

RESEARCH ARTICLE

Open Access



Virulence genes and phylogenetic groups of uropathogenic *Escherichia coli* isolates from patients with urinary tract infection and uninfected control subjects: a case-control study

Seyedeh Elham Rezatofghi^{1*} , Mahsa Mirzarazi¹ and Mansour Salehi²

Abstract

Background: Urinary Tract Infection (UTI) is one of the most common bacterial infectious diseases which causes considerable morbidity and costly health problems. Uropathogenic *Escherichia coli* (UPEC), the most common pathogen causing UTI, is a highly heterogeneous group of extraintestinal pathogenic *E. coli* (ExPEC) which may carry a variety of virulence factors and belonging to different phylogenetic backgrounds. The current study aimed to investigate the frequency and association between various virulence factors (VFs) and phylogenetic groups of UPEC and commensal isolates.

Methods: UPEC and commensal *E. coli* strains isolated from UTI and feces of healthy humans were compared for the presence of VFs and phylogenetic groups. Association between virulence genes was investigated and cluster analysis was employed.

Results: According to the results, among a 30 virulence markers tested, the pathogenicity-associated island (PAI), *papAH*, *papEF*, *fimH*, *fyuA*, and *traT* genes prevalence were statistically significant in UPEC isolates. A strong association was found between the B2 and D phylogenetic groups and clinical isolates of UPEC; while, commensal isolates were mostly associated with phylogenetic group A. The aggregated VFs scores were more than twice higher in the UPEC isolates in comparison with the commensal isolates. Interestingly, the B2 group in both UPEC and commensal isolates had the highest VF scores. A strong positive association was found between several virulence genes. The clustering results demonstrated that UPEC or commensal *E. coli* isolates were highly heterogeneous due to different composition of their virulence gene pool and pathogenicity islands.

(Continued on next page)

* Correspondence: e.tofighi@yahoo.com; e.tofighi@scu.ac.ir

¹Department of biology, Faculty of Science, Shahid Chamran University of Ahvaz, Ahvaz 6135743135, Iran

Full list of author information is available at the end of the article



© The Author(s). 2021 **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

(Continued from previous page)

Conclusion: Genetic structure and VFs of UPEC strains vary from region to region; therefore, to control the UTI, the epidemiological aspects and characterization of the UPEC isolates need to be investigated in different regions. Since UPEC isolates are generally originate from the commensal strains, it may be feasible to reduce the UTI burden by interfering the intestinal colonization, particularly in the highly pathogenic clonal lineages such as B2.

Keywords: Urinary tract infection, Uropathogenic *Escherichia coli*, Virulence factor, Phylogenetic group, Extraintestinal Pathogenic *E. coli*

Background

Escherichia coli inhabit the large intestine of healthy humans and other warm-blooded animals, but in some instances it can produce a wide range of extraintestinal infections. This bacterium can easily acquire virulence factors (VFs) and mobile genetic elements from related bacteria that leads to different pathogenicity [1]. Extraintestinal pathogenic *E. coli* (ExPEC) isolates are highly complex and have a variety of VFs and may belong to different phylogenetic lineages. These strains cause complicated urinary tract infections (UTIs), bacteremia, and sepsis [2].

UTI is one of the most common infectious diseases accounting for approximately 40% of all nosocomial infections and 10–20% of hospital-acquired infections [3]. UTI is associated with considerable morbidity and costly health problems. They cause a variety of clinical signs from asymptomatic bacteriuria to pyelonephritis, cystitis, and septic shock with multi-organ systems failure [4]. The most common pathogen causing UTI is a heterogeneous group of ExPEC, named uropathogenic *E. coli* (UPEC) [5]. UPEC strains cause 75–95% of uncomplicated and 40–50% of complicated UTIs [3]. Based on the available literature, UPEC strains evolve from non-pathogenic strains by acquiring new VFs through horizontal gene transfer (HGT) [6]. The *E. coli* genome consists of a main core genome and a mobile gene pool that determine pathotype or ecotype specific traits [3]. Various VFs have been attributed to UPEC pathogenesis; however, there is no general agreement regarding the definitive discriminatory virulence factors within this pathotype. UPEC isolates need VFs for colonizing or invading host cells, escaping or disrupting hosts' immune systems, damaging host tissues, and/or stimulating inflammatory responses. Among variety of VFs, some are generally accepted to be more associated with UPEC [7]. *fim* operon, *pap* operon, and *sfa* genes encode type I fimbriae, P fimbriae and S fimbriae respectively [8, 9]. These structural VFs are the main attachment factors associated with colonization of organism to host cells [7]. Apart from adhesins, some virulence genes encode toxins such as hemolysin (*hly* gene), cytotoxic-necrotizing-factor (*cnf1* gene), and siderophores (*fyuA* gene) that are mainly involved in intracellular survival,

iron-acquisition, escape from immune system, inflammatory response, and host tissue damage [8–10]. UPEC isolates may carry pathogenicity-associated islands (PAIs) which carry sets of different virulence associated genes [8]. Some of these virulence genes are also found in the commensal isolates and are not specific to pathogenic isolates. However, most available studies have only investigated the prevalence of virulence genes in UPEC associated isolates.

Due to variability in the gene content and the possibility of HGT among different *E. coli* isolates, it is vital to understand the genetic basis of differences between commensal and UPEC isolates, to be able to prevent ExPEC and UPEC infections more effectively. This information can be acquired through case-control epidemiological studies [4, 11]. Therefore, in the present study, we investigated the frequency and relationship between different VFs and phylogenetic groups of UPEC and commensal isolates.

Methods

Sample collection and analysis

In a case-control study, 702 midstream specimens of urine were collected from patients with an age range from 1 month to 93 years (mean age: 37.07 ± 22.2 years). Sampling was performed randomly. The samples were obtained from patients referred by physicians to medical centers in Isfahan city, Iran, to diagnose urine infection. Urine samples were cultured on a MacConkey (Merck, Germany), Eosin Methylene Blue (Merck, Germany) and blood agar (Merck, Germany) plates. Positive urine cultures with at least 10⁵ cfu/mL and an evaluated white blood cell counts (≥10⁴ leukocyte/mL of urine) were considered UTI positive. UPEC isolates were confirmed using standard biochemical tests including EMB, Methyl Red - Voges-Proskauer (MR-VP), Triple Sugar Iron agar (TSI), and Simmons Citrate agar testing. To confirm the isolates to be *E. coli* and also to evaluate of the quality of extracted DNA, the presence of *uspA* gene was investigated with PCR [11]. This gene encodes the highly specific *E. coli* universal stress protein A. Out of 702 urine specimens, 138 samples were positive for UTI caused by UPEC.

In addition, 30 commensal *E. coli* isolates collected from feces of healthy humans were considered as control. The control samples were collected from volunteers who had no symptoms of disease and not taken antibiotics in the last three months. The study was approved by Ethics Committee of Shahid Chamran University of Ahvaz (Ethics statement No 63/21/8/90). Patients and volunteers were asked to read, accept and sign an informed consent form before any information was collected. Written consent was obtained from the parents for sampling children.

Virulence genes (VGs)

DNA of UPEC and commensal isolates were extracted using the boiling method. A group of 30 VGs was analyzed. Five separate multiplex PCR were used for the presence of VFs including PAI, *papAH*, *fimH*, *kpsMT III*, *papEF*, *ibeA*, *fyuA*, *bmaE*, *sfa/focDE*, *iutA*, *papG*, *papG* allele I, I^a, II, and III, *kpsMT K1*, *hlyA*, *rfaE*, *kpsMT II*, *papC*, *gafD*, *cvaC*, *cdtB*, *focG*, *traT*, *afa/draBC*, *cnf1*, *sfaS*, and *kpsMT K5*. PCR conditions and primers were performed as described by Johnson and Stell [11]. Each positive gene was confirmed with separate PCR. The sum of VFs for each isolate was calculated and regarded as a VF score.

Phylogenetic analysis

Phylogenetic groups of isolates were investigated using the method of Clermont et al. based on the presence of two genes of *chuA* and *yjaA*, as well as, a DNA fragment TSPE4.C2 [12]. According to the amplification results, the *E. coli* isolates were classified into one of the major phylogenetic groups: A, B1, B2, or D.

Cluster analysis

Similarity relationships based on composite genomic profiles of the isolates were used to create a dendrogram according to the unweighted pair group method with averaging (UPGMA) supported by the Numerical Taxonomy and Multivariate Analysis System (NTSYS) package version 2.02pc.

Statistical methods

The sample size was calculated by the Epi Info, a program developed by the Centers for Disease Control and Prevention available via the link: <https://www.cdc.gov/epiinfo/index.html>. According to a pilot study, Odds Ratio, percent of control exposed, power, alpha, and ratio of controls to cases were considered 5, 12, 80%, 0.05, and 0.3, respectively. Therefore, at least 94 case samples and 29 control samples were required. The Goodman and Kruskal tau coefficient was used to measure the strength of the associations for the cross-tabulation of virulence genes of the *E. coli* isolates from both UPEC

and commensal strains. The association between different groups and the presence of the investigated genes was assessed using the Pearson Chi-square test or Fisher's exact test with SPSS 21.0 software. Moreover, the scores were compared using the Mann–Whitney *U*-test. Results were considered statistically significant at $p < 0.05$.

Results

Virulence characteristics

Compared to commensal isolates, the UPEC isolates had a significantly higher prevalence of several VFs including *papAH* and *papEF* (P fimbria), *fimH* (type 1 fimbriae), *fyuA* (siderophore), *traT*, and PAI ($P < 0.05$) (Table 1). Although other genes were also more prevalent in the UPEC isolates than commensal, but their prevalence was not statistically significant ($P > 0.05$) and were considered normal distribution. (Table 1).

Phylogenetic groups of isolates

The phylogenetic distribution analysis showed a strong association between B2 and D phylogenetic groups and clinical isolates of UPEC; while, commensal isolates were more associated with A phylogenetic group (Table 2). Regarding the strong associations of UTI status with both phylogenetic background and some virulence factors, in the next step, the relationship between phylogenetic groups and virulence factors was investigated.

Phylogenetic distribution of VFs

The distribution of VFs in each phylogenetic group was compared to other phylogenetic groups combined. The results indicated that *papC*, *fimH*, *fyuA*, *iutA*, *kpsMT II*, *kpsMT K5*, and PAI genes were positively associated with group B2, and *papG* allele I^a, PAI and *fimH* were also positively associated with group D; while, the *fimH*, *snf1*, and *fyuA* genes were negatively associated with group A, and PAI with group B1, respectively. Other virulence genes had a normal distribution among phylogenetic groups, although group B2 accounted more for most of the virulence genes (Table 3).

Aggregate VF scores among UPEC and commensal isolates

Aggregate VF scores were calculated by summing the number of the virulence genes present in each commensal or UPEC isolates belonging to each phylogenetic group. Results appeared that aggregate VF scores were more than twice as high among the UPEC isolates (mean: 4.93; range: 2–12) than among the commensal isolates (mean: 2.2; range: 0–8). Different phylogenetic groups also had various aggregate VF scores, so that group B2 had the highest VF score, group D exhibited an intermediate VF score, and scores of the groups A

Table 1 Distribution of virulence-associated traits of uropathogenic and commensal *Escherichia coli* isolates

Gene	UPEC isolates (n=138)		Commensal isolates (n=30)		P value
	N (%)	N (%)	N (%)	N (%)	
<i>papAH</i>	27 (19.5)	0 (0)	0 (0)	0 (0)	█
<i>papC</i>	30 (21.7)	5 (16.6)	5 (16.6)	5 (16.6)	...
<i>papEF</i>	29 (21)	1 (3.3)	1 (3.3)	1 (3.3)	█
<i>papG</i> I,II,III	32 (23.2)	3 (10)	3 (10)	3 (10)	...
allele I	2 (1.4)	0 (0)	0 (0)	0 (0)	...
allele I ^a	5 (3.6)	1 (3.3)	1 (3.3)	1 (3.3)	...
allele I+ I ^a	7 (5)	1 (3.3)	1 (3.3)	1 (3.3)	...
allele II	9 (6.5)	0 (0)	0 (0)	0 (0)	...
allele III	16 (11.5)	2 (6.6)	2 (6.6)	2 (6.6)	...
<i>sfa/focDE</i>	20 (14.4)	3 (10)	3 (10)	3 (10)	...
<i>sfaS</i>	17 (12.3)	1 (3.3)	1 (3.3)	1 (3.3)	...
<i>focG</i>	1 (0.7)	0 (0)	0 (0)	0 (0)	...
<i>afa/draBC</i>	22 (15.9)	3 (10)	3 (10)	3 (10)	...
<i>bmaE</i>	6 (4.3)	0 (0)	0 (0)	0 (0)	...
<i>gafD</i>	0 (0)	0 (0)	0 (0)	0 (0)	...
<i>nfaE</i>	2 (1.4)	0 (0)	0 (0)	0 (0)	...
<i>fimH</i>	86 (62.3)	7 (23.3)	7 (23.3)	7 (23.3)	█
<i>hlyA</i>	2 (1.4)	0 (0)	0 (0)	0 (0)	...
<i>cnfI</i>	17 (12.3)	1 (3.3)	1 (3.3)	1 (3.3)	...
<i>cdtB</i>	0 (0)	0 (0)	0 (0)	0 (0)	...
<i>fyuA</i>	83 (60.1)	10 (33.3)	10 (33.3)	10 (33.3)	█
<i>iutA</i>	81 (58.6)	12 (40)	12 (40)	12 (40)	...
<i>kpsMT II</i>	45 (32.6)	9 (30)	9 (30)	9 (30)	...
<i>kpsMT III</i>	20 (14.4)	1 (3.44)	1 (3.44)	1 (3.44)	...
<i>kpsMT K1</i>	11 (7.9)	5 (16.6)	5 (16.6)	5 (16.6)	...
<i>kpsMT K5</i>	18 (13)	1 (3.3)	1 (3.3)	1 (3.3)	...
<i>rfc</i>	1 (0.7)	1 (3.3)	1 (3.3)	1 (3.3)	...
<i>ibeA</i>	8 (5.7)	1 (3.3)	1 (3.3)	1 (3.3)	...
<i>cvaC</i>	6 (4.3)	0 (0)	0 (0)	0 (0)	...
<i>traT</i>	58 (42)	3 (10)	3 (10)	3 (10)	█
PAI	45 (32.6)	0 (0)	0 (0)	0 (0)	█

P values (by χ^2 test or Fisher's exact test) are shown only if $P < 0.05$. The values significantly higher than among the other groups are indicated as follows: █ $P < 0.001$, █ $P < 0.01$, █ $P < 0.05$

and B1 were significantly lower than other phylogenetic groups. Also, VF scores of all phylogenetic groups of commensal isolates were lower than phylogenetic groups of UPEC isolates (Table 4). Therefore, commensal isolates generally had lower VFs than UPEC isolates.

Table 2 Phylogenetic distribution of uropathogenic and commensal *Escherichia coli* isolates

Phylogenetic groups (no of isolates)	Prevalence of phylogenetic groups No (%)			P value
	UPEC isolates	Commensal isolates		
	N (%)	N (%)		
A (26)	17 (12.3)	9 (30)		0.02
B1 (16)	16 (11.7)	0 (0)		...
B2 (85)	76 (55)	9 (30)		0.01
D (41)	29 (21)	12 (40)		0.03
Total	138 (100)	30 (100)		

P values (by χ^2 test or Fisher's exact test) are shown only if $P < 0.05$. UPEC: Uropathogenic *Escherichia coli*

Prevalence of VFs among group B2 of UPEC and commensal isolates

So far, the results showed that the VFs were associated with UPEC isolates and phylogenetic group B2. To explore which origin source of phylogenetic group B2 have more virulence genes, the prevalence of VFs among group B2 of UPEC and commensal isolates were compared. The analysis revealed no large difference between VFs of UPEC isolates over commensal isolates within phylogenetic group B2. Out of 30 VFs only two individual VFs of *traT* and PAI were significantly more

Table 3 Phylogenetic distribution of virulence-associated traits among uropathogenic and commensal *Escherichia coli* isolates

Gene (no)	Prevalence of trait in phylogenetic groups			
	No (%)			
	Group A (n=26)	Group B1 (n=16)	Group B2 (n=85)	Group D (n=41)
<i>papAH</i> (27)	1 (4)	2 (12.5)	18 (21.2)	6 (14.6)
<i>papC</i> (35)	2 (8)	1 (6.25)	27 (31.8)	5 (12.2)
<i>papEF</i> (30)	5 (19)	2 (12.5)	18 (21.2)	5 (12.5)
<i>papG</i> I, II, III (32)	3 (11.5)	1 (6.25)	17 (20)	6 (14.6)
allele I (2)	0 (0)	0 (0)	1 (1.2)	1 (2.4)
allele I ^a (6)	1 (4)	0 (0)	1 (1.2)	4 (9.7)
allele II (9)	0 (0)	1 (6.25)	6 (7)	2 (4.9)
allele III (18)	3 (11.5)	0 (0)	11 (13)	4 (9.7)
<i>sfa/focDE</i> (23)	5 (19)	0 (0)	15 (17.6)	3 (7.3)
<i>sfaS</i> (18)	1 (4)	2 (12.5)	13 (15.3)	2 (4.9)
<i>focG</i> (1)	0 (0)	0 (0)	1 (1.2)	0 (0)
<i>afa/draBC</i> (25)	3 (11.5)	0 (0)	17 (20)	5 (12.2)
<i>bmaE</i> (6)	1 (4)	0 (0)	3 (3.5)	2 (4.9)
<i>gafD</i> (0)	0 (0)	0 (0)	0 (0)	0 (0)
<i>nfaE</i> (2)	0 (0)	0 (0)	2 (2.4)	0 (0)
<i>fimH</i> (93)	8 (30.7)	11 (69)	58 (68.2)	16 (39)
<i>hlyA</i> (2)	0 (0)	0 (0)	2 (2.4)	0 (0)
<i>cnf1</i> (18)	2 (8)	1 (6.25)	12 (14.1)	3 (7.3)
<i>cdtB</i> (0)	0 (0)	0 (0)	0 (0)	0 (0)
<i>fyuA</i> (93)	8 (30.7)	7 (44)	56 (66)	22 (53.6)
<i>iutA</i> (93)	12 (46)	9 (56.2)	54 (63.5)	18 (44)
<i>kpsMT</i> II (54)	5 (19)	2 (12.5)	37 (43.5)	10 (24.4)
<i>kpsMT</i> III (21)	3 (11.5)	2 (12.5)	13 (15.3)	4 (9.7)
<i>kpsMT</i> K1 (16)	1 (4)	1 (6.25)	10 (11.8)	4 (9.7)
<i>kpsMT</i> K5 (19)	1 (4)	0 (0)	16 (18.8)	2 (4.9)
<i>rfc</i> (2)	2 (8)	0 (0)	0 (0)	0 (0)
<i>ibeA</i> (9)	0 (0)	0 (0)	7 (8.2)	2 (4.9)
<i>cvaC</i> (6)	0 (0)	1 (6.25)	5 (5.9)	0 (0)
<i>traT</i> (61)	8 (30.7)	2 (12.5)	37 (43.5)	14 (34.1)
PAI (45)	3 (11.5)	0 (0)	36 (42.4)	6 (14.6)

P values were calculated by χ^2 test or Fisher's exact test for comparison of trait of isolates in each group with all other combined groups. The values significantly higher than among the other groups are indicated as follows: ■ $P < 0.001$, ■ $P < 0.01$, ■ $P < 0.05$. The values significantly lower than among the other groups are indicated as follows: ■ $P < 0.01$

prevalent among group B2 of UPEC isolates than group B2 of commensal isolates (Table 5).

Association between virulence genes

The association between the various virulence genes of the *E. coli* isolates is shown as a heat map (Fig. 1). A strong positive association was found between the genes of (*fyuA* and *iutA*), (*kpsMT* II and *papC*), (PAI and *fimH*) and (*sfaS* and both genes of *afa/draBC* and *kpsMT* K5). A moderate to weak association was found between the remaining genes.

Cluster analysis

Cluster analysis using the UPGMA was employed to agglomerate individual isolates into larger clusters and form a dendrogram describing the relationships

among UPEC and commensal isolates. Based on the presence or absence of the virulence genes among individual isolates and visualizing genetic associations, one major cluster and several minor clusters were distinguished (Fig. 2). Nearly 80% of commensal isolates were located in the major cluster with the smallest amount of virulence genes. Most of these commensal isolates were correspond to the phylogenetic group A or D. However, some commensal isolates were placed in the other clusters next to the pathogenic isolates with high virulence genes content. These isolates were corresponded to phylogenetic group B2, except for one case. About half of the UPEC isolates were placed in the minor clusters and were related to phylogenetic group B2. These isolates exhibited higher virulence genes content.

Table 4 Virulence factor (VF) scores for uropathogenic and commensal *Escherichia coli* isolates, by phylogenetic group

Phylogenetic group	Mean VF scores N (Mean)			P value UPEC vs. Commensal isolates
	All isolates (n=168)	UPEC isolates (n=138)	Commensal isolates (n=30)	
Total	(4.45)	(4.93)	(2.2)	
Group A	26 (2.9)	17 (4.05)	9 (0.66)	
Group B1	16 (2.9)	16 (2.9)	0 (0)	
Group B2	85 (5.76)	76 (5.95)	9 (4.2)	
Group D	41 (3.3)	29 (3.9)	12 (1.8)	

UPEC Uropathogenic *Escherichia coli*. The values significantly higher than among the other groups are indicated as follows: ■ $P < 0.001$, ■ $P < 0.01$, ■ $P < 0.05$. The values significantly lower than among the other groups are indicated as follows: ■ $P < 0.001$, ■ $P < 0.01$, ■ $P < 0.05$

Discussion

To have a better knowledge on the pathogenesis of the UPEC, it is necessary to identify virulence markers of various strains that cause UTIs. VFs, as the potential clinical predictors, help clinicians to manage patients and anticipate the evolution of infection in the host body [6, 13]. However, except for the genetic characteristics of the virulence strains, host factors play an important role in the incidence and outcomes of the infection [14, 15].

Fimbriae and adhesins are frequently reported as VFs in the UPEC isolates [11, 16, 17]. Fimbriae have an important role in establishing and progression of UTI. P-fimbrial adhesins with binding capacity to renal cell receptors cause the specific signaling pathways that trigger mucosal inflammation and tissue damage [18, 19]. Although, in the current study all investigated P fimbriae genes were more prevalent in the UPEC isolates, than the commensal isolates, but the prevalence of *papAH* and *papEF* genes was statistically significant. Correlation association analysis between virulence genes revealed a positive association between *papAH* and *papEF* genes. The *pap* genes are usually chromosomal [2]; therefore, these two genes were probably transmitted together through the chromosome.

Among adhesin genes, *fimH*, the gene that encodes type 1 fimbriae, was common among the UPEC isolates. This VF is attributed to cystitis-associated UPEC strains and helps to adhere, invade, and form the intracellular bacterial communities (IBCs) [14, 16, 20]. The prevalence of *fimH* was found to be 62.3% that was lower than some studies in Iran [16, 21], but it's consistent with some previous reports from Mexico, Tunisia, and Iran

[6, 22, 23]. In the current study, the *fimH* gene had the most prevalence among virulence genes, which may indicate its critical role in producing UTI. Therefore, FimH could be considered as a potential vaccine candidate. Besides, some studies are investigating this issue. For instance, it has been previously shown that antibodies against FimH can prevent the colonization of UPEC in urinary tract system [24, 25]. The *fimH* gene showed a positive correlation with PAI, indicating the genetic linkages between them.

Although, both siderophore genes of *fyuA* and *iutA* were prevalent in more than 50% of UPEC isolates; the frequency of yersiniabactin (*fyuA* gene) was statistically significant in the UTI producing isolates more than commensal isolates. Moreno et al. also reported a strong association of *fyuA* with urine versus fecal sources [26]. Drawn heat map revealed a strong positive association between two *fyuA* and *iutA* siderophore genes, which indicates the importance of iron absorption systems in pathogenic isolates.

The *traT* was another gene that was statistically significant in UPEC isolates. It expresses a transfer protein that inhibits the classical pathway of complement activation [2]. This gene is a part of *tra* operon of the F-like conjugative plasmids and leads to serum survival [2]. In the other study from Iran, consistent with our results, *traT*, *fyuA*, and *fimH* genes were the most frequently detected VFs in UPEC isolates [27].

ExPEC strains, particularly UPEC isolates, usually contain multiple PAIs with a distinctive combination of VFs; therefore, some isolates may have multiple copies of a VF [13]. In this way, PAIs can play an important role in increasing the pathogenicity of bacteria.

Table 5 Distribution of virulence-associated traits among uropathogenic and commensal *Escherichia coli* isolates within phylogenetic groups B2

Gene	Prevalence of trait N (%)		P value
	UPEC isolates (N=76)	Commensal isolates (N=9)	
<i>papAH</i>	18 (23.7)	0 (0)	...
<i>papC</i>	23 (30.3)	4 (44.4)	...
<i>papEF</i>	18 (23.7)	0 (0)	...
<i>papG</i> I, II, III	17 (22.4)	0 (0)	...
allele I	1 (1.3)	0 (0)	...
allele I ^a	1 (1.3)	0 (0)	...
allele II	6 (7.9)	0 (0)	...
allele III	10 (13.1)	1 (11.1)	...
<i>sfa/focDE</i>	13 (17.1)	2 (22.2)	...
<i>sfaS</i>	13 (17.1)	0 (0)	...
<i>focG</i>	1 (1.3)	0 (0)	...
<i>afa/draBC</i>	16 (21)	1 (11.1)	...
<i>bmaE</i>	3 (4)	0 (0)	...
<i>gafD</i>	0 (0)	0 (0)	...
<i>nfaE</i>	2 (2.6)	0 (0)	...
<i>fimH</i>	52 (68.4)	6 (66.6)	...
<i>hlyA</i>	2 (2.6)	0 (0)	...
<i>cnf1</i>	12 (15.8)	0 (0)	...
<i>cdtB</i>	0 (0)	0 (0)	...
<i>fyuA</i>	49 (64.5)	7 (77.7)	...
<i>iutA</i>	46 (60.5)	8 (88.8)	...
<i>kpsMT II</i>	31 (40.8)	6 (66.6)	...
<i>kpsMT III</i>	12 (15.8)	1 (11.1)	...
<i>kpsMT K1</i>	8 (10.5)	2 (22.2)	...
<i>kpsMT K5</i>	15 (19.7)	1 (11.1)	...
<i>rfc</i>	0 (0)	0 (0)	...
<i>ibeA</i>	7 (9.2)	0 (0)	...
<i>cvaC</i>	5 (6.6)	0 (0)	...
<i>traT</i>	37 (48.7)	0 (0)	0.004
PAI	36 (47.4)	0 (0)	0.008

P values (by χ^2 test or Fisher's exact test) are shown only if $P < 0.05$. UPEC Uropathogenic *Escherichia coli*

Although, the frequency of other genes was not statistically significant between two groups of UPEC and commensal isolates; but, in the current study, gene expression levels were not investigated. The rate of genes transcription, expression, or the copy number of each gene might be different in these two groups of isolates. Also, to detect virulence genes the PCR method was used, however, due to mutation, some virulence genes may not be accurately detected. Therefore, positive PCR results indicated the presence of genes; but, a negative result does not essentially equivalent to the absence of the corresponding genes, although this phenomenon is scarce [6].

Phylogenetic analysis revealed that the B2 and D (to a lesser extent) were dominant phylogroups of UPEC isolates. The prevalence of VFs was higher among group B2 isolates taken from the urine of patients with acute cystitis than fecal isolates of healthy people [28, 29], which is in line with previous studies conducted in Ethiopia [240], Denmark [30], Pakistan [31], South Korea [18], Poland [32], and Mexico [22]. Therefore, in the producing-UTI isolates, most of the VFs were more frequent than commensal isolates. On the other hand, the UPEC isolates mostly belonged to group B2. Thus, the association between VFs and phylogenetic groups was investigated.

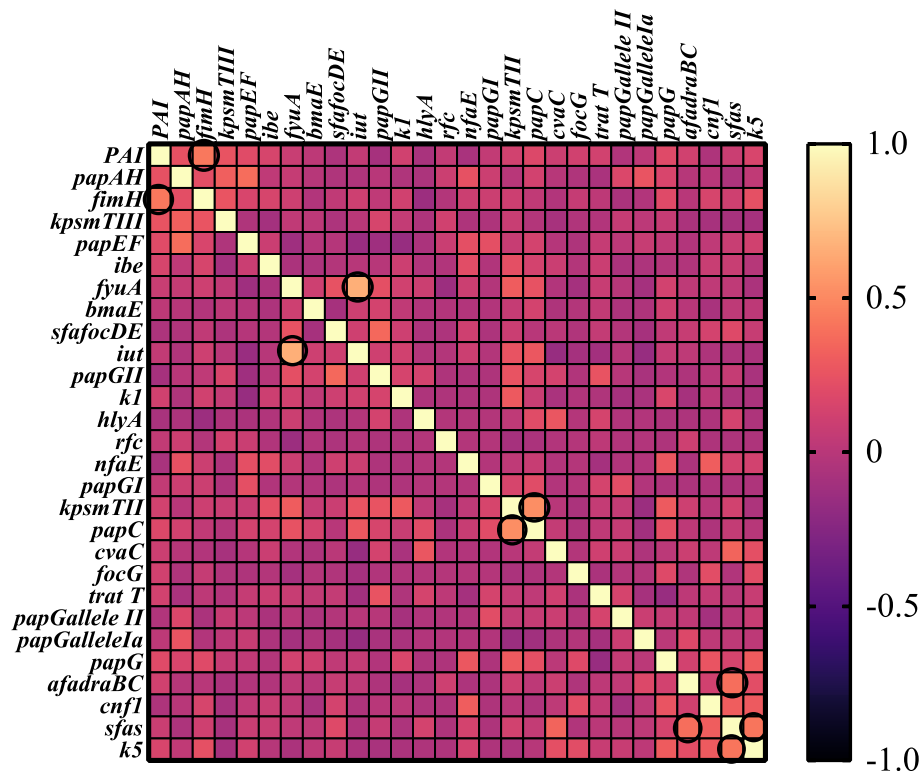


Fig. 1 Heat map generated according to statistical association between virulence genes of the *Escherichia coli* isolates derived from uropathogenic and commensal *E. coli* isolates. The strong associations between genes are indicated in the circle. No values were introduced in the case of undetected genes

Distribution of VFs in the phylogenetic groups indicated the presence of some genes, including *papC*, *fyuA*, *iutA*, *kpsMT* K5, and *kpsMTII* were positively associated with group B2, *fimH* and PAI with both B2 and D groups, and *papG* allele I^a was associated with group D. Such associations are also reported in other studies [14, 33–35]. Using such evidence scientists can investigate these phylogenetic groups (B2 and D) for VFs, which as yet is undefined [13].

VF scores of UPEC isolates were higher than commensal isolates. Group B2 in both commensal and UPEC isolates had the highest aggregative VF scores, followed by group D (intermediate), and groups A and B1 (the lowest). Based on the obtained results, it can be concluded that isolates related to the B2 phylogenetic group had a pathogenic potential, regardless of their origin. In this way, these isolates can be a possible candidate for developing a vaccine or drugs.

Co-selection or direct genetic linkage of VFs leads to the common simultaneous appearance of certain VFs [11, 13, 36]. In the current study, the Goodman and Kruskal tau coefficient method and heat map were used to find the strength of the associations of virulence genes. If different studies investigate and identify the associations between genes; it would be feasible to have a better

understanding of the role of these genes in how pathogenesis occurs.

UPEC isolates are diverse due to the presence of different virulence genes carried by plasmids, transposons, PAIs, and bacteriophages. These genetic elements may carry antibiotic resistance genes in addition to virulence genes. Plasmids belonging to the IncF incompatibility group were found to encode both VFs and antibiotic resistance genes [37]. However, some researchers believe that multidrug resistant isolates are significantly less virulent than susceptible isolates since antibiotic resistance and virulence do not usually co-evolve simultaneously [38]. Although some studies have found the positive association between virulence traits including iron scavenger receptors with antibiotic resistance [39, 40], there are many contradictions in this regard. For example, in some studies the presence of *hly* gene was associated with sensitivity to fluoroquinolones [41–43] and in other study resistance to fluoroquinolones [8]. Thus, future comprehensive studies are necessary to elucidate the relationship between VFs and antibiotic resistance and the evolutionary direction of bacteria.

The clustering results demonstrated that the UPEC or commensal *E. coli* isolates are genotypically highly heterogeneous. These different patterns of clusters are

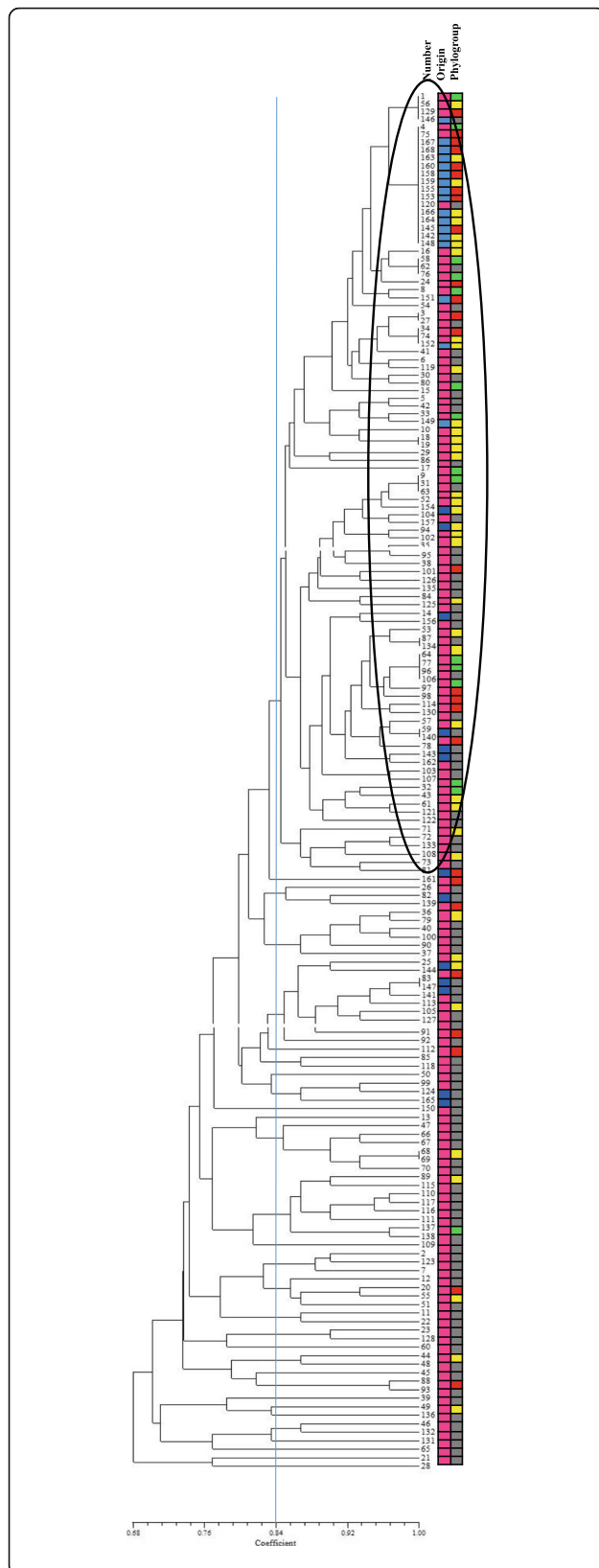


Fig. 2 Similarity relationships based on composite genomic profiles and phylogenetic groups of *Escherichia coli* isolates. Commensal and uropathogenic *E. coli* isolate are shown as blue and purple colors respectively. A, B1, B2, and D groups are shown as red, green, gray, and yellow colors respectively. Major cluster is marked on an oval

probably due to chromosomal or plasmid location of virulence genes and vertical (within-lineage) or horizontal (among-lineage) gene transfer phenomena [13]. Additionally, these data emphasize the findings relay on the fact that many UPEC isolates originate from commensal strains without considerable virulence genes content as previously reported in some women, and UTI may cause by a high prevalence of relatively low virulence *E. coli* strains in the fecal reservoir [4]. Also, commensal strains can be potentially pathogenic, when colonizing extra-intestinal tissues. Therefore, by reducing the intestinal colonization of UPEC strains and dealing with *E. coli* virulence mechanisms, the UTIs may be prevented [26]. The epidemiological aspects of the UPEC in different regions need further investigation to find the spread of different isolates and to understand the dissemination of these pathogens to hosts [17].

Conclusion

In conclusion, in this study, most UPEC isolates were related to group B2, followed by group D. UPEC isolates carried sets of important fimbriae adhesion associated virulence genes especially *papAH*, *papEF*, and *fimH*. The siderophore genes were another noticeable VF in UPEC isolates and a strong positive association were found between *fyuA* and *iutA* siderophore genes. The results of the present study confirmed that the UPEC or commensal *E. coli* isolates are highly heterogeneous and have complex genetic backgrounds. Therefore, it seems that the epidemiological aspects and characterization of the UPEC isolates needs to be investigated in different regions in different time frames.

Abbreviations

VF: Virulence factor; ExPEC: Extraintestinal pathogenic *E. coli*; UTI: Urinary tract infection; PAI: Pathogenicity associated island; UPEC: Uropathogenic *E. coli*; HGT: Horizontal gene transfer; WBC: White blood cell; VG: Virulence gene; UPGMA: Unweighted pair group method with averaging; NTSYS: Numerical Taxonomy and Multivariate Analysis System; IBC: Intracellular bacterial communities

Acknowledgements

The authors are very thankful to Shahid Chamran University of Ahvaz for the facilities to accomplish the present research project. Mahsa Mirzarazi is currently a Ph. D student and his affiliation is follow: Department of Biotechnology; School of Advanced technologies in medicine; Shahid Beheshti University of medical science; Tehran; Iran.

Authors' contributions

All authors contributed to the design of the experiment. SER designed and supervised the research study. MR carried out the experiments. MS participated in the design of the study and data analysis. All authors read and approved the final manuscript.

Funding

No funding.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations**Ethics approval and consent to participate**

The study was approved by Ethics Committee of Shahid Chamran University of Ahvaz (Ethics statement No 63/21/8/90). Volunteers will be asked to read, accept and sign an informed consent form before any information was collected. Written consent was obtained from the parents for sampling children.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹Department of biology, Faculty of Science, Shahid Chamran University of Ahvaz, Ahvaz 6135743135, Iran. ²Department of Genetics and Molecular Biology, Medical School, Isfahan University of Medical Sciences, Isfahan, Iran.

Received: 19 June 2020 Accepted: 5 April 2021

Published online: 17 April 2021

References

- Harwalkar A, Gupta S, Rao A, Srinivasa H. Lower prevalence of *hlyD*, *papC* and *cnf-1* genes in ciprofloxacin-resistant uropathogenic *Escherichia coli* than their susceptible counterparts isolated from southern India. *J Infect Public Heal*. 2014;7(5):413–9. <https://doi.org/10.1016/j.jiph.2014.04.002>.
- Sarowska J, Futoma-Koloch B, Jama-Kmieciak A, Frej-Madrzak M, Ksiazczyk M, Bugla-Ploskonska G, et al. Virulence factors, prevalence and potential transmission of extraintestinal pathogenic *Escherichia coli* isolated from different sources: recent reports. *Gut Pathog*. 2019;11(1):10. <https://doi.org/10.1186/s13099-019-0290-0>.
- Tan CW, Chlebicki MP. Urinary tract infections in adults. *Singap Med J*. 2016;57(9):485–90. <https://doi.org/10.11622/smedj.2016153>.
- Agarwal J, Srivastava S, Singh M. Pathogenomics of uropathogenic *Escherichia coli*. *Indian J Med Microbiol*. 2012;30(2):141–9. <https://doi.org/10.4103/0255-0857.96657>.
- Chakraborty A, Adhikari P, Shenoy S, Saralaya V. Molecular characterisation of uropathogenic *Escherichia coli* isolates at a tertiary care hospital in South India. *Indian J Med Microbiol*. 2017;35(2):305–10. https://doi.org/10.4103/ijmm.IJMM_14_291.
- Tarchouna M, Ferjani A, Ben-Selma W, Boukadida J. Distribution of uropathogenic virulence genes in *Escherichia coli* isolated from patients with urinary tract infection. *Int J Infect Dis*. 2013;17(6):e450–3. <https://doi.org/10.1016/j.ijid.2013.01.025>.
- Katongole P, Nalubega F, Florence NC, Asiimwe B, Andia I. Biofilm formation, antimicrobial susceptibility and virulence genes of Uropathogenic *Escherichia coli* isolated from clinical isolates in Uganda. *BMC Infect Dis*. 2020;20(1):453. <https://doi.org/10.1186/s12879-020-05186-1>.
- Yazdanpour Z, Tadjrobehkar O, Shahkhal M. Significant association between genes encoding virulence factors with antibiotic resistance and phylogenetic groups in community acquired uropathogenic *Escherichia coli* isolates. *BMC Microbiol*. 2020;20(1):241. <https://doi.org/10.1186/s12866-020-01933-1>.
- Terlizzi ME, Gribaudo G, Maffei ME. UroPathogenic *Escherichia coli* (UPEC) infections: virulence factors, bladder responses, antibiotic, and non-antibiotic antimicrobial strategies. *Front Microbiol*. 2017;8:1566. <https://doi.org/10.3389/fmicb.2017.01566>.
- Khairy RM, Mohamed ES, Abdel Ghany HM, Abdelrahim SS. Phylogenetic classification and virulence genes profiles of uropathogenic *E. coli* and diarrhegenic *E. coli* strains isolated from community acquired infections. *Plos One*. 2019;14(9):e0222441.
- Johnson JR, Stell AL. Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. *J Infect Dis*. 2000;181(1):261–72. <https://doi.org/10.1086/315217>.
- Clermont O, Bonacorsi S, Bingen E. Rapid and simple determination of the *Escherichia coli* phylogenetic group. *Appl Environ Microbiol*. 2000;66(10):4555–8. <https://doi.org/10.1128/AEM.66.10.4555-4558.2000>.
- Johnson JR, Russo TA. Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. *Int J Med Microbiol*. 2005;295(6–7):383–404. <https://doi.org/10.1016/j.ijmm.2005.07.005>.
- Ali I, Rifaqat Z, Ahmed I, Tariq F, Graham SE, Salzman E, et al. Phylogeny, sequence-typing and virulence profile of uropathogenic *Escherichia coli* (UPEC) strains from Pakistan. *BMC Infect Dis*. 2019;19(1):620. <https://doi.org/10.1186/s12879-019-4258-y>.
- Takahashi A, Kanamaru S, Kurazono H, Kunishima Y, Tsukamoto T, Ogawa O, et al. *Escherichia coli* isolates associated with uncomplicated and complicated cystitis and asymptomatic bacteriuria possess similar phylogenies, virulence genes, and O-serogroup profiles. *J Clin Microbiol*. 2006;44(12):4589–92. <https://doi.org/10.1128/JCM.02070-06>.
- Malekzadegan Y, Khashei R, Sedigh Ebrahim-Saraie H, Jahanabadi Z. Distribution of virulence genes and their association with antimicrobial resistance among uropathogenic *Escherichia coli* isolates from Iranian patients. *BMC Infect Dis*. 2018;18(1):572. <https://doi.org/10.1186/s12879-018-3467-0>.
- Kim DH, Subhadra B, Kang HY, Woo K, Kim J, Son YJ, et al. Virulence properties of uropathogenic *Escherichia coli* isolated from children with urinary tract infection in Korea. *Genes Genom*. 2018;40(6):625–34. <https://doi.org/10.1007/s13258-018-0664-6>.
- Lee JH, Subhadra B, Son YJ, Kim DH, Park HS, Kim JM, et al. Phylogenetic group distributions, virulence factors and antimicrobial resistance properties of uropathogenic *Escherichia coli* strains isolated from patients with urinary tract infections in South Korea. *Lett Appl Microbiol*. 2016;62(1):84–90. <https://doi.org/10.1111/lam.12517>.
- Mabbett AN, Ulett GC, Watts RE, Tree JJ, Totsika M, Ong CL, et al. Virulence properties of asymptomatic bacteriuria *Escherichia coli*. *Int J Med Microbiol*. 2009;299(1):53–63. <https://doi.org/10.1016/j.ijmm.2008.06.003>.
- Sheldon IM, Rycroft AN, Dogan B, Craven M, Bromfield JJ, Chandler A, et al. Specific strains of *Escherichia coli* are pathogenic for the endometrium of cattle and cause pelvic inflammatory disease in cattle and mice. *Plos One*. 2010;5(2):e9192. <https://doi.org/10.1371/journal.pone.0009192>.
- Momtaz H, Karimian A, Madani M, Safarpour Dehkordi F, Ranjbar R, Sarshar M, et al. Uropathogenic *Escherichia coli* in Iran: serogroup distributions, virulence factors and antimicrobial resistance properties. *Ann Clin Microbiol Antimicrob*. 2013;12(1):8. <https://doi.org/10.1186/1476-0711-12-8>.
- Paniagua-Contreras GL, Monroy-Pérez E, Rodríguez-Moctezuma JR, Domínguez-Trejo P, Vaca-Paniagua F, Vaca S. Virulence factors, antibiotic resistance phenotypes and O-serogroups of *Escherichia coli* strains isolated from community-acquired urinary tract infection patients in Mexico. *J Microbiol Immunol Infect*. 2017;50(4):478–85. <https://doi.org/10.1016/j.jmii.2015.08.005>.
- Tabasi M, Asadi Karam MR, Habibi M, Yekaninejad MS, Bouzari S. Phenotypic assays to determine virulence factors of Uropathogenic *Escherichia coli* (UPEC) isolates and their correlation with antibiotic resistance pattern. *Osong Public Health Res Perspect*. 2015;6(4):261–8. <https://doi.org/10.1016/j.phrp.2015.08.002>.
- 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. <https://doi.org/10.1186/s12879-020-4844-z>.
- Flores-Mireles AL, Walker JN, Caparon M, Hultgren SJ. Urinary tract infections: epidemiology, mechanisms of infection and treatment options. *Nat Rev Microbiol*. 2015;13(5):269–84. <https://doi.org/10.1038/nrmicro3432>.
- Moreno E, Andreu A, Pérez T, Sabaté M, Johnson JR, Prats G. Relationship between *Escherichia coli* strains causing urinary tract infection in women and the dominant faecal flora of the same hosts. *Epidemiol Infect*. 2006;134(5):1015–23. <https://doi.org/10.1017/S0950268806005917>.
- Haghighatpanah M, Mojtahedi A. Characterization of antibiotic resistance and virulence factors of *Escherichia coli* strains isolated from Iranian inpatients with urinary tract infections. *Infect Drug Resist*. 2019;12:2747–54. <https://doi.org/10.2147/IDR.S219696>.
- Sannes MR, Kuskowski MA, Owens K, Gajewski A, Johnson JR. Virulence factor profiles and phylogenetic background of *Escherichia coli* isolates from

- veterans with bacteremia and uninfected control subjects. *J Infect Dis.* 2004; 190(12):2121–8. <https://doi.org/10.1086/425984>.
29. Zhang L, Foxman B, Marrs C. Both urinary and rectal *Escherichia coli* isolates are dominated by strains of phylogenetic group B2. *J Clin Microbiol.* 2002; 40(11):3951–5. <https://doi.org/10.1128/JCM.40.11.3951-3955.2002>.
 30. Ejrnæs K, Stegger M, Reisner A, Ferry S, Monsen T, Holm SE, et al. Characteristics of *Escherichia coli* causing persistence or relapse of urinary tract infections: phylogenetic groups, virulence factors and biofilm formation. *Virulence.* 2011;2(6):528–37. <https://doi.org/10.4161/viru.2.6.18189>.
 31. Bashir S, Haque A, Sarwar Y, Ali A, Anwar MI. Virulence profile of different phylogenetic groups of locally isolated community acquired uropathogenic *E coli* from Faisalabad region of Pakistan. *Ann Clin Microb Antimicrob.* 2012; 11:23.
 32. Kot B, Wicha J, Gruzewska A, Piechota M, Wolska K, Obrębska M. Virulence factors, biofilm-forming ability, and antimicrobial resistance of urinary *Escherichia coli* strains isolated from hospitalized patients. *Turk J Med Sci.* 2016;46(6):1908–14. <https://doi.org/10.3906/sag-1508-105>.
 33. Er DK, Dundar D, Uzuner H, Osmani A. Relationship between phylogenetic groups, antibiotic resistance and patient characteristics in terms of adhesin genes in cystitis and pyelonephritis isolates of *Escherichia coli*. *Microb Pathogen.* 2015;89:188–94. <https://doi.org/10.1016/j.micpath.2015.10.014>.
 34. Hancock V, Ferrières L, Klemm P. The ferric yersiniabactin uptake receptor FyuA is required for efficient biofilm formation by urinary tract infectious *Escherichia coli* in human urine. *Microbiology (Reading, Engl.)*. 2008;154(Pt 1):167–75.
 35. Rowe MC, Withers HL, Swift S. Uropathogenic *Escherichia coli* forms biofilm aggregates under iron restriction that disperse upon the supply of iron. *FEMS Microbiol Lett.* 2010;307(1):102–9. <https://doi.org/10.1111/j.1574-6968.2010.01968.x>.
 36. Kanamaru S, Kurazono H, Ishitoya S, Terai A, Habuchi T, Nakano M, et al. Distribution and genetic association of putative uropathogenic virulence factors *iroN*, *iha*, *kpsMT*, *ompT* and *usp* in *Escherichia coli* isolated from urinary tract infections in Japan. *J Urol.* 2003;170(6 Pt 1):2490–3. <https://doi.org/10.1097/01.ju.0000094185.48467.dc>.
 37. Rijavec M, Erjavec MS, Avguštin JA, Reissbrodt R, Fruth A, Krizan-Hergouth V, et al. High prevalence of multidrug resistance and random distribution of mobile genetic elements among uropathogenic *Escherichia coli* (UPEC) of the four major phylogenetic groups. *Curr Microbiol.* 2006;53(2):158–62. <https://doi.org/10.1007/s00284-005-0501-4>.
 38. Johnson JR, Kuskowski MA, Gajewski A, Sahn DF, Karlowsky JA. Virulence characteristics and phylogenetic background of multidrug-resistant and antimicrobial-susceptible clinical isolates of *Escherichia coli* from across the United States, 2000–2001. *J Infect Dis.* 2004;190(10):1739–44. <https://doi.org/10.1086/425018>.
 39. Asadi Karam MR, Habibi M, Bouzari S. Relationships between virulence factors and antimicrobial resistance among *Escherichia coli* isolated from urinary tract infections and commensal isolates in Tehran, Iran. *Osong Public Health Res Perspect.* 2018;9(5):217–24. <https://doi.org/10.24171/j.phrp.2018.9.5.02>.
 40. Koczura R, Mokracka J, Barczak A, Krysiak N, Kaznowski A. Association between the presence of class 1 integrons, virulence genes, and phylogenetic groups of *Escherichia coli* isolates from river water. *Microb Ecol.* 2013;65(1):84–90. <https://doi.org/10.1007/s00248-012-0101-3>.
 41. Hozzari A, Behzadi P, Kerishchi Khiabani P, Sholeh M, Sabokroo N. Clinical cases, drug resistance, and virulence genes profiling in Uropathogenic *Escherichia coli*. *J Appl Genet.* 2020;61(2):265–73. <https://doi.org/10.1007/s13353-020-00542-y>.
 42. López-Banda DA, Carrillo-Casas EM, Leyva-Leyva M O-HG, Manjarrez-Hernández AH, Arroyo-Escalante S, et al. Identification of virulence factors genes in *Escherichia coli* isolates from women with urinary tract infection in Mexico. *Biomed Res Int.* 2014;2014:959206.
 43. Yun KW, Kim HY, Park HK, Kim W LIS. Virulence factors of uropathogenic *Escherichia coli* of urinary tract infections and asymptomatic bacteriuria in children. *J Microbiol Immunol Infect.* 2014;47(6):455–61. <https://doi.org/10.1016/j.jmii.2013.07.010>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

