Evaluation of PAISP subtyping to characterize uropathogenic *E. coli* isolates

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

¹Department of Pathobiological and Medical Diagnostics, Faculty of Medicine and Health Sciences, Universiti Malaysia Sabah, Kota Kinabalu, Malaysia
²Department 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 PAISP 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 VF genes of UPEC, namely alpha hemolysin (*a-hly*) and cytotoxic necrotizing factor (*cnf-1*) 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 PAISP 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 *a-hly* and *cnf-1* genes.

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

Key words: *usp* gene; uropathogenic *Escherichia coli*; PAISP subtyping; characterization.


(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 *uspl* 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 µL 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 µL of overnight broth culture and 300 µL 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.

PCRs were 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 was 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.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer</th>
<th>Primer sequence (5’–3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>α-hly</td>
<td>F</td>
<td>AAC AAG GAT AAG CAC TGT TCT GGC T</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>ACC ATA TAA GCG GTC ATT CCC GTC A</td>
</tr>
<tr>
<td>cnf-1</td>
<td>F</td>
<td>AAG ATG GAG TTT CCT ATG CAG GAG</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>CAT TCA GAG TCC TGC CCT CAT TAT T</td>
</tr>
<tr>
<td>usp</td>
<td>F</td>
<td>ATG CTA CTG TTC CCG AGT AGT GTG T</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>TTA TCT CCT GTC AAG TTT CAT CAT G</td>
</tr>
<tr>
<td>orfU1</td>
<td>F</td>
<td>ATG CTA CTG TTC CCG AGT AGT GTG T</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>AAT GAT CCA CTA AGG TAT CGA G</td>
</tr>
<tr>
<td>orfU2</td>
<td>F</td>
<td>ATG CTA CTG TTC CCG AGT AGT GTG T</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>AGT ACA TTT CTA GCT CCT TA</td>
</tr>
<tr>
<td>orfU3</td>
<td>F</td>
<td>ATG CTA CTG TTC CCG AGT AGT GTG T</td>
</tr>
<tr>
<td></td>
<td>R</td>
<td>CAA GCA ACT CAA TGA ATT</td>
</tr>
</tbody>
</table>
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 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 VF s 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, Ila, and IIB. Among these subtypes, type Ila 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
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 PAIusp subtyping method. In comparison to a-hly and cnf-I 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, PAIusp 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 PAIusp 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 PAIusp 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 PAIusp 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 PAIusp subtypes for characterization of UPEC strains isolated from two hospitals located around Kota Kinabalu, Sabah, East Malaysia. The prevalence rate of PAIusp 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


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.