

Original Article

Evaluation of PAI_{usp} subtyping to characterize uropathogenic *E. coli* isolates

Yun Mei Lai¹, Myo Thura Zaw¹, Shamsul Bahari Shamsudin², Zaw Lin¹

¹ Department of Pathobiological and Medical Diagnostics, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia

² Department of Community and Family Medicine, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia

Abstract

Introduction: Uropathogenic virulence factors have been identified by comparing the prevalence of these among urinary tract isolates and environmental strains. The uropathogenic-specific protein (USP) gene is present on the pathogenicity island (PAI) of uropathogenic *Escherichia coli* (UPEC) and, depending on its two diverse gene types and the sequential patterns of three open reading frame units (*orfUs*) following it, there is a method to characterize UPEC epidemiologically called PAI_{usp} subtyping.

Methodology: A total of 162 UPEC isolates from Sabah, Malaysia, were tested for the presence of the *usp* gene and the sequential patterns of three *orfUs* following it using polymerase chain reaction (PCR). In addition, by means of triplex PCR, the prevalence of the *usp* gene was compared with other two VFs of UPEC, namely alpha hemolysin (*α-hly*) and cytotoxic necrotizing factor (*cnf-1*) genes encoding two toxins.

Results: The results showed that the *usp* gene was found in 78.40% of UPEC isolates, indicating that its prevalence was comparable to that found in a previous study in Japan. The two or three *orfUs* were also associated with the *usp* gene in this study. All the PAI_{usp} subtypes observed in Japan were present in this study, while subtype IIa was the most common in both studies. The *usp* gene was observed in a higher percentage of isolates when compared with *α-hly* and *cnf-1* genes.

Conclusions: The findings in Japan and Sabah, East Malaysia, were similar, indicating that PAI_{usp} subtyping is applicable to the characterization of UPEC strains epidemiologically elsewhere in the world.

Key words: *usp* gene; uropathogenic *Escherichia coli*; PAI_{usp} subtyping; characterization.

J Infect Dev Ctries 2016; 10(10):1053-1058. doi:10.3855/jidc.6944

(Received 31 March 2015 – Accepted 27 November 2015)

Copyright © 2016 Lai *et al.* This is an open-access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Introduction

Urinary tract infections (UTIs) cause morbidity and significant mortality globally. *Escherichia coli* is the most common bacterium causing complicated as well as uncomplicated UTIs. Uropathogenic virulence factors (VFs) have been identified by comparing the prevalence among urinary tract isolates with that of environmental strains. VFs include adhesins, iron uptake systems, cytotoxins, and specific O:H:K serotypes [1,2].

The understanding of uropathogenic VFs, their attributes that cause UTIs, and the horizontal transfer of VF genes among the related species as pathogenicity islands, was achieved through molecular epidemiological studies. The information provided the understanding that virulence genes are the potential molecular markers for epidemiological studies [2].

Zonula occludens toxin (ZOT) represents a novel family of toxin. While searching for a homologue to the

zot gene of *Vibrio cholerae* in UPEC, Kurazono *et al.* identified a 4,167-bp putative pathogenicity island (PAI) that was commonly associated with UPEC strains. By using molecular methods, they observed an open reading frame encoding a protein that was named USP on that PAI. The open reading frame of the gene encoding USP is 1,038 bp long and it encodes a protein that has 346 amino acids. Surprisingly, the *usp* gene is followed by three small open reading frame units (designated as *orfU1*, *orfU2*, and *orfU3*) of 98, 97, and 96 amino acids, respectively [3].

USP was found significantly more often in UPEC strains than in fecal isolates from healthy individuals, and it existed in all common serogroups of UPEC [4]. Another study reported that USP enhanced the infectivity of *E. coli* in the mouse UTI model; *E. coli* isolates that were positive with USP had higher virulence to be lethal to mice when compared with other UPEC strains [5]. However, the precise mechanism of

USP on the urinary tract was not clear. Parret and De Mot suggested USP protein and small ORF following this protein showed high homology to the S-type pyocins produced by *Pseudomonas aeruginosa* strains and the immunity proteins following these bacteriocins [6,7]. In a recent study, Zaw *et al.* observed that USP was successfully purified after co-expression of the *usp* gene and *orfU1*, and USP has been observed to have non-specific DNase activity [8,9].

The previous investigator proposed that PAI_{usp} subtyping could be used as a molecular epidemiological method for UTI isolates for a number of reasons. First, the predominance of the *usp* gene was shown among the UTI isolates in a study by Kanamaru *et al.* Second, there are two gene types (namely *uspI* and *uspII*) according to the nature of the gene sequence at the 3' region 230 bp upstream of the stop codon. Third, the various sequential positions of the three associated *orfUs* immediately downstream of the *usp* gene were not identical (*i.e.*, *orfU1*, *orfU2*, and *orfU3* were in sequence, or *orfU2*, *orfU3*, and *orfU1* were in sequence, or sometimes only two *orfUs* were present in different combinations) [1,4].

Alpha hemolysin (α -Hly) and cytotoxic necrotizing factor 1 (CNF-1) are two well-known toxins of UPEC proven to have direct cytotoxicity to host tissues [10,11]. HlyA is usually encoded by PAIs of UPEC together with or without CNF-1 [12]. A CNF-1-positive UPEC strain caused greater acute inflammatory response of the bladder in a mouse model of ascending UTI than did its isogenic CNF-1 mutant [13].

In this study, the prevalence of the *usp* gene and its associated three *orfUs* were investigated in UPEC strains isolated from two hospitals located around Kota Kinabalu district of Sabah. The objectives of this study were to evaluate whether PAI_{usp} subtyping can be applicable to characterize UPEC isolates of Sabah and to investigate which PAI subtypes are circulating in the region. The prevalence of subtypes was compared to that reported in a previous study in Japan. This will provide information about whether the UPEC isolates studied were different from those in other geographical regions in terms of PAI_{usp} subtyping, which has significant epidemiological value. In addition, because the *usp* gene was assumed to have nuclease activity, the correlation of prevalence of *usp* and two other toxin genes, *cnf-1* and *α -hly*, was investigated by triplex polymerase chain reaction (PCR).

Methodology

Ethical clearance

Ethical clearance was obtained from the university ethical committee, Universiti Malaysia Sabah and National Medical Research Registration, Malaysia.

Bacterial strains and culture media

A total of 162 UPEC isolates from Hospital Queen Elizabeth and Hospital Papar that had significant bacteriuria along with 12 environmental *E. coli* isolates were investigated for *usp* and two other toxin genes in this study. Significant bacteriuria in the urine samples, which were sent to the microbiology laboratory of these two hospitals for culture and sensitivity tests, were examined.

The positive control for this study were the isolates carrying the *usp* gene which was confirmed by DNA sequencing and the nucleotides sequence was aligned with the reference sequence from NCBI website. Culture media used were nutrient agar, MacConkey agar (Oxoid, Hampshire, England), Luria-Bertani broth (Oxoid, Hampshire, England) and novobiocin mEC broth (Eiken Chemical, Tochigi, Japan).

Study of significant bacteriuria

One milliliter of urine sample was diluted to 10⁻⁶ dilution and 50 μ L of each dilution was inoculated on to nutrient agar by micropipette. The agar plates were incubated at 37°C for 18 hours every time. The significant bacteriuria was determined by counting the colonies on the plates that had 30–300 colonies. To get the bacterial count, the colony count on the plates was multiplied by the dilution factor.

Identification of Escherichia coli in urine samples

One hundred microliters of urine was inoculated into 1 mL of novobiocin mEC broth and incubated for 6 hours at 37°C. This bacterial suspension was subcultured onto MacConkey agar by the streak inoculation method and incubated at 37°C for 18 hours. The next day, lactose-fermenting colonies on the plates were confirmed for *E. coli* by IMViC tests.

Stocking of E. coli isolates

Isolates with typical biochemical test for *E. coli* were stocked on nutrient agar slant and glycerol stock. For the preparation of the glycerol stock, 700 μ L of overnight broth culture and 300 μ L of glycerol were mixed. Nutrient agar slants were stocked at 4°C and glycerol stocks were stored at -80°C in a freezer.

PCR for single-gene and triplex PCR

The nucleotide sequences of the oligonucleotide primers used in this study are shown in Table 1. Bacterial stocks were cultured on MacConkey agar plates and incubation was done for 18 hours at 37°C. The next day, the isolated colonies were incubated in LB broth for 18 hours at 37°C. One mL overnight broth culture was taken out and centrifuged at 6,149 g for 10 minutes. The pellet was suspended in 50 µL sterile distilled water. The DNA was extracted by the boiling method. Centrifugation was done at 6,149 g for 10 minutes and supernatant was taken by micropipette to be used as DNA template. Five microliters of DNA template was added to the PCR reaction mixture of 20 µL containing 10 µM of each primer, dNTPs 10 mmol, 1 × buffer, 1 unit of Taq polymerase (Takara Bio Inc., Shiga, Japan). For single-gene PCR, for the *usp* and *α-hly* genes, the conditions were 94°C × 5 minutes, 30 cycles of 95°C for 30 seconds, 65°C × 30 seconds, 72°C × 1 minute followed by 72°C × 10 minutes and reactions were stopped by 4°C. For the *cnf-1* gene, annealing temperature was 50°C instead of 65°C. In the case of triplex PCR, the PCR condition was changed to 60°C for annealing and 1.5 minutes for extension.

PCR was done in an Applied Biosystems Thermocycler (Applied Biosystems, Foster City, USA). The size of PCR products was checked by 1.5% TAE agarose gel. The gel was stained by ethidium bromide and visualized with gel documentation apparatus Alpha Imager HP System (Alpha Inotech Corp., San Leandro, USA).

PCR to investigate the sequential pattern of three orfUs

Three types of PCR were done simultaneously. PCR using *usp* forward primer and *orfU1* reverse primer, *usp* forward primer and *orfU2* reverse primer, *usp* forward primer and *orfU3* reverse primer were performed separately for the presence of *orfU1*, *orfU2*,

and *orfU3* in each reaction. The PCR conditions were 94°C × 5 minutes, 30 cycles of 95°C × 30 seconds, 55°C × 30 seconds, 72°C × 1.5 minutes followed by 72°C × 10 minutes, and reactions were stopped by 4°C. The sizes of PCR products were checked as mentioned above in the single-gene PCR.

Figure 1 A. Gel electrophoresis picture showing PCR product for the *usp* gene. Lane 2 is the positive control and lane 3 is the negative control. Lanes 5, 6, 14, 15, and 16 were UPEC isolates positive for the *usp* gene with a PCR product of 1,041 bp. Lanes 4, 7–13, 17 were the environmental isolates that did not contain the *usp* gene. A 100 bp ladder was used as a molecular marker as shown in lane 1. **B. Gel electrophoresis picture of triplex PCR.** Lanes 1 and 2 were two isolates positive for all three genes investigated by triplex PCR. DNA fragments were 1,041 bp, 489 bp, 1,177 bp, PCR products of *usp*, *cnf-1*, and *α-hly* genes, respectively. Lane 3 is 100 bp ladder molecular marker.

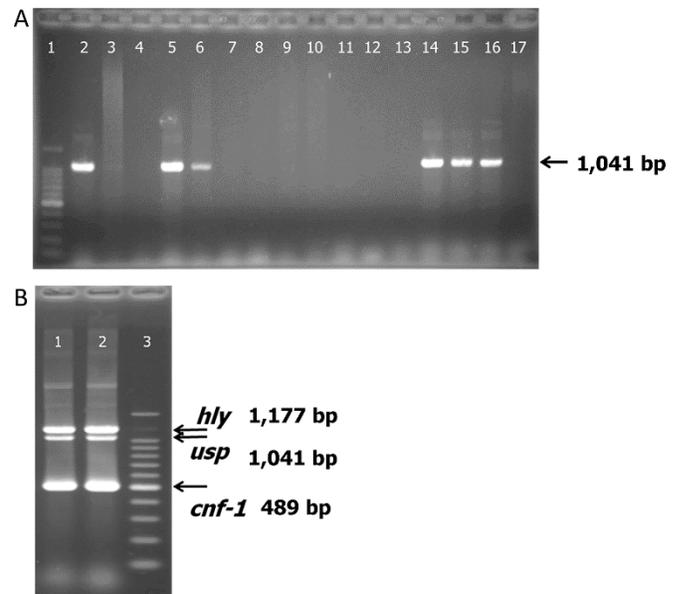


Table 1. Nucleotide sequences of primer sets used in this study.

Gene	Primer	Primer sequence (5'–3')
<i>α-hly</i>	F	AAC AAG GAT AAG CAC TGT TCT GGC T
	R	ACC ATA TAA GCG GTC ATT CCC GTC A
<i>cnf-1</i>	F	AAG ATG GAG TTT CCT ATG CAG GAG
	R	CAT TCA GAG TCC TGC CCT CAT TAT T
<i>usp</i>	F	ATG CTA CTG TTC CCG AGT AGT GTG T
	R	TTA TCT CCT GTA GAA TTT CAT CAT G
<i>orfU1</i>	F	ATG CTA CTG TTC CCG AGT AGT GTG T
	R	AAT GAT CCA CTA AGG TAT CGA G
<i>orfU2</i>	F	ATG CTA CTG TTC CCG AGT AGT GTG T
	R	AGT ACA TTT CTA GCT CCT TA
<i>orfU3</i>	F	ATG CTA CTG TTC CCG AGT AGT GTG T
	R	CAA GCA ACT CAA TGA ATT

Results

PCR for single-gene and triplex PCR

PCR products for *usp*, *α-hly*, and *cnf-1* genes were 1,041, 1,170, and 489 bp, respectively. The results of single-gene PCR for the *usp* gene are shown in Figure 1A, and the results of triplex PCR for all the three genes of two isolates are shown in Figure 1B. Among the isolates, 127 were *usp* positive, 31 were *cnf-1* positive, and 44 were *α-hly* positive. Among the isolates, 7 did not carry any of the genes, while 16 isolates were found to be positive for all 3 genes.

Result of PCR to investigate the sequential pattern of orfU1, orfU2, and orfU3

All *usp* gene-positive isolates were investigated for the *orfU* sequential position, as this can help in molecular epidemiological analysis of UPEC. Depending on the *orfU* sequence, the first one will have the smallest DNA fragment, the second one will be in the middle position, and the third one will have the largest fragment size. Eighty-five of the isolates included in the study had an *orfU2, orfU3, orfU1* sequential pattern, and 22 isolates had an *orfU1, orfU2, orfU3* pattern. Twenty isolates with two *orfUs* were observed; half were *orfU1, orfU3* and the rest were *orfU2, orfU3*. The results of these experiments for PCR product in gel electrophoresis analysis are shown in Figure 2. The comparison of percentages of PAI_{usp} subtypes in this study and the previous study by Kanamaru et al. are shown in Figure 3.

Discussion

In this study, the *usp* gene encoding one of the VFs of UPEC was studied using the PCR method. A previous study by Kanamaru et al. showed that UTI isolates were highly positive for *usp* and that out of 332 *usp*-positive strains, 324 (97.6%) could be classified into PAI_{usp} subtypes Ia, Ib, IIa, and IIb. Among these subtypes, type IIa was the most common subtype and type Ia was the second-most common type in their study [1].

According to Kanamaru et al., the *usp* gene was mainly classified into two types (*uspI* and *uspII*) depending on the difference in DNA sequence at the 3' 230 bp region before the stop codon. Depending on the presence of *orfU1, orfU2, orfU3* genes in it downstream and the sequential position of these three *orfUs*, it was further classified into four PAI_{usp} subtypes as mentioned above. While some isolates have three *orfU* genes, some have two *orfU* genes. The predominance of the *usp* gene in UPEC isolates and the presence of these PAI_{usp} subtypes indicated that the *usp* gene is a

Figure 2. Gel electrophoresis picture showing the sequential pattern of orfUs for 5 isolates. Lanes 2,3,4 show isolate with *orfU* sequence 1,2,3; lanes 5,6,7 show *orfU* sequence 2,3,1; lanes 8,9,10 show *orfU* sequence with 2,3 patterns; lanes 11,12,13 show 2,3,1; lane 14,15,16 show 1,3 pattern. Lanes 1 and 17 were 100 bp molecular markers. The size of 2,500 bp and 1,800 bp were estimated by other gel using 1 kb molecular marker (figure not shown).

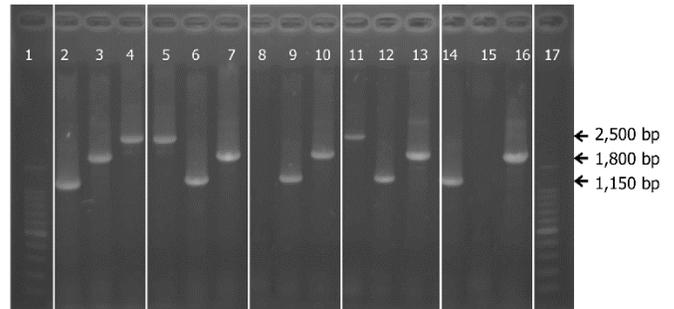


Figure 3. Diagram showing 4 PAI_{usp} subtypes classified according to orfU sequential patterns and their percentage of prevalence in this study. *Usp* gene types were determined by *orfU* sequential patterns. The prevalence of each PAI_{usp} subtypes in this study were compared with that reported previously by Kanamaru et al. [1].

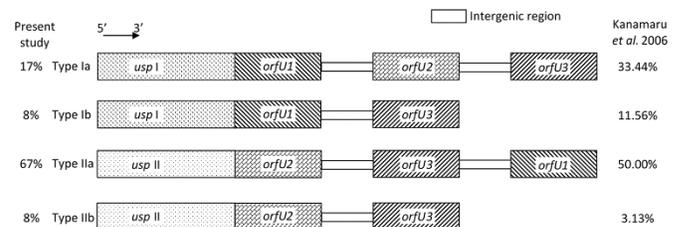
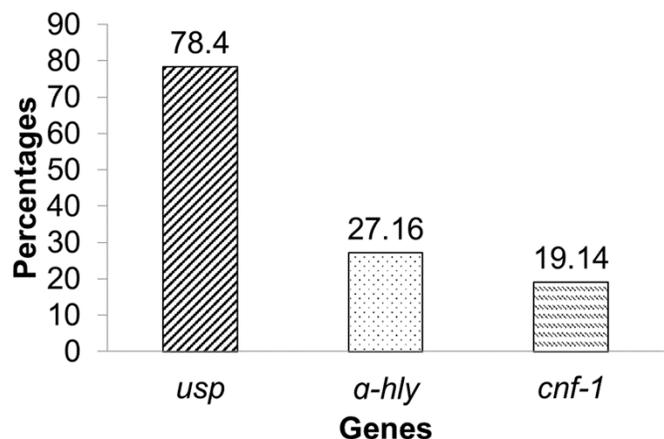


Figure 4. Histogram showing prevalence of *usp* gene and *α-hly* and *cnf-1* genes among UPEC isolates.



potential VF to be used as a molecular epidemiological marker for such isolates [1].

Most of the bacteriocins have immunity protein in their downstream region. These immunity proteins prevent the host bacteria from the suicidal effects of the bacteriocin. These proteins protect the effect of nuclease-bacteriocin on their own cell. Most bacteriocins have only one immunity protein, and this protein is enough to protect the host cell from the effect of bacteriocin. USP is assumed to be a bacteriocin-like nuclease because the DNA sequence of the *usp* gene has high homology with colicins and pyocins [6]. Consistent with this assumption, there are three *orfUs* in its downstream that have high homology among themselves and 40%–47% homology when compared with other immunity proteins. These *orfUs* recombined in a different pattern; the sequential nature of these *orfUs* was different among isolates. In one study, co-expression of the *usp* gene and only the *orfU1* gene was successful [7,8]. This led to the information that *orfU2* and *orfU3* are not essential to be present in the downstream region of the *usp* gene. The *orfU* at the immediate downstream of the *usp* gene is the probable important gene encoding immunity protein, which must be expressed together with the *usp* gene.

Though all three *orfUs* are not important for the expression of the *usp* gene, the three *orfU* sequential patterns are very useful as a PAI*usp* subtyping method. In comparison to *α-hly* and *cnf-1* genes, the *usp* gene was positive at a rate of up to 80% in this study (Figure 4). Kanamaru *et al.* also observed that the *usp* gene was found in 84.9% of their UPEC strains [1]. Taken together, these data show that this gene can characterize a higher percentage of UPEC isolates for epidemiology studies.

The distinct epidemiological features found in this study were that the prevalence of the *usp* gene was comparable to that of previous study in Japan, the *uspII* gene type was more common than in the previous study, PAI*usp* IIa was the most common subtype in both studies and this subtype was the most prevalent subtype in our study, and all the *orfU* sequential patterns found in the study in Japan were observed in this study.

However, in this study, UPEC isolates were classified into four PAI*usp* subtypes from the molecular level. For the discriminatory power to be stronger from the epidemiological point of view, it will be necessary to use it in combination with other typing methods such as plasmid profiles, virulence gene typing, antimicrobial resistance testing, multi-locus sequence typing, *fimH* single-nucleotide polymorphism typing, and phylogenetic grouping [14–18]. When there is

change in the subtype of the PAI*usp* in the locality, this will trigger the attention of the public health officer to trace the source of infections of this new subtype. Consequently, this will benefit the local health authorities by helping in the control of UTI outbreaks.

The novelty of this study is the finding that PAI*usp* subtyping can be applicable in the UPEC isolates of Sabah, Malaysia, supporting the data reported by Kanamaru *et al.* Researchers can use this method for the molecular epidemiological study of UPEC isolates elsewhere in the world where the *usp* gene is as highly prevalent as in these two studies.

Conclusions

This study evaluated PAI*usp* subtypes for characterization of UPEC strains isolated from two hospitals located around Kota Kinabalu, Sabah, East Malaysia. The prevalence rate of PAI*usp* subtypes were similar in two geographical areas of Asia, and the existence of the same subtypes indicates that this typing method will be valuable in further characterization of UPEC isolates in any region of the world.

Acknowledgements

We would like to thank Professor Dr. D Kamarudin D Mudin, Dean, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, for the continuous support throughout the whole research project. This study was supported by grant 2012-RMRU13-001, Rural Medicine Research Unit, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah. We would like to acknowledge Biotechnology Research Institute, Universiti Malaysia Sabah for allowing us to use their research equipment. We would like to thank Dr. Ahmed Toha from Hospital Queen Elizabeth 1 and Mr. Morris Hilarian from Hospital Papar for allowing us to collect the urine samples from their respective hospitals.

Authors' contributions

ZL conceived and designed the experiment. YML performed the experiments ZL and YML analyzed the data. ZL and MTZ wrote the paper. MTZ edited the paper. SBS encouraged the research.

References

1. Kanamaru S, Kurazono H, Nakano M, Terai A, Ogawa O, Yamamoto S (2006) Subtyping of uropathogenic *Escherichia coli* according to the pathogenicity island encoding uropathogenic-specific protein: comparison with phylogenetic groups. *Int J Urol* 13: 754-760.
2. Yamamoto S (2007) Molecular epidemiology of uropathogenic *Escherichia coli*. *J Infect Chemother* 13: 68-73.
3. Kurazono H, Yamamoto S, Nakano M, Nair GB, Terai A, Chaicumpa W, Hayashi H (2000) Characterization of a putative virulence island in the chromosome of uropathogenic

- Escherichia coli* possessing a gene encoding a uropathogenic-specific protein. Microb Pathog 28: 183-189.
4. Nakano M, Yamamoto S, Terai A, Ogawa O, Makino SI, Hayashi H, Nair GB, Kurazono H (2001) Structural and sequence diversity of the pathogenicity island of uropathogenic *Escherichia coli* which encodes the USP protein. FEMS Microbiol Lett 205: 71-76.
 5. Yamamoto S, Nakano M, Terai A, Yuri K, Nakata K, Nair GB, Kurazono H, Ogawa O (2001) The presence of the virulence island containing the *usp* gene in uropathogenic *Escherichia coli* is associated with urinary tract infection in an experimental mouse model. J Urol 165: 1347-1351.
 6. Parret AH, De Mot R (2002) *Escherichia coli*'s uropathogenic-specific protein: a bacteriocin promoting infectivity? Microbiology 148: 1604-1606.
 7. Sano Y, Matsui H, Kobayashi M, Kageyama M (1993) Molecular structures and functions of pyocin S1 and S2 in *Pseudomonas aeruginosa*. J Bacteriol 175: 2907-2916.
 8. Zaw MT, Yamasaki E, Yamamoto S, Nair GB, Kawamoto K, Kurazono H (2013) *Uropathogenic specific protein* gene, highly distributed in extraintestinal uropathogenic *Escherichia coli*, encodes a new member of H-N-H nuclease superfamily. Gut Pathog 5: 13. doi:10.1186/1757-4749-5-13.
 9. Zaw MT, Lai YM, Lin Z (2013) Uropathogenic specific protein: epidemiologic marker of uropathogenic *Escherichia coli* as well as non-specific DNase. Int J Coll Res Int Med Public Health 5: 630-640.
 10. Keane WF, Welch R, Gekker G, Peterson PK (1987) Mechanism of *Escherichia coli* alpha-hemolysin-induced injury to isolated renal tubular cells. Am J Pathol 126: 350-357.
 11. De Rycke J, Guillot JF, Boivin R (1987) Cytotoxins in non-enterotoxigenic strains of *Escherichia coli* isolated from feces of diarrheic calves. Vet Microbiol 15: 137-150
 12. Davis JM, Rasmussen SB, O'Brien AD (2005) Cytotoxic necrotizing factor type 1 production by uropathogenic *Escherichia coli* modulates polymorphonuclear leukocyte function. Infect Immun 73: 5301-5310.
 13. Mills M, Meysick KC, O'Brien AD (2000) Cytotoxic necrotizing factor type 1 of uropathogenic *Escherichia coli* kills cultured human uroepithelial 5637 cells by an apoptotic mechanism. Infect Immun 68: 5869-5880.
 14. Farshad S, Ranjbar R, Japoni A, Hosseini M, Anvarinejad M, Mohammadzadegan R (2012) Microbial susceptibility, virulence factors, and plasmid profiles of uropathogenic *Escherichia coli* strains isolated from children in Jahrom, Iran. Arch Iran Med 15: 312-316.
 15. Gundogdu A, Long YB, Vollmerhausen TL, Katouli M (2011) Antimicrobial resistance and distribution of *sul* genes and integron-associated *intI* genes among uropathogenic *Escherichia coli* in Queensland, Australia. J Med Microbiol 60: 1633-1642.
 16. Bidet P, Mahjoub-Messai F, Blanco J, Blanco J, Dehem M, Aujard Y, Bingen E, Bonacorsi S (2007) Combined multilocus sequence typing and O serogrouping distinguishes *Escherichia coli* subtypes associated with infant urosepsis and/or meningitis. J Infect Dis 196: 297-303.
 17. Dias RCS, Moreira BM, Riley LW (2010) Use of *fimH* Single-Nucleotide Polymorphisms for strain typing of clinical isolates of *Escherichia coli* for epidemiologic investigation. J Clin Microb 48: 483-488.
 18. Clermont O, Bonacorsi S, Bingen E (2000) Rapid and simple determination of the *Escherichia coli* phylogenetic group. Appl Environ Microbiol 66: 4555-4558.

Corresponding author

Zaw Lin, MD, PhD

Pathobiological and Medical Diagnostics Department, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Jalan UMS, Kota Kinabalu

88400, Sabah, Malaysia

Phone: 088-320000 Ext: 611428

Fax: 088-321373

Email: 56dr.zawlin@gmail.com

Conflict of interests: No conflict of interests is declared.