Medical Journal of Babylon
Vol. 14- No. 3 : 470 - 477 , 2017
http://www.medicaljb.com
ISSN 2312-6760©2017 University of Babylon

Original Research Article

Prevalence and Molecular Characterization of *Fim* H Gene in *Escherichia Coli* Isolates Recovered From Patients With Utis

Narmin Saeed Merza
College of Science, University of Duhok, Duhok, IRAQ

E-mail:narmin.merza@gmail.com

Accepted 6 September, 2017

Abstract

In this study, the prevalence of *fim* H gene was studied among 105 *E. coli* isolates obtained from urine samples of patients attended Azadi hospital in Duhok City. The intended gene was detected in94.3% of the isolates. Triplex PCR assay was applied and according to which the studied isolates were assigned into four groups namely A, B1, B2, and D groups which constituted 20.95 %, 3.8 %, 54.28 %, 20.95 %, respectively. Ten randomly selected isolates were subjected to SNPs *fim*H analysis with 3 reference strains of *E. coli*. The results revealed that 44 SNPs observed at 42 polymorphic sites accounting for 5.59%. All mutations were of substitutions and 29.5 % of mutations were transversions while transition type mutations constituted of 70.5 %. Ten SNPs accounting for 22.7 % of mutations gave rise to amino-acid changes (sense mutation) while the rest 34 (77.3 %) resulted in silent mutations. Moreover, twelve SNPs were singletons and among them five were with amino acid replacements. Amino acid replacements due to SNPs accounted for 1.27% of whole sequenced fragment of *fim*H.It can be concluded that there is no relationship inferred between the isolates of *E.coli* when the two phylotyping techniques are compared but the results of both can serve the purpose of genotypic characterization of uropathogenic *E. coli*.

Key Words: *E. coli*, *Fim*H, Single nucleotide polymorphisms, Phylogeny, Sequencing.

الخلاصة

تم دراسة انتشار جين *FimH* بين ١٠٥ عزلة من بكتريا الاشيريشيا كولاي والمعزولة من عينات الادرار في مستشفى ئازادي التعليمي في مدينة دهوك. وقد تبين ان ٩٤,٣ % من هذه العينات تمتلك هذا الجين. وباستخدام تقنية التسلسل التضاعفي المتبلمر الثلاثية (Triplex PCR) تم تصنيف العزلاتالى اربع مجاميع رئيسة وهي كل من A, B1, B2, D بنسب 20.95%, 54.28 %, 3.8 %, 20.95 %على التوالي. وتم تطبيق تقنية تباين النيكلوتيد الواحد لجين FimH (SNPs analysis) على عشر عزلات مختارة عشوائيا بالاضافة الى ثلاث سلالات قياسية لبكتريا الاشيريشيا كولاي، وقد تبين ان هناك ٤٤ نيكليوتيد متباين ظهرت على ٤٢ موقع مشكلة نسبة ٥,٥٩%. وقد كانت جميع الطفرات من نوع الاستبدالي (Substitutions) وكان٢٩,٥% منهذه الطفرات من نوع البديلة(Transversion) اما الطفرات من النوع الانتقالي (Transition) فكانت بنسبة ٧٠,٥%. وقد أظهرت النتائج انعشرة (٢٢,٧%) من الطفرات الناتجة منتقنية النيوكلوتيد الواحد المتابية SNPs قد ادت الى تغيير الحامض الاميني (الطفرات الحسية)و٣٤ من الطفرات لم تؤدي الى تغيير الحامض الاميني (الطفرات الصامتة). بالاضافة الى ان١٢ من الطفرات النيوكلوتيد الواحد المتابية هي فردية (Singletons) وان خمس من هذا الطفرات الفردية قد ادت الى تغيير الحامض الاميني.وشكلت نسبة استبدال الاحماض الامينية بسببتباين نيكلوتيد المفرد 1.27%من التسلسل الكامل للجزء المدروس من الجين. ويمكنأن نستنتج بأنه لم نجد هناك علاقة في تصنيف (Phylotyping) العزلات لبكتريا الايشريشيا كولاي عندما قورنت نتائج التقنيتين لكن يمكن استخدام التقنيتين معا لإيجاد الخصائص الجينية لهذه العزلات.

Introduction

The genotypic characterization of pathogens has become a major aim for searching the genetic relatedness between clinical pathogenic strains and standard strains and for epidemiological investigation and public health strategies [1] Clermont *et al.* [2] proposed a triplex PCR assay for phylotyping of *E. coli* isolates into one of the major phylo-groups, A, B1, B2 or D depending on presence or absence of three markers namely *huA, yjaA* genes and a DNA fragment TspE4.C2 which is later characterized as a putative lipase esterase gene [3] Single-nucleotide polymorphism (SNP) analysis of *fimH* is another cheap, screening tool that has been widely used for epidemiological typing and genotypic analyses of Uropathogeneic *Escherichia coli* (UPEC) [4]. It has been found that during the course of bacterial evolution, genomic changes take place as a result of point mutations or horizontal transfer of genetic information resulting in strains divergence [5]. SNP analysis takes advantage of these changes at multiple loci to differentiate isolates [6]. *Fim*H is a unique adhesin found at the top of type 1 fimbriae that specifies mannose-sensitive binding of bacteria to eukaryotic cells [7]. The usefulness of *fim*H single-nucleotide polymorphisms (SNPs) for *E. coli* typing has been proven by many studies [4,8,9].

The aim of the present study is to apply SNP *fim*H analysis for characterization and genotyping of *E. coli* isolates assigned to different phylogroups, recovered from urine samples.

Materials and Methods

Sample collection and Phylogenetic grouping

One hundred five*E. coli*isolates have been recovered from urine samples of individuals attending Azadi Hospital in Dohuk City. Phenotypic identification and characterization have been done by other work [10]. Lyophilized master mix (Bioneer/South Korea) was reconstituted by adding 12µl grade distal water, 1µl of each primer including forward and reverse in a concentration of 20 pmol/µl of *chuA, yjaA* and TspE4.C2 genes, and 2µl of genomic DNA (50-100ng/µl) [2]. Regarding the screening of *fim*H gene, the amplification reaction consisted of 16µl PCR-grade water along with 1 µl of each forward and reverse primers (10 pmol/µl) of *fim*H (*Fim*H-F 5'-CGAGTTATTACCCTGTTTGCTG-3') and *Fim*H-R 5'-ACGCCAATAATCGATTGCAC-3') and 2 µl DNA template (50-100ng/µl) [9]. The PCR products were electro-phoresed by running through 2% (w/v) agarose in TBE Buffer [11].

Sequencing *fim*H gene

For sequencing reaction of *fimH* gene, 50ul of PCR product was prepared as follows;Illustra GFX PCR DNA and Gel Band Purification Kit(manufactured in UK) were used for amplicons purification.The Big Dye Terminator V3.1 cycle sequencing Kit was utilized for sequencing the intended DNA fragments; sequencing reaction consisted of 1µl of BigDye, 3µl of 5X Buffer reaction, 1µl of primers (3.2 pmol/µl) either forward or reverse and 10 µl of distilled water. Sequencing program created as follows: 1 cycle at 96^{o}C/1minute, 25 consecutive cycles of 96 ^{o}C /10 seconds, 50 ^{o}C / 5 seconds and 60 ^{o}C /4 minutes, holding at 4^{o}C. Capillary electrophoresis sequencer **(**ABI 3130 DNA sequencer, Singapore) was used to determine sequences of the reaction products. The resulting raw sequences were visualized by Chromas3.5V software to create counting of each individual gene using forward and reverse sequences. The sequences of each segment were trimmed to a specified length of 786bp correspondent to the region used to determine the target gene. All DNA sequences deposited into

NCBI/Blastn to identify corresponding region of the gene of interest.

SNP analysis was performed by multiple sequence alignments of the corresponding *fim*H gene for 10 *E. coli* isolates with3 reference stains including; CFT073, *E. coli* UTI89 and *E coli* MM_1856/ Iran with accession number AE014075.1, CP000243.1 and JX847135.1 respectively using BioEdit (Version 6). *E. coli* CFT037 was used as standard to compare with other *E. coli* stains. MEGA4was used for phylogenetic analysis and construction of neighboring trees.

Results

Using PCR assay, *fimH* gene was detected in 99 isolates accounting for 94.3% of the studied *E. coli* isolates. These isolates were successfully amplified 878 bp fragment of *fimH* gene while 6 isolates (5.7%) lacked it (Figure 1).

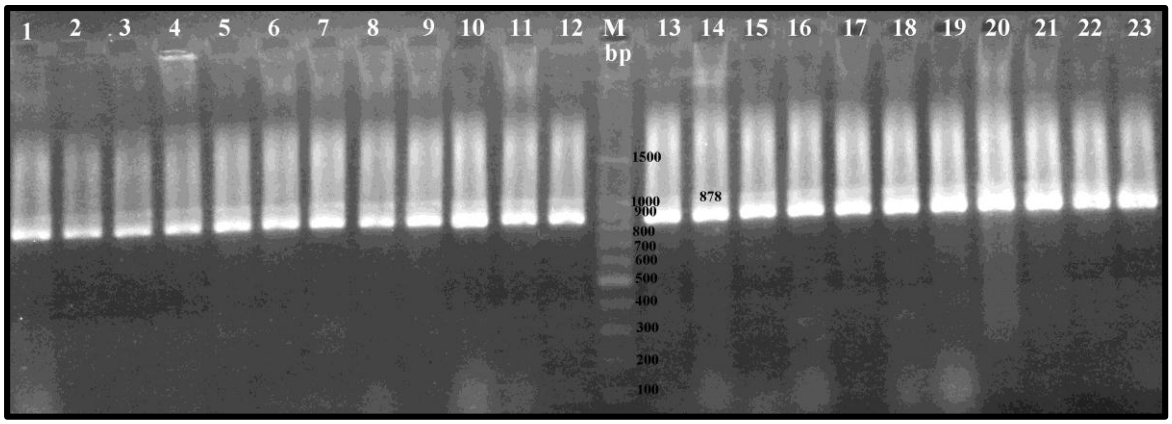

Figure 1: PCR amplification of *fimH*gene. Amplicon molecular weight was 878bp. Agarose gel (1.2%)was used for electrophoresing PCR-products. Power was5V/Cm, for 2hours. Lane M consisted of 100 bp molecular weight marker.

Among the 105 *E. coli* analyzed isolates, it has been found that57of these isolates belonged to phylogroupB2 accounting for 54.28%, 22 isolates belonged for both phylo groups A and D representing 20.95% of the isolates, while only 4 isolates (3.8%)assigned to phylogroup B1 (Table 1).

Table 1: Presence of *fim*H gene among Phylogenic group of *E. coli* strains isolated from Urine

	No.	PhylogroupA	PhylogroupB2	PhylogroupB1	PhylogroupD
E. coli	105	22	57	4	22
fimH (+)	99(94.28%)	20	55	3	21

*fim*H SNPs analysis involved 10 *fim*H gene positive clinical *E. coli* isolates and 3 references sequence strains (*E. coli* CFT073, *E. coli* UTI89 and E coli MM_1856/ Iran with accession number AE014075.1, CP000243.1, and JX 847135.1, respectively). Comparisons of the sequences of 786-bp *fim*H DNA fragments of the ten clinical isolates of *E.coli*, *E. coli* UTI89, and *E coli* MM_1856/ Iran were done in relation to that of CFT073 *E coli fim*H.

Sequencing revealed the presence of 44SNPs observed at 42 polymorphic sites accounting for 5.59% of the whole sequenced fragment of *fimH* gene (Table 2). All the detected SNPs mutations were of a substitution type; transversion observed in 13 out of 44 (29.5%)mutations while 31 out of 44 (70.5%) showed transition type SNPs. Ten SNPs (22.7 %)

resulted in amino acid changes known as sense mutation while 34 SNPs (77.3 %)out of 44 were silent substitutions. Nine amino acid substitutions out of a total of ten were caused by transitions. Twelve SNPs were identified as singletons meaning they were detected in one *fim*H type only, among them 5 were with amino acid replacements.

Table 2: DNA Polymorphism in *fimH* genes of the studied *E. coli* isolates

SNP	No.
Total	44 (5.59%)
No. of mutations/site	44/42
Point substitutions (Transitions)	31/44
Point substitutions (Transversions)	13/44
Amino acid replacement	10/44
Silent substitutions	34/44
Singletons	12 (5 with a.a. replacements)

Amino acid replacements and silent SNPs were compared among randomly selected *E coli* isolates and reference *E. coli* strains (*E. coli*_CFT073) (Table 3). It was revealed that all isolates shared the same amino acid substitutions C497T (Ala166Val); clinical isolates of *E. coli* in the order of Eco2 through Eco8 shared the same amino acids substitutions at positions G218A(Ser73Asn), A242G (Asn81Ser), and C497T (Ala166Val). Unique amino acids substitutions observed in ECO1-B2 at C230T (Thr77ILeu), ECO2 at C326T (Ala109Val) and EC08-B2 at C89T (Ala30Val), C313T (Pro105Ser) (Table 3).

Table 3: Polymorphism (s) due to alteration of single nucleotide with amino-acid replacements and the phylo-grouping of *E. coli* isolates

Strain	Phylogroup	*FimH S*NPs
CFT073	B2	None
ECO1-Duhok	B2	**C230T(Thr77ILeu)**
ECO2- Duhok	B2	G218A(Ser73Asn), A242G(Asn81Ser), **C326T(Ala109Val),** C497T (Ala166Val)
ECO3A-Duhok	A	G218A(Ser73Asn), A242G(Asn81Ser), C365T (Ala122Val), C497T (Ala166Val)
ECO4-Duhok	A	G218A(Ser73Asn), A242G(Asn81Ser), C497T (Ala166Val)
ECO5-Duhok	D	G218A(Ser73Asn), A242G(Asn81Ser), C497T (Ala166Val), G506A (Arg169His)
ECO6-D/Duhok	D	G218A(Ser73Asn), A242G (Asn81Ser), C497T (Ala166Val), G506A (Arg169His)
ECO7-B2/Duhok	B2	G218A(Ser73Asn), A242G(Asn81Ser), C497T (Ala166Val), G506A (Arg169His)
ECO8-B2/Duhok	B2	**C89T(Ala30Val),** G218A(Ser73Asn), A242G(Asn81Ser), **C313T(Pro105Ser),** C497T (Ala166Val)
ECO9-B2/Duhok	B2	C365T (Ala122Val), C497T (Ala166Val)
ECO10-B2/Duhok	B2	C497T (Ala166Val)
E_coli_MM_1856/_Iran		G218A(Ser73Asn), A242G(Asn81Ser), C497T (Ala166Val)
Ecoli_UTI89	B2	T193G(Ser65Ala), C497T (Ala166Val)

Dendogram analysis of the studied isolates showed that the isolates fallen into two main groups, the first group divided into three subgroups; (ECO2/B2; ECO4 and *E. coli* MMI-1856/Iran); (ECO5-D, ECO6-D and ECO7-B2/Duhok) and (ECO3-A,

ECO8-B2); the second group; ECO1, ECO9 and ECO10; Eco9 was found to be the closest strain to CFT073. The construction of a phylogram helped the determination of a particular *fim*H variant (Figure-1).

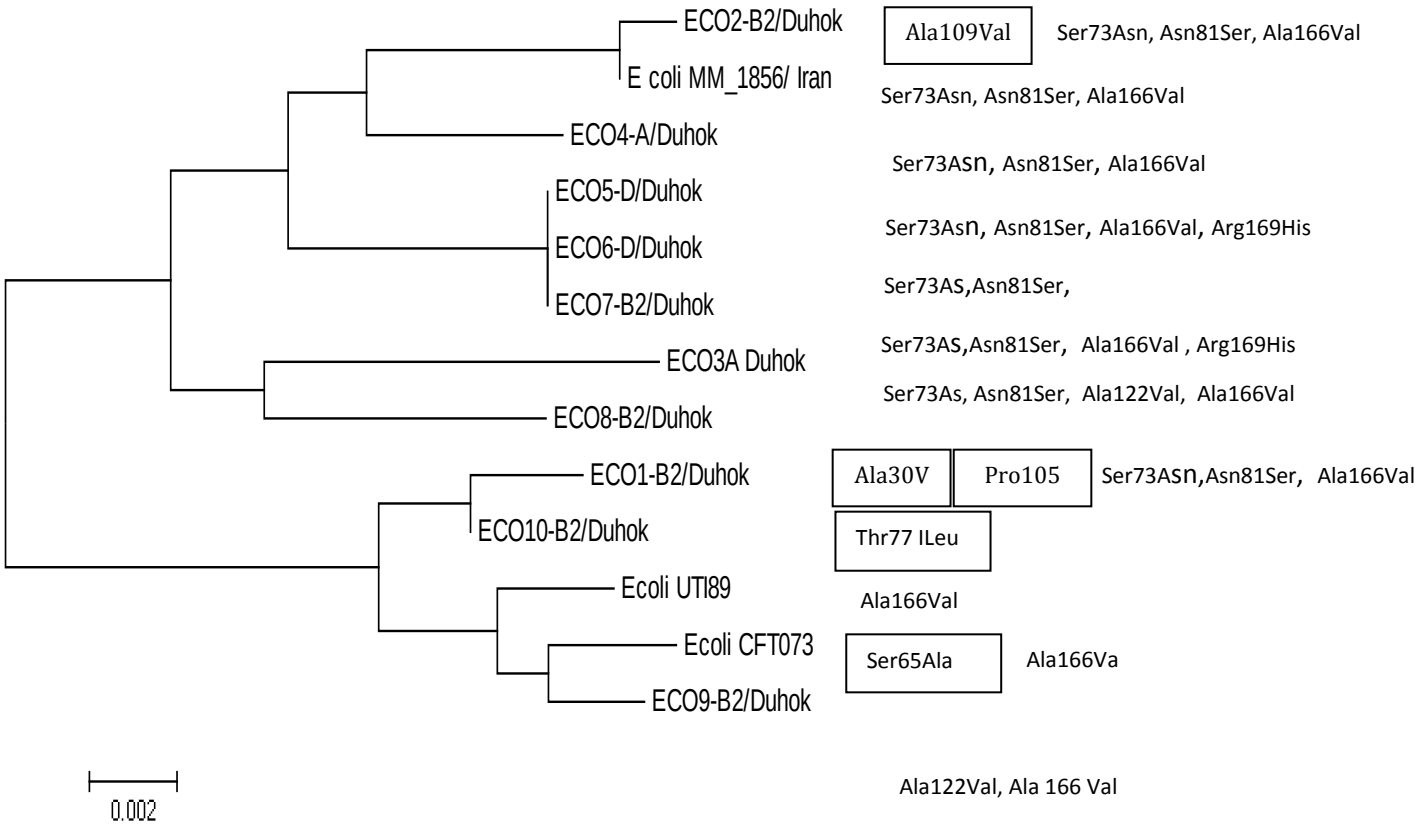

Figure 1. Phylogenetic tree obtained from 13 *Fim*H sequence variants chosen from 10 *fim*H containing clinical isolates of *E. coli* and reference *E. coli* CFT073, *E coli* MM_1856/ Iran and UTI89 (specified by MEGA 4, by the application of Kimura 2-parameter model [12]. *fim*H sequences were interpreted by comparison with the *E. coli* CFT073 sequences. Boxes resembled amino acid substitutions found in an individual strain (singleton) which culminate in 5 *fim*H protein variants deduced from analyzing nucleotide sequences. Non-singleton amino acids are those non-boxed in the figure. The independent alteration of amino acids in the same position can be found in different strains. The phylogroup capital letters (A, B1, B2, and D) are determined by Triplex-PCR.

Discussion

Urinary tract infection is one of the most common widely distributed hospital-acquired infections. The ability of *Escherichia coli* to expresses a number of different adhesive organelles as essential requirement for colonizing the urinary tract including P, type 1, S, and F1C fimbriae makes it to be the predominant pathogen responsible for urinary tract infections[13, 14].

Type 1 fimbriae are produced by more than 80% of all UPEC, it is very well established that the production of type 1 fimbriae by *E. coli* is a determining virulence trait for pathogenic strains [13].The adhesive subunit of type 1 fimbriae (*Fim*H) is an essential determinant, which has increased tropism for urinary tract cell receptors; thus, *Fim*H

adhesin is critical for colonizing various sites by *E. coli* [4].

The prevalence of *fim*H gene was previously investigated by various researchers in different countries. The results of the present study revealed consistent observations regarding the prevalence of *fim*H gene among UPEC isolates with the results of others [15] who stated that 97.5% of the UPEC were positive of the existence of *fim*H gene. This high prevalence was also noted by another group of scientists[9]who showed that *fim*H gene was found in 316 (92%) among 345 UPEC isolates and others [4] who reported a percentage of 92.8 % of UPEC isolates which harbored *fim*H gene. Genotyping tests had improved our knowledge about the epidemiology of UTI due to UPEC, these tests greatly improve our abilities to characterize modes of transmission of a pathogen, determine the source and risk factors of an infection [8]. Phylogenetic grouping of *E. coli* based on Triplex-PCR assay has been used extensively worldwide as a simple and relatively inexpensive method for assigning *E. coli* isolates into different phylogroups including; A, B1, B2, and D and submitted further clues that strains of the phylogroups diverse in their phenotypic and genotypic traits, their ecological niche and potential capability to result in disease [16]. This strategy based on different combinations of the presence and/or absence of three genetic markers namely, *chuA* and *yjaA* genes and TSPE4.C2 fragment as proposed by group of researchers [2]. It have been demonstrated that virulent extra-intestinal *E. coli* strains belong mainly to phlogroups B2 and D while groups A and B1 represent the most commensal strains [17]. Interestingly, [18] gave conceivable interpretation behind the predominance of phylogroup B2 that this phylogroup is usually of human source and proposed that UTIs usually is attributed to contamination by feces. Furthermore, since most *E. coli* strains belonged to phylogroup A and B1 considered as commensal and usually originate from non-human sources such as soil, water or animals [8], the probability of contamination by these isolates is more likely accepted as a source of infection [10].

Sequence-based genotyping approaches are becoming increasingly popular in epidemiological investigations of infectious diseases such as UTIs. Since *fimH* gene is frequently associated with UPEC strains and it is more likely to be altered or modified due to selective pressure, the phenotypic variants of *fim*H gene is earnestly associated with genetic variations thus, SNPs may contribute to the ability of organisms to cause illness conferring epidemic distribution or long term evolution of virulence [15]. SNP *fim* H analysis has discriminating power for this locus and it may be accurate enough for investigating UTI caused by UPEC that occurs over limited time periods or in confined geographical settings [19]. This study may the first attempt as our knowledge for molecular characterization of *E coli* based on SNP of *Fim*H analysis in Dohuk Province and involved a limited number of clinical samples.

It's certain that the majority of studied clinical isolates of *E. coli* shared the same sense mutations which gave rise to amino acid replacements (Ser73Asn, Asn81Ser, Ala166Val), despite they belonged to different phylogroups according to Triplex PCR analysis. By comparing the results of both genotyping based on sequencing of *fim*H gene with phylogenetic grouping based on Triplex PCR analysis, it's obvious that the two types of results are irrelevant, for instance, although Eco/5, Eco6 and Eco7were assigned to one allelotype, Eco5, ECO6 belonged to Phylogroup D while Eco7/Duhok strain belonged to phylogroup B2. Moreover, most clinical isolates which belong to B2 Phylogroup

had unique amino acid replacements, but all other groups lacked them, this observation led the way for suggesting an effect of these mutations on *fim*H gene related virulence among these isolates. Horizontal transmission of virulence genes like *fim*H among isolates regardless of phylogroups may explain such results. Interestingly, phylogroup B2 isolates are more virulent than others and are mostly predominant among UPEC strain thus, these isolates are supposedly exposed to selective pressure more than isolates which belong to other phylogroups. It's mentionable that most isolates shared reference strain *E coli*/MM-1865 isolated from Iran with most non-sense mutations in addition to three sense mutations (Ser73Asn, Asn81Ser, Ala166Val). This technique may has effectively role for finding the track of relevant strains in geographical studied area.

The results of the current study were in line with these of others who emphasized that a combination of Triplex PCR and SNP *fim* H analysis being practiced with extra intestinal *E. coli* strains could serve as a fast, highly reproducible typing tool for epidemiological explorations of ExPEC [19,1].

Discrepancy between SNP and PCR typing maybe attributed to horizontal gene transfer that alter the SNP phylogeny, in particular when a difference of the two techniques showed among closely related groups, such as groups A and B1 or groups D and B2. Alternatively, as the PCR typing is founded on only three genomic DNA segments, Taken results from both techniques together may serve for epidemiological studies and evolution of ExPEC strains [1].

References

1. Hommais F, Pereira S, Acquaviva C, Escobar-Pa´ramo P, Denamur E. Single-Nucleotide Polymorphism Phylotyping of *Escherichia coli*, App Environ Microbiol. 2005; 71(8):4784–4792.
2. Clermont, O, Bonacorsi, S, and Bingen, E. Rapid and simple determination of the *Escherichia coli* phylogenetic group. Appl Environ Microbiol. 2000; 66(10):4555-4558.
3. Gordon David M., Clermont Olivier, Tolley Heather and Denamur Erick. Assigning *Escherichia coli* strains to phylogenetic groups: multi-locus sequence typing versus the PCR triplex method. Environ. Microbiol.,2008; 10(10): 2484–2496.
4. Hojati Z, Zamanzad B, Hashemzadeh M, Molaie R, Gholipour A. The *FimH* Gene in Uropathogenic *Escherichia coli* Strains Isolated From Patients With Urinary Tract Infection. Jundishapur J Microbiol. 2015; 8(2): e17520.
5. Wiedenbeck J, Cohan FM. Origins of bacterial diversity through horizontal genetic transfer and adaptation to new ecological niches, FEMS Microbiol Rev.2011; 35: 957–976
6. Langille M. GI and Brinkman F. SL. Bioinformatic detection of horizontally transferred DNA in bacterial genomes Morgan GI Langille and Fiona SL Brinkman. Biology Reports.2009;1:25
7. Klemm Per. Molecular Characterization of the *Escherichia coli FimH* Adhesin; J Infect Dis. 2001; 183(Suppl 1):S28–S31.
8. Tartof SY, Solberg OD, Riley LW. Genotypic analyses of uropathogenic *Escherichia coli* based on *fim*H single nucleotide polymorphisms (SNPs). J Med Microbiol. 2007; 56(10):1363-1369.
9. Dias Rubens CS, Moreira Beatriz M., Riley Lee. *fim*H single nucleotide poly-morphisms to strain type clinical isolates of *Escherichia coli* for 6 epidemiologic investigation, J. Clin. Microbiol., 2009; doi:10.1128.
10. Merza NS, Jubrael JMS. Phylogenetic Grouping of Uropathogenic *Escherichia coli* Using Different Molecular Typing Methods in Kurdistan Region-Iraq. Int. J chem. BioMol. Sci.; 2015; 51-55.

11. Sambrook J, Russell RW. Molecular cloning: A laboratory manual, 3rd ed. Cold spring harbor laboratory press, Cold spring harbor, N.Y.2001.
12. Tamura K, Dudley J, Nei M, Kumar S. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 2007;24:1596–1599.
13. Schembri MA, Kjaergaard K, Sokurenko EV, Klemm P. Molecular Characterization of the *Escherichia coli* FimH Adhesin. J Infect Dis., 2001; 183(Suppl 1):S28–S31.
14. Valiquette, L. Urinary tract infections in women. Can J Urol 8 (Suppl. 1), 2001; 6–12.
15. Lai Yun Mei, Norgainathai Rina, Zaw Myo Thura and Lin Zaw. A New Primer Set for Detection of fimH Gene in *Escherichia coli* Isolates. Int J Collaborative Res on Internal Med & Pub Health, 2015; 7(4).
16. Clermont O, Christenson JK., Denamur E,Gordon DM. The Clermont *Escherichia coli* phylotyping method revisited: improvement of specificity and detection of new phylo-groups, Environ Microbiol Reports.2013;5(1): 58–65.
17. Iranpour D, Hassanpour M, Ansari H, Tajbakhsh S, et al. Phylogenetic Groups of *Escherichia coli* Strains from Patients with Urinary Tract Infection in Iran Based on the New Clermont Phylotyping Method, Biomed Res Int. 2015; 846219.
18. Carlos C, Pires M.M, Stoppe NC, Hachich, EM, Sato MI, Gomes TA, Amaral LA, Ottoboni, LM. *Escherichia coli* phylo-genetic group determination and its application in the identification of the major animal source of fecal contamination. BMC Microbiol.2010; 10: 161.
19. Abdallah KS, Cao Y, Wei DJ.Epidemiologic Investigation of Extraintestinal pathogenic E. coli (ExPEC) based on PCR phylogenetic group and fimH single nucleotide polymorphisms (SNPs) in China. Int J Mol Epidemiol Genet. 2011; 2(4):339-53.