Iran J Public Health, Vol. 49, No.12, Dec 2020, pp.2409-2417

Original Article

Multilocus Variable-Number Tandem-Repeat Analysis for Genotyping of *Escherichia coli* Strains Isolated from Hospital Wastewater, Tehran, Iran

Omid FARAHANI [1]***, *Reza RANJBAR*** [2]***, Sahar HONARMAND JAHROMY*** [1]***, Bahareh ARABZADEH*** [2]

1. *Department of Microbiology, Islamic Azad University, Varamin-Pishva Branch, Varamin, Iran*
2. *Molecular Biology Research Center, Systems Biology and Poisonings Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran*

***Corresponding Author:** Email: ranjbarre@gmail.com

(Received 14 Apr 2019; accepted 17 Jun 2019)

Abstract

Background: *Escherichia coli* is one of the most frequent causes of many common bacterial infections. As a potential reservoir, hospital wastewater is considered for the dissemination of bacterial pathogens such as *E. coli*. Therefore, research on hospital waste's bacteria by low-cost, rapid and easy molecular typing methods such as multilocus variable-number tandem-repeat analysis (MLVA) can be helpful for the study of epidemics.

Methods: *E. coli* strains were isolated from hospital wastewater sources in Tehran, Iran, over a 24-month sampling period (Jun 2014- Jun 2016) and identified by standard bacteriological methods. The diversity of repeated sequences of seven variable-number tandem-repeat (VNTR) loci was studied by MLVA method base on polymerase chain reaction (PCR).

Results: Overall, 80 *E. coli* isolates were discriminated into 51 different genotypes. Analysis of the MLVA profiles using a minimum spanning tree (MST) algorithm showed two clonal complexes with 71 isolates and only nine isolates were stayed out of clonal complexes in the form of a singleton. High genotypic diversity was seen among *E. coli* strains isolated from hospital wastewaters; however, a large number of isolates showed a close genetic relationship.

Conclusion: MLVA showed to be a rapid, inexpensive and useful tool for the analysis of the phylogenetic relationships between *E. coli* strains under the study.

Keywords: *Escherichia coli;* Multilocus variable number tandem repeat analysis; Hospital wastewater; Genotyping; Phylogenetic relationships

Introduction

Hospital wastewater is hazardous for organisms' life since it contains many pollutants such as chemical products, drugs, and pathogenic microorganisms (1). Contamination of surface waters with hospital wastewater will become a threat to public health (2). If the hospital wastewater is discharged into the environment, it can spread pathogen microorganisms such as *Escherichia coli* (3, 4). Generally, *E. coli* bacteria are known as a part of the natural gut microflora of humans, but they also have pathogenic strains (5, 6). Some *E. coli* classes

including enteroaggregative *E. coli* (EAEC), enterohemorrhagic *E. coli* (EHEC), enteropathogenic *E. coli* (EPEC), enteroinvasive *E. coli* (EIEC) and enterotoxigenic *E. coli* (ETEC) are a significant cause of diarrhea in the world, especially in developing countries and some others are associated with extraintestinal infections like uropathogenic *E. coli* (UPEC) and meningitis-associated *E. coli* (MAEC) that cause urinary tract infection meningitis in neonates, respectively (7-11). *E. coli* has different strains with genetic and genotype diversity (6).

Therefore, their genotyping is essential. There are some methods for genotyping, such as pulsed-field gel electrophoresis (PFGE), multilocus sequence typing (MLST), repetitive sequence-based PCR (Rep-PCR) and multilocus variable-number tandem-repeat (VNTR) analysis (MLVA) (12- 14). Although PFGE is a gold standard method for determining the diversity of bacteria, as an excellent tool for outbreaks analyzing, it is expensive, time-consuming with professional skill requirements. The ability of MLVA in discrimination of bacterial strains is not less than PFGE, and its discriminatory power is much higher than MLST. Moreover, compared to other methods, MLVA is a low-cost and straightforward method (15, 16).

The current study aimed to use the MLVA method for molecular typing of *E. coli* strains isolated from hospital wastewater sources in Tehran, Iran.

Materials and Methods

Wastewater sample collection and E. coli isolation

The present study included non-duplicate *E. coli* isolates originated from hospital wastewaters of twenty hospitals in Tehran, Iran over a 24-month sampling period (Jun 2014- Jun 2016).

After sampling, bottles were rapidly placed in a lightproof insulated box including icepacks to ensure rapid cooling. The samples were shipped to the laboratory for further analysis. To isolate the *E. coli*, conventional biochemical methods were used according to Standard Methods for the Examination of Water and Wastewater 22nd edition (17) and then samples were cultured in different environments. *E. coli* colonies from agar plates were picked and streaked for purity on EMB agar. Identification of *E. coli* strains was made by standard bacteriological procedures. Well-isolated colonies of purified *E. coli* were resuspended in trypticase soy broth with 20% (v/v) glycerol and stored in -70 °C for long-term storage, as described previously (18).

DNA preparation

A pure culture of *E. coli* was plated on LB agar and incubated overnight at 37 °C. A single colony was removed from the plate, suspended in 200 µl of sterile deionized water, and boiled for 15 minutes. After centrifugation at 8,000 gr for 6 min, the supernatant was transferred into a new tube for subsequent PCR analysis.

MLVA assay

These seven VNTR loci were selected: ms06, ms07, ms09, ms11, ms21, ms23 and ms32. The primer sets for PCR amplification of these VNTR loci were previously reported (Table 1) (19).

Table 1: MLVA primers and annealing temperatures for PCR reactions

MLVA locus	*Forward primer 5'to 3'*	*Reverse primer 5'to 3'*	*Ta°C*
ms06	AAACGGGAGAGCCGGTTATT	TGTTGGTACAACGGCTCCTG	55°C
ms07	GTCAGTTCGCCCAGACACAG	CGGTGTCAGCAAATCCAGAG	55°C
ms09	GTGCCATCGGGCAAAATTAG	CCGATAAGGGAGCAGGCTAGT	55°C
ms11	GAAACAGGCCCAGGCTACAC	CTGGCGCTGGTTATGGGTAT	55°C
ms21	GCTGATGGCGAAGGAGAAGA	GGGAGTATGCGGTCAAAAGC	55°C
ms23	GCTCCGCTGATTGACTCCTT	CGGTTGCTCGACCACTAACA	55°C
ms32	TGAGATTGCCGAAGTGTTGC	AACTGGCGGCGTTTATCAAG	55°C

PCR was performed in 25 µl volume containing 1X PCR buffer (50mmol/L KCL, 10 mmol/L Tris, pH9), 2.5 mmol/L $MgCl_2$, 0.2 mmol/L of each primer with 0.5 U TaqDNA polymerase (CinnaGen Co., Iran), and 3 µl of DNA extract. Cycling conditions for PCR reactions were 93 °C for 5 min, followed by 34 cycles of 93 °C for 30 sec, 55 °C for 1 min, and 72 °C for 1 minute. The PCR products were run on agarose gels, stained with ethidium bromide, and visualized under UV transillumination by gel doc.

Data analysis

The number of repetitions was calculated according to the size of the amplicon by manual reading (Fig.1).

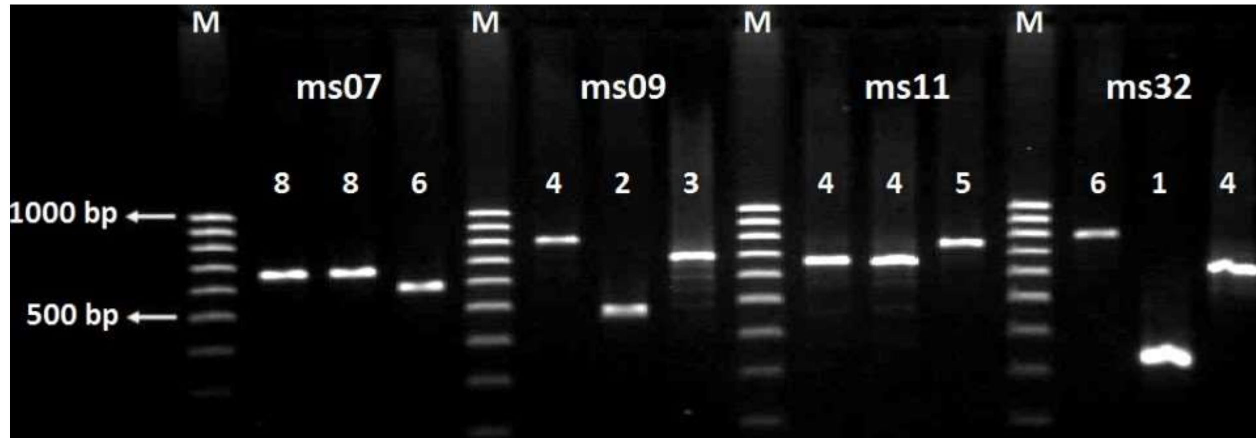

Fig. 1: Polymorphism of 3 VNTR loci in different E. coli isolates. This image illustrates how the number of repeats can be directly deduced by manual reading

Amplicon sizes were converted into quantities of repeats based on the formula: the number of repeats (bp) = PCR product size (bp) − flanking regions (bp)/repeat size (bp).

Repeat numbers of each locus were imported into Microsoft Excel, and allelic profiles were obtained. The minimum spanning tree (MST) was constructed with a categorical coefficient based on allelic profiles of the *E. coli* strains using 7.6.1 version of bionumerics software (Applied Maths NV, Sint-Martens-Latem, Belgium). MST is a convenient complimentary tool to cluster multiple isolates and visualize the relative diversity within different lineages. A dendrogram of genetic relationships was also generated using the unweighted pair group method with arithmetic averages (UPGMA) method (20). Furthermore, Simpson's index of diversity (D) and 95% confidence intervals (CI) for each VNTR locus were calculated using version 2.0 phyloviz software (http://www.phyloviz.net)

Results

Overall, 91 isolates suspected to *E. coli* were recovered based on culture characteristics. Eighty isolates were definitely identified as *E. coli* by standard biochemical and serogical procedures and subjected to the study. As the result of PCR, the size distributions of the 80 *E. coli* strains for each VNTR were 301–379 bp (ms06), 392–821 bp (ms07), 445–1432 bp (ms09), 718–958 bp (ms11), 233–1492 bp (ms21), 584–959 bp (ms23), and 355–810 bp (ms32).

Simpson's index of diversity, confidence Interval 95%, number of different alleles per locus and Tandem repeat (TR) size were counted (Table 2).

Analysis of VNTR showed different genotypes in isolates. 38 *E. coli* isolates were divided into 2 clonal complexes, and nine isolates were out of the clonal complexes in the form of Singleton (Fig. 2).

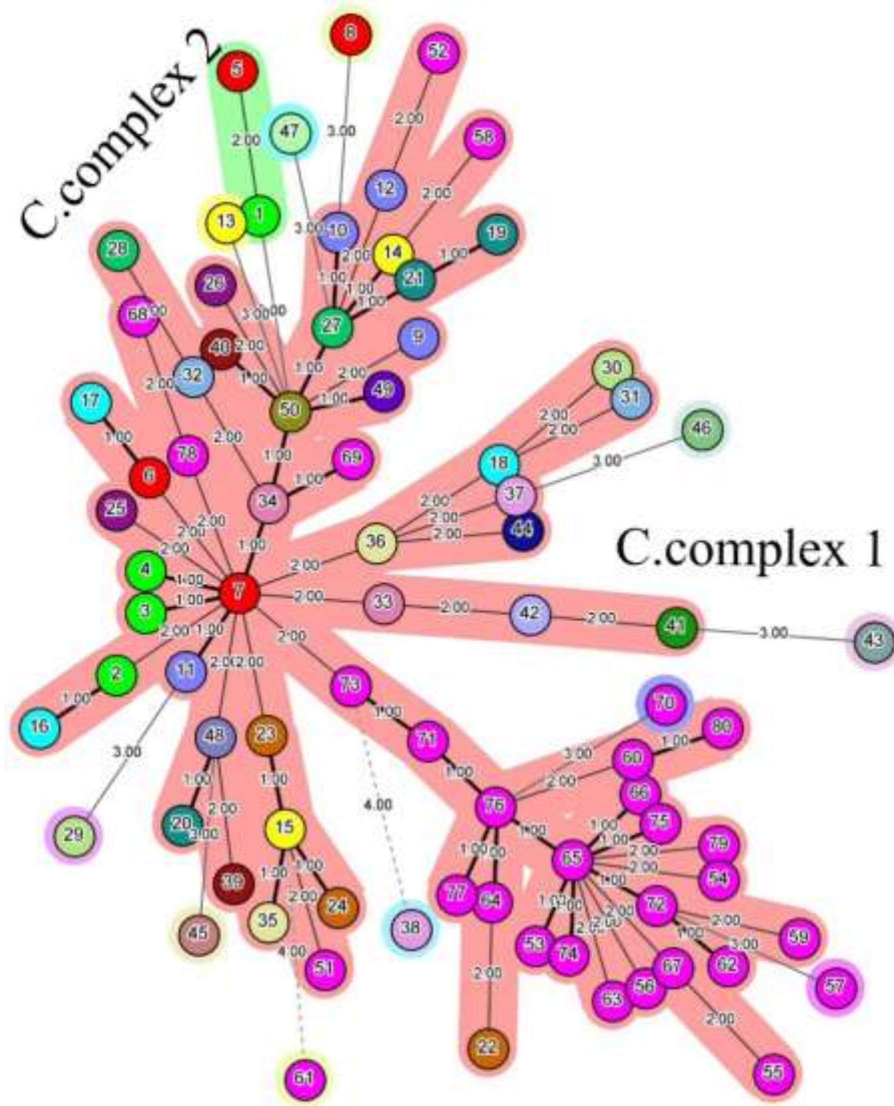

Fig. 2: Data analysis by using the Minimum Spanning Tree (MST) algorithm with a difference of 2 loci. Each Circle is showing an isolate with a unique profile. Lines representing isolates relationship and numbers on them refers to the number of genetic loci difference. 2 clonal complexes are shown with haloes. The colors of the circles correspond to the region of isolation

According to the UPGMA method and Dice similarity coefficient 80, *E. coli* isolates were classified in three clusters, namely cluster number 1(C1) contained 7 isolates which none of them were from a same hospital, cluster number 2(C2) included the maximum numbers of isolates (70 out of 80) and all hospitals had isolated in this cluster, finally cluster number 3(C3) that consisted of just 3 isolates from two different hospitals. Most of the isolates were in cluster 2(C2), for this reason, this cluster was divided into two sub-clusters C1a with 25 isolates and C1b with 45 isolates (Fig. 3).

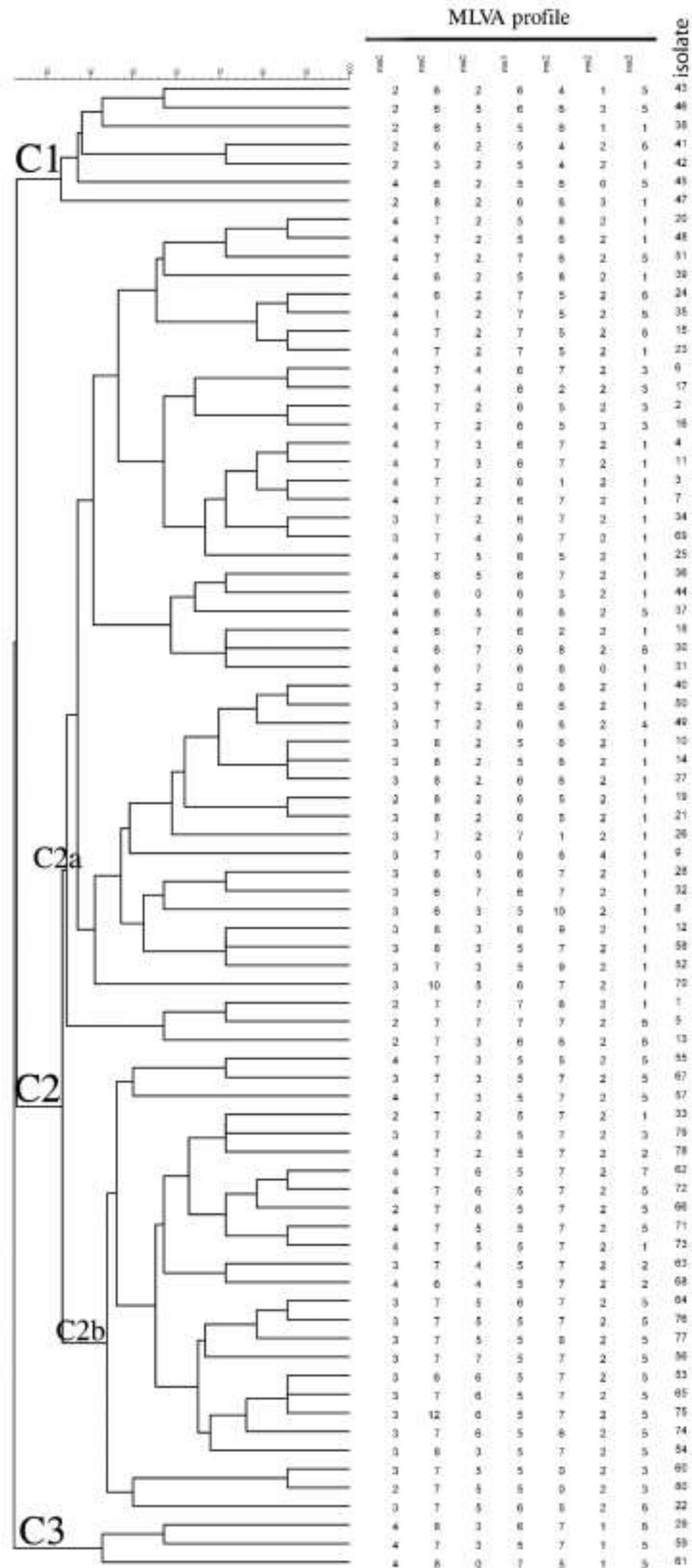

Fig. 3: Clustering of the MLVA profiles by UPGMA with the categorical coefficient of similarity

Discussion

Now, it is the first study of using MLVA method among *E. coli* strains isolated from hospital wastewaters. Accordingly, seven VNTR loci were analyzed by MLVA method for genotyping of the strains.

One of the achievements of this research refers to describing the ability of MLVA to differentiate the tested isolates. It can offer valuable information about the relationship between hospital wastewater *E. coli* strains.

In recent years, various studies based on VNTR have been performed on several bacteria, such as *Shigella* (13, 14, 21), *Brucella* (22, 23), *Salmonella* (24, 25), *Coxiella* (26) and *E. coli*, especially O157:H7 serotype (27-30). However, according to the source of isolation, no VNTR research has been done on *E. coli* bacteria isolated from hospital wastewater.

One most important step in subtyping protocols is the TR loci choosing for analysis. In France, study on 19 *Shigella flexneri*, 26 *S. boydii*, 9 *S. sonnei*, 23 *S. dysenteriae*, 1 *Salmonella enterica* and 11 *E. coli* isolates, 15 VNTR loci such as ms06, ms07, ms09, ms11, ms21, ms23 and ms32 were chosen and evaluated by designed primers. The selected loci VNTR in their study had suitable repeat sizes, so it was easy to estimate repeat sizes by electrophoresis on agarose gel (19). The mentioned genes and primers also were used by others (13, 18, 21, 31) and they reached to their goals. We targeted these genes by using mentioned primers in MLVA. The results of this study showed that these primers are precise and appropriately selected.

Typing of *E. coli* strains contains older and newer objectives. More former purposes are related to epidemiological studies, local and regional strains identification and the possible changes in the native strains. In new targets, the study of evolution and phylogenetic tree were added to the previous aims. In older studies, selected loci had small repeat size. For this reason, accurate calculation of PCR amplicon sizes is needed to capillary electrophoresis or DNA sequencing that they are costly. In this study, the selected sizes are perfectly adequate, and results are accurate.

Simplicity, low cost, rapid and the ability to store information and comparison with results from other researchers are the advantages of the MLVA technique (32). The results of this study confirmed the above-mentioned reports. Based on the obtained data of the study, the diversity of seven TR loci was calculated with values ranging from 0.253 to 0.807 (Table 2). In addition, 51 different alleles were observed in this study. The highest level of diversity was related to *ms21* and this was the most polymorphic locus in the current study with 14 different alleles. The lowest level of diversity was related to ms23. Similar observations were reported (19).

Table 2: VNTR locus, tandem repeat size, Simpson's index of diversity and confidence interval 95%

VNTR locus	*TR size (bp)*	*Simpson's index of diversity*	*confidence Interval95%*	*No. of alleles*	*Null alleles*
ms06	39	0.6301	0.587, 0.6732	3 (2, 3, 4)	No
ms07	39	0.5921	0.498, 0.6862	7 (1, 3, 6, 7, 8, 10,12)	No
ms09	179	0.8066	0.7453, 0.868	9 (0, 1.5, 2.5, 3, 4, 4.5, 5, 6, 7)	Yes
ms11	96	0.681	0.6234, 0.7387	6 (0, 4.5, 5, 5.5, 6, 7)	Yes
ms21	141	0.8073	0.747, 0.8676	14(0, 1, 2, 3, 4, 5, 5.5, 6, 6.5, 7, 7.5, 8, 9, 10)	Yes
ms23	375	0.2532	0.1271, 0.3792	5 (0, 1, 2, 3, 4)	Yes
ms32	101	0.7038	0.6315, 0.7761	7 (1, 2, 3, 4, 4.5, 5, 5.5)	No

In a study on EPEC and STEC in Norway 2011, seven VNTR loci were investigated (33). Although it's obtained data were not similar to the data of the current study but both studies were common in null

allele observation so that null alleles were seen in both studies. This can occur because of a mutation in the primer-binding site on the DNA.

One of the essential aims of this study was pointed to investigate genetic relationship between *E. coli* strains isolated from hospitals' wastewater in Tehran, Iran. As the result of this study, 86% of *E. coli* isolates had a high molecular affinity and gathered in a common clonal complex (CC1), and only 11% of isolates did not collect in a group. The remaining isolates owned another clonal complex (CC2) (Fig. 2) and were closely related isolates.

Since similar studies did not perform on *E. coli* strains recovered from hospital wastewaters, this study can be considered as a primary base for other studies focusing on the genetic relationship between *E. coli* and other non-*E. coli* isolates recovered from hospital wastewater sources in the future.

Conclusion

This study showed low genotypic diversity among *E. coli* strains isolated from hospital wastewater in Tehran, Iran. MLVA showed to be comfortable, high-operational, rapid and dependable method for genotyping of our strains. Future studies should carefully compare MLVA with older genotyping techniques such as PFGE that is the gold standard genotyping method.

Ethical considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Acknowledgements

We would like to thank from the "Clinical Research Development Center of Baqiyatallah hospital" for their kindly cooperation. This study was financially supported in part by "Clinical Research Development Center of Baqiyatallah hospital".

Conflict of interest

The authors declare that there is no conflict of interests.

References

1. Moges F, Endris M, Belyhun Y, Worku W (2014). Isolation and characterization of multiple drug resistance bacterial pathogens from waste water in hospital and non-hospital environments, Northwest Ethiopia. *BMC Res Notes,* 7:215.
2. McLellan SL (2004). Genetic diversity of *Escherichia coli* isolated from urban rivers and beach water. *Appl Environ Microbiol,* 70:4658-4665.
3. Yang C, Lin M, Liao P, et al (2009). Comparison of antimicrobial resistance patterns between clinical and sewage isolates in a regional hospital in Taiwan. *Lett Appl Microbiol,* 48:560-565.
4. Ranjbar R, Masoudimanesh M, Dehkordi FS, et al (2017). Shiga (Vero)-toxin producing *Escherichia coli* isolated from the hospital foods; virulence factors, o-serogroups and antimicrobial resistance properties. *Antimicrob Resist Infect Control,* 6:4.
5. Anvarinejad M, Farshad S, Ranjbar R, et al (2012). Genotypic analysis of *E. coli* strains isolated from patients with cystitis and pyelonephritis. *Iran Red Crescent Med J,* 14(7):408-16.
6. Timmons C, Trees E, Ribot EM, et al (2016). Multiple-locus variable-number tandem repeat analysis for strain discrimination of non-O157 Shiga toxin-producing *Escherichia coli*. *J Microbiol Methods,* 125:70-80.
7. Trabulsi LR, Keller R, Gomes TAT (2002). 10.321/eid0805.Typical and Atypical Enteropathogenic *Escherichia coli*. *Emerg Infect Dis,* 8(5):508-13.
8. Giuffrè M, Geraci DM, Bonura C, et al (2016). The increasing challenge of multidrug-resistant Gram-negative bacilli: results of a 5-year active surveillance program in a neonatal intensive care unit. *Medicine (Baltimore),* 95(10):e3016.
9. Sabaté M, Prats G, Moreno E, et al (2008). Virulence and antimicrobial resistance profiles among *Escherichia coli* strains isolated from human and animal wastewater. *Res Microbiol,* 159:288-293.

10. Momtaz H, Karimian A, Madani M, et al (2013). Uropathogenic *Escherichia coli* in Iran: serogroup distributions, virulence factors and antimicrobial resistance properties. *Ann Clin Microbiol Antimicrob,* 12:8.
11. Miquel S, Peyretaillade E, Claret L, et al (2010). Complete genome sequence of Crohn's disease-associated adherent-invasive *E. coli* strain LF82. *PLoS One,* 5:e12714.
12. Sadeghifard N, Ranjbar R, Zaeimi J, et al (2010). Antimicrobial susceptibility, plasmid profiles, and RAPD-PCR typing of Acinetobacter bacteria. *Asian biomedicine*, 4(6):901-911.
13. Ranjbar R, Memariani M, Memariani H (2015). Diversity of variable number tandem repeat loci in *Shigella* species isolated from pediatric patients. *Int J Mol Cell Med,* 4(3):174-81.
14. Filliol-Toutain I, Chiou CS, Mammina C, et al. (2011). Global distribution of Shigella sonneiclones. *Emerg Infect Dis,* 17: 1910-1912.
15. Malachowa N, Sabat A, Gniadkowski M, et al (2005). Comparison of multiple-locus variable-number tandem-repeat analysis with pulsed-field gel electrophoresis, spa typing, and multilocus sequence typing for clonal characterization of *Staphylococcus* aureus isolates. *J Clin Microbiol,* 43:3095-3100.
16. Behzadi P, Najafi A, Behzadi E, et al (2016). Microarray long oligo probe designing for Escherichia coli: An in-silico DNA marker extraction. *Cent European J Urol*, 69(1): 105-11.
17. Association APH, Association AWW, Federation WPC, Federation WE (1915). *Standard methods for the examination of water and wastewater*. ed. American Public Health Association.
18. Memariani M, Peerayeh SN, Salehi TZ, Mostafavi SKS (2015). Occurrence of SHV, TEM and CTX-M β-lactamase genes among enteropathogenic *Escherichia coli* strains isolated from children with diarrhea. *Jundishapur J Microbiol,* 8(4):e15620.
19. Gorgé O, Lopez S, Hilaire V, et al (2008). Selection and validation of a multilocus variable-number tandem-repeat analysis panel for typing *Shigella* spp. *J Clin Microbiol,* 46:1026-1036.
20. Weniger T, Krawczyk J, Supply P, et al (2010). MIRU-VNTR plus: a web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. *Nucleic Acids Res,* 38(Web Server issue):W326-31.
21. Ranjbar R, Memariani M (2015). Multilocus variable-number tandem-repeat analysis for genotyping of *Shigella sonnei* strains isolated from pediatric patients. *Gastroenterol Hepatol Bed Bench,* 8(3):225-32.
22. Yang Y, Wang Y, Poulsen E, et al (2017). Genotyping *Brucella canis* isolates using a highly discriminatory multilocus variable-number tandem-repeat analysis (MLVA) assay. *Sci Rep,* 7(1):1067.
23. Menshawy A, Perez-Sancho M, Garcia-Seco T, et al (2014). Assessment of genetic diversity of zoonotic *Brucella* spp. recovered from livestock in Egypt using multiple locus VNTR analysis. *BioMed Res Int,* 2014: 353876.
24. Wang H, Diao B, Cui Z, Yan M, Kan B (2016). Genotyping of *Salmonella typhi* using 8-loci multi locus VNTR analysis. *Gut Pathog,* 8:14.
25. Ghaderi R, Tadayon K, Avagyan S, et al (2013). The population structure of *Salmonella* enterica Enteritidis in Iran analyzed by multiple-locus variable-number tandem repeat analysis. *Trop Anim Health Prod,* 45:889-894.
26. Santos AS, Tilburg JJ, Botelho A, et al (2012). Genotypic diversity of clinical *Coxiella burnetii* isolates from Portugal based on MST and MLVA typing. *Int J Med Microbiol,* 302:253-256.
27. Swirski A, Pearl D, Williams M, et al (2014). Spatial epidemiology of *Escherichia coli O157: H7* in dairy cattle in relation to night roosts of Sturnus vulgaris (European starling) in Ohio, USA (2007–2009). *Zoonoses Public Health,* 61:427-435.
28. Byrne L, Elson R, Dallman TJ, et al (2014). Evaluating the use of multilocus variable number tandem repeat analysis of Shiga toxin-producing *Escherichia coli O157* as a routine public health tool in England. *PLoS One,* 9:e85901.
29. Perry N, Cheasty T, Dallman T, et al (2013). Application of multilocus variable number tandem repeat analysis to monitor Verocytotoxin-producing *Escherichia coli O157* phage type 8 in England and Wales: emergence of a profile associated with a national outbreak. *J Appl Microbiol,* 115:1052-1058.
30. Dyet K, Robertson I, Turbitt E, Carter P (2011). Characterization of *Escherichia coli O157: H7* in New Zealand using multiple-locus variable-number tandem-repeat analysis. *Epidemiol Infect,* 139:464-471.

31. Nave HH, Mansouri S, Moghadam MT, Moradi M (2016). Virulence Gene Profile and Multilocus Variable-Number Tandem-Repeat Analysis (MLVA) of Enteroinvasive *Escherichia coli* (EIEC) Isolates From Patients With Diarrhea in Kerman, Iran. *Jundishapur J Microbiol,* 9(6):e33529.
32. Nadon C, Trees E, Ng L, et al (2013). Development and application of MLVA methods as a tool for inter-laboratory surveillance. *Euro Surveill,* 18:20565.
33. Sekse C, Sunde M, Lindstedt B, et al (2011). Potentially human pathogenic *Escherichia coli O26* in Norwegian sheep flocks. *Appl Environ Microbiol*, 77(14):4949-58.