1.000

Data is presented in count and percent.

* not Uroisolates include wound swabs, blood and sputum.

^{MC}p: p-value obtained by Chi-square test with Monte-Carlo simulation.

Bold p-value presents a statistically significant difference between groups at alpha level < 0.05.

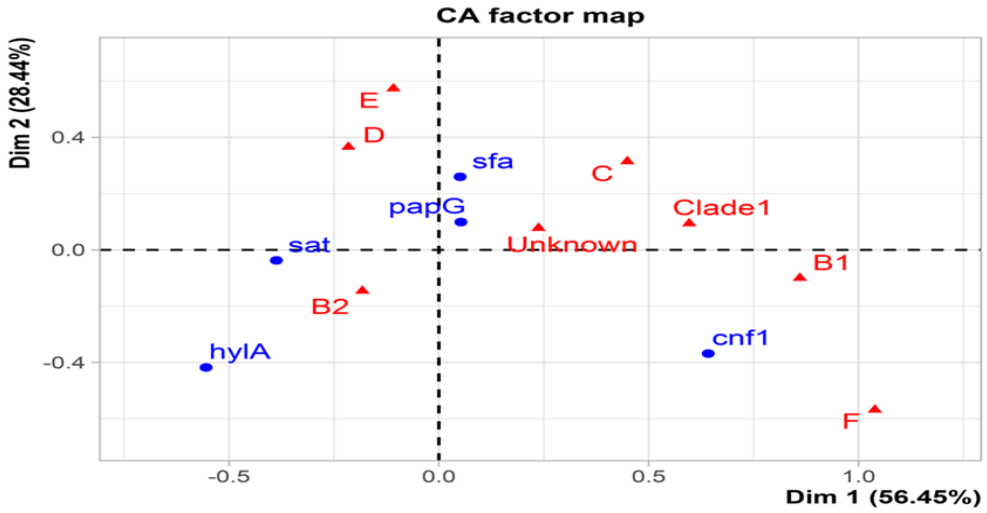
Table 5. Distribution of various detected genes among phylotypes of *E. coli* isolates (N= 50).

Genes \ Phylotype		B1, n= 3	B2, n= 26	C, n= 4	Clade, n= 2	D, n= 5	E, n= 4	F, n= 1	Unknown, n= 5	^{FE}P-value
papG	Present	1 (33.3)	20 (76.9)	3 (75.0)	2 (100.0)	5 (100.0)	2 (50.0)	1 (100.0)	5 (100.0)	0.253
	Absent	2 (66.7)	6 (23.1)	1 (25.0)	0 (0.0)	0 (0.0)	2 (50.0)	0 (0.0)	0 (0.0)	
cnf1	Present	2 (66.7)	7 (26.9)	1 (25.0)	1 (50.0)	0 (0.0)	0 (0.0)	1 (100.0)	2 (40.0)	0.206
	Absent	1 (33.3)	19 (73.1)	3 (75.0)	1 (50.0)	5 (100.0)	4 (100.0)	0 (0.0)	3 (60.0)	
hlyA	Present	0 (0.0)	10 (38.5)	0 (0.0)	0 (0.0)	1 (20.0)	0 (0.0)	0 (0.0)	0 (0.0)	0.369
	Absent	3 (100.0)	16 (61.5)	4 (100.0)	2 (100.0)	4 (80.0)	4 (100.0)	1 (100.0)	5 (100.0)	
sfa	Present	2 (66.7)	12 (46.2)	2 (50.0)	1 (50.0)	4 (80.0)	2 (50.0)	0 (0.0)	2 (40.0)	0.899
	Absent	1 (33.3)	14 (53.8)	2 (50.0)	1 (50.0)	1 (20.0)	2 (50.0)	1 (100.0)	3 (60.0)	
sat	Present	0 (0.0)	11 (42.3)	0 (0.0)	0 (0.0)	2 (40.0)	1 (25.0)	0 (0.0)	2 (40.0)	0.642
	Absent	3 (100.0)	15 (57.7)	4 (100.0)	2 (100.0)	3 (60.0)	3 (75.0)	1 (100.0)	3 (60.0)	

Data is presented in count and percent. ^{FE}p: Fisher Exact test.

Bold p-value presents a statistically significant difference between groups at alpha level < 0.05.

Figure 1. Biplot by correspondence analysis (CA) showing relationships between phylotypes and virulence genes. Blue circles represent names of genes while red triangles represent names of phylotypes. CA: correspondence analysis, Dim: dimension.



.

Discussion

Escherichia coli (E. coli), a highly adaptable bacterium, primarily resides in the mammalian gut but can also thrive in diverse environments. This versatile microorganism is capable of causing a broad range of infections, from localized intestinal illnesses to systemic diseases. *E. coli*'s pathogenic potential extends beyond the gastrointestinal tract, enabling it to induce extraintestinal infections such as bloodstream infections, urinary tract infections, and meningitis. Its ability to adapt to various ecological niches and its diverse virulence factors contribute to its significance as both a commensal and pathogenic organism [15].

This bacterium has a great genomic plasticity which permits it to obtain and share genetic materials that improve its abilities to withstand the environmental stresses. In particular, *E. coli* has been demonstrated to harbor several mobile genetic elements (MGEs) that enhance their resistance [16]. The pathogenic *E. coli* has been proposed to originate from normal *E. coli* through the acquiring of chromosomal and/or extra-chromosomal genes that encode virulence [15]. *E. coli* with elevated levels of resistance and genetic variety was stated in various regions of Africa and worldwide [5, 17].

In order to more effectively address the epidemiological significance of pathogenic *E. coli, some* strains have previously been classified into phylotypes according to the presence of particular gene virulence factors in clinical *E. coli* isolates. The current investigation aligns with emerging research paradigms focused on unraveling E. coli's pathogenic mechanisms and formulating personalized treatment strategies against antibiotic-resistant variants.

In the light of the goal, this research studied 50 *E. coli* strains isolated from clinical specimens from hospitalized patients. Pediatric samples formed the primary source of isolates, with a gender distribution showing a notable female predominance of 60% compared to 40% male representation. *E. coli*'s prominence as a urinary tract pathogen is reflected in this study's sample composition, with urine specimens constituting the majority (76.0%) of isolates. This predominance of urinary samples aligns with findings from comparable research, where urine has consistently been the primary source of *E. coli* isolates [17]. In contrast to these results, **Čurová et al.** conducted research on 310 isolates, (41.9%) were from blood followed by urine (37.1%) and (21%) from patients suffering of skin and soft tissue diseases [18]. Different sample types provide insight into the varied clinical manifestations of *E. coli* infections beyond its common association with urinary tract infections.

Rising levels of resistance to many antibiotic types provide significant health concern by reducing effectiveness of therapy and restricting choices [19]. This study identified a significant prevalence of antimicrobial resistance among *E. coli* isolates, with 62% classified as MDR and 8% as XDR. These findings are consistent with recent researches, as exemplified by **El-Baz's** study, which

reported an even higher MDR rate of 92.67% and a lower XDR rate of 4% [15]. Similarly, **Masoud et al.** observed MDR in 73% of isolates in Egypt [20], while **Khairy et al.** detected MDR in 55% of diarrheal *E. coli* isolates from children in Upper Egypt [21]. In a comparable study, **Čurová et al.** identified MDR phenotypes in 53.2% of their isolates [18]. Collectively, these results highlight the growing challenge of antibiotic resistance in *E. coli* across various regions, underscoring the urgent need for improved antimicrobial stewardship and the development of novel treatment approaches.

Grouping *E. coli* into different types helps scientists understand how *E. coli* populations are organized and how these types relate to their ability to cause disease and resist antibiotics. While various methods exist for *E. coli* phylotyping, the PCR-based approach developed by **Clermont et al.** has become a standard in the field due to its efficiency and accessibility. This method assigns *E. coli* isolates to phylogroups (A, B1, B2, D, C, E, F, and clade 1) based on the presence or absence of four genetic alleles. This PCR-based method has become widely adopted due to its simplicity and efficiency in phylogenetic classification of *E. coli* strains [4].

In this study, 50 *E. coli* isolates from different sample types have been grouped phylogenetically by the presence or absence of the 4 genes by real time PCR using classification described earlier by **(Clermont)** [4]. Considering this classification method, 26(52%) were assigned to phylogroup B2 followed descendingly by group D 5(10%) and each of phylogroup C and E 4(8%), B1 3(6%), clade I 2(4%), and phylogroup F 1(2%). However, 5(10%) *E. coli* isolates were unknown. Group A was absent from our study. Unclassified strains (unknown) are expected to be resulted from extensive recombination events from two diverse groups or the extremely fluctuating gene content of *E. coli*, which is caused by the recurrent gaining and losing of genetic materials [4].

Numerous studies aimed to identify phylogroups in *E. coli* isolated from clinical specimens. Similar to our study **Byarugaba et al.** observed in 95 *E. coli* isolates that phylogroup B2 was the most prevalent among other phylotypes (30.523%) followed by B1 (12.63%), D (11.58%), F (10.53%), and C (5.26%), unlike our study group A represented (27.37%) [17]. **El-Baz et al.** inspected 332 *E. coli* isolates collected from different specimens and reported that all 8 phylogroups have been observed in clinical isolates of *E. coli*: 40 (26.67%) were B2, followed by 38 (25.33%) B1, 15 (4.5%) A, 12 (3.6%) D phylotype, whereas group C was 15 (4.5%), group E 17 (5.1%), group F 7 (2.1%), and 5 clade I (1.5%) isolates. 1 (0.31%) isolate did not belong to any group (unknown) [15]. **Olalekan et al.** reported that out of the 113 strains of *E. coli* isolate B2 was dominant group 65 (58.0%), clades I & II 16 (14.1%), groups E, F, A, C, and B1 were 12 (10.6.%), 6 (5.3%), 3 (2.7%), 2 (1.8%), and 1 (0.9) respectively. However, 8 (7.1%) of the isolates were unknown [22]. In Iraq **Mohammed et al.** reported that among 112 *E. coli* isolates group B2 exhibited the highest prevalence rate 39 (34.8%), followed by phylogenetic groups C 24 (21.4%), D 19 (16.9%), and A 12 (10.7%). The remaining groups, namely B1 6 (5.3%), E 5 (4.4%), F 3 (2.6%), and unknown 4 (3.5%) [23]. Contrary to previous data **Deku et al.** detected phylogenetic groups A in,(74.8%) of *E. coli* isolates followed by B1 (2.2%), B2 (14.8%), C (3.0%), clade I (1.5%), and D (3.7%) [24].

In our study, among the uroisolates phylogroup B2 were the most prevalent 20 (52%) of 38 *E. coli* isolates followed by phylogroup D 4(10.5%), each group C and B1 3(7.9%), group E 1(2.6%), clade I 2(5.3%) and group F 1(2.6%). Four isolates were unknown (10.5%). Consistently, several researches revealed that among UPEC the most prevalent phylogroup was B2 [5, 25-27]

In general, both B2 and D phylogroups are classified to be the most virulent strains associated with UTIs [5,28]. In this study, given that urine samples constituted the majority of specimens, phylogroup B2 emerged as the predominant phylogenetic group among the *E. coli* isolates. Following B2, phylogroup D was identified as the second most common lineage. This distribution pattern aligns with previous research findings, which have consistently associated phylogroups B2 and D with uropathogenic *E. coli* strains [5, 26, 27].

The variations in the distribution of detected phylotypes reported in numerous studies can be attributed to several factors, including differences in host health conditions, dietary habits, hygiene practices, geographic locations, climate, antimicrobial usage, and host genetic factors. Additionally, discrepancies may arise from variations in sampling methods, sample sizes, and research designs employed in these investigations. These elements collectively influence the observed prevalence and diversity of *E. coli* phylotypes across different studies [27, 29].

Escherichia coli (E. coli) pathogens have a various range of virulence traits which assist in establishing infection and development of diseases. Furthermore, *E. coli* pathotypes possess greater content of virulence factors compared to normal flora *E. coli* [30]. These factors facilitate the escape from host defenses, invasion and destruction of host cells, and eventually incite inflammations in the host [31]. Fimbrial and afimbrial adhesins participate in adhesion and colonizing of host epithelial cells [32]. It also produces many different types of toxins that play a crucial role in its pathogenesis [32].

PapG gene; virulence gene dedicated to adhesion; was the most virulence gene prevalent among the 50 *E. coli* isolates (78%). Varying prevalence of this gene was detected in other studies. Several studies reported prevalence of *papG* gene [18, 15, 33, 34].

Sfa (S fimbriae) enables *E. coli* to bind to the host kidney cells. Also, it mediates adhesion to brain microvascular endothelium, so, S fimbriae are usually detected in *E. coli* isolated from meningitis patients [32]. In this study *Sfa* was noticed in 25 (50%) among *E. coli* isolates collected from different specimen types, while other studies declared different distributions of this gene among their isolates [18, 35, 36].

Secreted autotransporter toxin (*Sat*) is a class-1serine protease autotransporter of *Enterobacteriaceae* (SPATE) proteases. It has potential for bacterial nutrients uptake, avoiding immunogenic responses and colonizing the infection sites. Sat has been firstly identified in a pyelonephritis associated UPEC strain, this toxin acts cytotoxically on bladder and kidney cells [37]. In the present study *sat* virulence gene was detected in 16(32%) of different *E. coli* samples. Similarly, **Turton**, revealed consistent prevalence of virulence gene *Sat* (34%) [38]. However, **Kim et al.** detected higher rate of *Sat* gene (46%) [36], on the other hand **Duan et al.** detected lower prevalence of *sat* gene (16.8%) [23].

Cnf1 (cytotoxic necrotizing factor 1) is a Rho GTPase activating proteins that stimulates invasion into host cells. It induces actin rearrangements and increases the phagocytic-like activity of human epithelial cells, causing digestion of apoptotic cells. It is commonly detected in ExPEC and it is encountered in UTIs, bacteremia, and meningitis [15,39].

In the present study the *Cnf1*gene was detected in 14 (28%) isolates of different *E. coli* sample types. **Kim et al**. detected similar proportion of *Cnf1* gene (28%) [36] while in **El-Baz et al.** study, *Cnf1* gene rate was higher (42%) [15]. On the othier hand, **Čurová et al., Gultekin et al.,** and **Turton** detected lower rates of *Cnf1* gene (18.1%, 14%, and 11.5%) respectively [18,33,38].

Alpha hemolysin is a pore-forming exotoxin secreted by various *E. coli* pathogenic strains promoting their infection [18]. Probable roles of α haemolysin in the pathogenesis of *E. coli* comprise exfoliation of epithelial cells to facilitate invasion of underlying tissues, alteration of host immunogenic responses and finally cell apoptosis [40, 41]. *hylA gene* was reported in this study in 11(22%) isolates from different sample types, similarily, **Čurová et al.** showed nearly consistent detection rate of *hylA gene* (19.7%) [18], while a previous study by **Abd El-Baky RM et al.** detected higher proportion of *hylA gene* (45.3%) [42] The variations in virulence genes observed in both our study and the previously mentioned researches may be attributed to differences in sample types, sample sizes, and geographical locations.

In our study group *hlyA* was only detected in phylogroup B2 and D, however, there were no statistical differences in the prevalence of the virulence genes detected among *E coli* phylogroups except for *HlyA* which was more detected in phylogroup B2 than non B2 phylogroups ($p= 0.01$). It was more detected in uroisolates than non-uroisolates ($p=0.05$).

Pattern of two virulence genes was detected in 17 isolates of 50 *E. coli* isolates in our study, with most common combinations (*papG* and *sat*) and (*papG* and *sfa*) detected in 5 isolates (each). This may be contributed to the fact that most of our isolates were from UTI. Similarly, **Turton** reported the presence of two virulence gene combination being *malX* and *sat* and *papC* and *sat* among 86 (14.5%) of *E. coli* isolates in their study [38].

Conclusion

Phylogroup B2 was the most prevalent (52%), followed by groups D, C, E, B1, clade 1, and F. We detected various virulence genes, with *papG* being the most common (78%). Notably, phylogroup B2 exhibited all investigated virulence genes and showed a stronger association with the *sat* and *hlyA* genes. These findings provide insights into the distribution of *E. coli* phylotypes and their

associated virulence factors in our clinical setting, which may inform future approaches to diagnosis and treatment of *E. coli* infections.

Competing interests

The authors have no conflicts of interest to declare.

Ethical consideration

The Ethics Committee at the Medical Research Institute at Alexandria University, (IORG#: IORG0008812) gave approval on this study. All participating patients gave informed written consent. It was conducted in accordance to Helsinki Declaration for studies on human subjects.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contributions

Shahinda Rezk: planned and supervised the experiments, discussed the results and writing of the final manuscript

Ahmed Jamal: carried out the experiments, performed the analysis, discussed the results and wrote the draft manuscript.

Abeer Ghazal: conceived the original idea, supervised the findings of this work, discussed the results, and critically revised the manuscript.

Aliaa Aboulela: supervised practical experiments.

References

1- **Abram K, Udaondo Z, Bleker C, Wanchai V, Wassenaar TM, Robeson MS, Ussery DW.** Mash-based analyses of Escherichia coli genomes reveal 14 distinct phylogroups. Communications Biology 2021; 4(1), 117. doi: 10.1038/s42003-020-01626-5

2- **El-Mahdy R, Mahmoud R, Shrief R.** Characterization of E. coli phylogroups causing catheter-associated urinary tract infection. Infection and Drug Resistance 2021; 3183-3193. doi: 10.2147/IDR.S325770

3- **Clermont O, Bonacorsi S, Bingen E.** Rapid and simple determination of the Escherichia coli phylogenetic group. Appl Environ Microbiol 2000; 66(10), 4555-4558. doi: 10.1128/aem.66.10.4555-4558.2000

4- **Clermont, O., Christenson, J. K., Denamur, E., & Gordon, D. M.** The Clermont Escherichia coli phylo-typing method revisited: improvement of specificity and detection of new phylo-groups. Environ Microbiol Rep 2013; 5(1), 58-65. doi: 10.1111/1758-2229.12019

5- **Baldiris-Avila R, Montes-Robledo A, Buelvas-Montes Y.** Phylogenetic Classification, Biofilm-Forming Capacity, Virulence Factors, and Antimicrobial Resistance in Uropathogenic Escherichia coli (UPEC). Curr Microbiol 2020; 77(11), 3361-3370. doi: 10.1007/s00284-020-02173-2

6- **Shah C, Baral R, Bartaula B, Shrestha LB.** Virulence factors of uropathogenic Escherichia coli (UPEC) and correlation with antimicrobial resistance. BMC microbiology 2019; 19, 1-6. doi: 10.1186/s12866-019-1587-3.

7- **Ahmed NA, Mahmood SS, Abbas AH.** A comparative study of some virulence factors and phylogenetic charactrization of Escherichia. coli isolates causing urinary tract infection and the commensal gut. The Iraqi Journal of Agricultural Science 2019; 50(3), 1193-1198. doi: 10.36103/ijas.v50i4.763

8- **Munkhdelger Y, Gunregjav N, Dorjpurev A, Juniichiro N, Sarantuya J.** Detection of virulence genes, phylogenetic group and antibiotic resistance of uropathogenic Escherichia coli in Mongolia. The Journal of Infection in Developing Countries 2017; 11(01), 51-57. doi: 10.3855/jidc.7903

9- **Lescat M, Clermont O, Woerther PL, Glodt J, Dion S, Skurnik D, Denamur E.** Commensal Escherichia coli strains in Guiana reveal a high genetic diversity with host-dependant population structure. Environ

Microbiol Rep 2013; 5(1), 49-57. doi: 10.1111/j.1758-2229.2012.00374.x

10- **Johnson JR, Stell AL.** Extended virulence genotypes of Escherichia coli strains from patients with urosepsis in relation to phylogeny and host compromise. The Journal of infectious diseases 2000; 181(1), 261-272. doi: 10.1086/315217

11- **Adamus-bialek W, Wojtasik A, Majchrzak M, Sosnowski M, Parniewski P.** (CGG) 4-Based PCR as a Novel Tool for Discrimination of Uropathogenic Escherichia coli Strains : Comparison with Enterobacterial Repetitive Intergenic Consensus-PCR, Journal of Clinical Microbiology 2009; 47 (12), 3937–3944. doi: 10.1128/JCM.01036-09

12- **Chapman TA, Wu XY, Barchia I, Bettelheim KA, Driesen S, Trott D, Chin JJC.** Comparison of virulence gene profiles of Escherichia coli strains isolated from healthy and diarrheic swine. Applied and Environmental Microbiology 2006; 72(7), 4782-4795. doi: 10.1128/AEM.02885-05

13- **Mladin C, Usein CR, Chifiriuc MC, Palade A, Slavu CL, Negut M, Damian M.** Genetic analysis of virulence and pathogenicity features of uropathogenic Escherichia coli isolated from patients with neurogenic bladder. Romanian Biotechnological Letters 2009; 14(6), 4906-4911. Corpus ID: 56397065.

14- **Restieri C, Garriss G, Locas MC, Dozois CM.** Autotransporter-encoding sequences are phylogenetically distributed among Escherichia coli clinical isolates and reference strains. Applied and environmental microbiology 2007; 73(5), 1553-1562. doi: 10.1128/AEM.01542-06

15- **El-Baz R, Said HS, Abdelmegeed ES, Barwa R.** Characterization of virulence determinants and phylogenetic background of multiple and extensively drug resistant Escherichia coli isolated from different clinical sources in Egypt. Appl Microbiol Biotechnol 2022; 106(3), 1279-1298. doi: 10.1007/s00253-021-11740-x

16- **Kocsis B, Gulyás D, Szabó D.** Emergence and Dissemination of Extraintestinal Pathogenic High-Risk International Clones of Escherichia coli. Life (Basel) 2022; 12(12). doi: 10.3390/life12122077

17- **Byarugaba DK, Erima B, Wokorach G, Alafi S, Kibuuka H, Mworozi E, Wabwire-Mangen F.** Resistome and virulome of high-risk pandemic clones of multidrug-resistant extra-intestinal pathogenic Escherichia coli (ExPEC) isolated from tertiary healthcare settings in Uganda. PLoS One 2023; 18(11), e0294424. doi: 10.1371/journal.pone.0294424

18- **Čurová K, Slebodníková R, Kmeťová M, Hrabovský V, Maruniak M, Liptáková E, Siegfried L.** Virulence, phylogenetic background and antimicrobial resistance in Escherichia coli associated with extraintestinal infections. J Infect Public Health 2020; 13(10), 1537-1543. doi: 10.1016/j.jiph.2020.06.032

19- **Monroy-Pérez E, Cerón AB, García Cortés LR, Alonso NN, Domínguez-Trejo P, Hernández-Jaimes T, Vaca S.** Virulence gene transcription, phylogroups, and antibiotic resistance of cervico-vaginal pathogenic E. coli in Mexico. PLoS One 2020; 15(6), e0234730. doi: 10.1371/journal.pone.0234730

20- **Masoud SM, Abd El-Baky RM, Aly SA, Ibrahem RA.** Co-Existence of Certain ESBLs, MBLs and Plasmid Mediated Quinolone Resistance Genes among MDR E. coli Isolated from Different Clinical Specimens in Egypt. Antibiotics 2021; 10(7), 835. doi: 10.3390/antibiotics10070835

21- **Khairy RMM, Fathy ZA, Mahrous DM, Mohamed ES, Abdelrahim SS.** Prevalence, phylogeny, and antimicrobial resistance of Escherichia coli pathotypes isolated from children less than 5 years old with community acquired- diarrhea in Upper Egypt. BMC Infect Dis 2020; 20(1), 908. doi: 10.1186/s12879-020-05664-6

22- **Olalekan A, Onwugamba F, Iwalokun B.** Phylogroup profiling and the genetic determinants of extended spectrum beta-lactamase producing Escherichia coli isolated from clinical samples in lagos, Nigeria. Ife Journal of Science 2020; 22(3), 159-166. doi: 10.4314/ijs.v22i3.13

23- **Mohammed HN, Al Marjani MF, Authman SH.** Molecular identification and characterization and phylogenetic study in Escherichia coli in Baghdad province. Reviews and Research in Medical Microbiology 2024; 10.1097. doi:10.1097/mrm 396

24- **Deku JG, Duedu KO, Kinanyok S, Kpene GE, Feglo PK.** Phylogenicity and virulence profiles of clinical Escherichia coli isolates in the Ho teaching hospital of Ghana. BioMed Research International 2022; 1-8. doi: 10.1155/2022/1347033

25- **Dahwash SM, Raheema RH, Albahadili MA, Maslat AH.** Distribution of phylogenetics and virulence genes of uropathogenic escherichia coli among urinary tract infection in pregnant women. Biochemical & Cellular Archives 2021; 21(1). doi: 10.1186/s12879-020-4844-z

26- **Wang MC, Fan YH, Zhang YZ, Bregente CJB, Lin WH, Chen CA, Kao CY.** Characterization of uropathogenic Escherichia coli phylogroups associated with antimicrobial resistance, virulence factor distribution, and virulence-related phenotypes. Infect Genet Evol 2023; 114, 105493. doi: 10.1016/j.meegid.2023.105493

27- **Dadi BR, Abebe T, Zhang L, Mihret A, Abebe W, Amogne W.** Distribution of virulence genes and phylogenetics of uropathogenic Escherichia coli among urinary tract infection patients in Addis Ababa, Ethiopia. BMC Infect Dis 2020; 20(1), 108. doi: 10.1186/s12879-020-4844-z

28- **Badger-Emeka LI, Kausar N, Estrella E, Angeles GB.** A Three-Year Look at the Phylogenetic Profile, Antimicrobial Resistance, and Associated Virulence Genes of Uropathogenic Escherichia coli. Pathogens 2022; 11(6). doi: 10.3390/pathogens11060631

29- **Aguirre-Sánchez JR, Valdez-Torres JB, Del Campo NC, Martínez-Urtaza J, Del Campo NC, Lee BG, Chaidez-Quiroz C.** Phylogenetic group and virulence profile classification in Escherichia coli from distinct isolation sources in Mexico. Infect Genet Evol 2022; 106, 105380. doi: 10.1016/j.meegid.2022.105380

30- **Sobhy NM, Yousef SGA, Aboubakr HA, Nisar M, Nagaraja KV, Mor SK, Goyal SM.** Virulence factors and antibiograms of Escherichia coli isolated from diarrheic calves of Egyptian cattle and water buffaloes. PLoS One 2020; 15(5), e0232890. doi: 10.1371/journal.pone.0232890

31- **El-Shaer S, Abdel-Rhman SH, Barwa R, Hassan R.** Virulence characteristics, serotyping and phylogenetic typing of clinical and environmental Escherichia coli isolates. Jundishapur Journal of Microbiology 2018; 11(12). doi: 10.5812/jjm.82835

32- **Ochoa SA, Cruz-Córdova A, Luna-Pineda VM, Reyes-Grajeda JP, Cázares-Domínguez V, Escalona G, Xicohtencatl-**

Cortes J. Multidrug- and Extensively Drug-Resistant Uropathogenic Escherichia coli Clinical Strains: Phylogenetic Groups Widely Associated with Integrons Maintain High Genetic Diversity. Front Microbiol 2016; 7, 2042. doi: 10.3389/fmicb.2016.02042

33- **Gultekin EO, Ulger ST, Delialioğlu N.** Distribution of pathogenicity island markers and virulence factors genes of extraintestinal pathogenic Escherichia coli isolates. Jundishapur Journal of Microbiology 2022; 15(4). doi: 10.5812/jjm-121044

34- **Bozcal E, Eldem V, Aydemir S, Skurnik M.** The relationship between phylogenetic classification, virulence and antibiotic resistance of extraintestinal pathogenic Escherichia coli in İzmir province, Turkey. PeerJ 2018; 6, e5470. doi: 10.7717/peerj.5470

35- **Tayh G, Nagarjuna D, Sallem RB, Verma V, Yahia HB, Gharsa H, Slama KB.** Identification of virulence factors among esbl-producing escherichia coli clinical isolates from gaza strip, palestine: Virulence genes among ESBL-producing E. coli. Journal of microbiology, biotechnology and food sciences 2021; 11(3), e2865-e2865. doi:10.15414/jmbfs.2865

36- **Kim B, Kim JH, Lee Y.** Virulence Factors Associated With Escherichia coli Bacteremia and Urinary Tract Infection. Ann Lab Med 2022; 42(2), 203-212. doi: 10.3343/alm.2022.42.2.203

37- **Vieira PCG, Espinoza-Culupú AO, Nepomuceno R, Alves MR, Lebrun I, Elias WP, Ruiz RC.** Secreted autotransporter toxin (Sat) induces cell damage during enteroaggregative Escherichia coli infection. PLoS One 2020; 15(2), e0228959. doi: 10.1371/journal.pone.0228959

38- **Turton JF.** Virulence factors among isolates of extraintestinal Esherichia coli (ExPEC) from hospitals in the United Kingdom. BioRxiv 2022; 31:2022-10. doi: 10.1101/2022.10.30.514400

39- **Chaoprasid P, Dersch P.** The cytotoxic necrotizing factors (CNFs)—A family of rho GTPase-activating bacterial exotoxins. Toxins. 2021; 15;13(12):901. doi: 10.3390/toxins13120901

40- **Caetano BDL, Domingos MdO, da Silva MA, da Silva JCA, Polatto JM, Montoni F, Vieira PCG.** In silico prediction and design of uropathogenic Escherichia coli alpha-hemolysin generate a soluble and hemolytic recombinant toxin. Microorganisms 2022; 10(1), 172. doi: 10.3390/microorganisms10010172.

41- **Kolenda R, Sidorczuk K, Noszka M, Aleksandrowicz A, Khan MM, Burdukiewicz M, Schierack P.** Genome placement of alpha-haemolysin cluster is associated with alpha-haemolysin sequence variation, adhesin and iron acquisition factor profile of Escherichia coli. Microb Genom 2021; 7(12). doi: 10.1099/mgen.0.000743

42- **Abd El-Baky RM, Ibrahim RA, Mohamed DS, Ahmed EF, Hashem ZS.** Prevalence of Virulence Genes and Their Association with Antimicrobial Resistance Among Pathogenic E. coli Isolated from Egyptian Patients with Different Clinical Infections. Infect Drug Resist 2020; 13, 1221-1236. doi: 10.2147/idr.s24107