JCM
Journals.ASM.org

Identification of *Escherichia coli* Genes Associated with Urinary Tract Infections

Bin-Hsu Mao,[a] Yung-Fu Chang,[b] Joy Scaria,[b] Chih-Ching Chang,[c] Li-Wei Chou,[d] Ni Tien,[e] Jiunn-Jong Wu,[f,j] Chin-Chung Tseng,[g] Ming-Cheng Wang,[g] Chao-Chin Chang,[h] Yuan-Man Hsu,[i] and Ching-Hao Teng[a,j]

Institute of Molecular Medicine, National Cheng Kung University Medical College, Tainan City, Taiwan[a]; Department of Population Medicine and Diagnostic Sciences, College of Veterinary Medicine, Cornell University, Ithaca, New York, USA[b]; Department of Environmental and Occupational Health, National Cheng Kung University Medical College, Tainan City, Taiwan[c]; Department of Physical Medicine and Rehabilitation, China Medical University Hospital, Taichung City, Taiwan[d]; Department of Laboratory Medicine, China Medical University Hospital, Taichung City, Taiwan[e]; Department of Medical Laboratory Science and Biotechnology, National Cheng Kung University Medical College, Tainan City, Taiwan[f]; Division of Nephrology, Department of Internal Medicine, National Cheng Kung University Hospital, Tainan City, Taiwan[g]; Graduate Institute of Microbiology and Public Health, School of Veterinary Medicine, National Chung Hsing University, Taichung City, Taiwan[h]; Department of Biological Science and Technology, China Medical University, Taichung City, Taiwan[i]; and Center of Infectious Disease and Signaling Research, National Cheng Kung University, Tainan City, Taiwan[j]

***Escherichia coli* is the most common cause of urinary tract infections (UTIs). *E. coli* genes epidemiologically associated with UTIs are potentially valuable in developing strategies for treating and/or preventing such infections as well as differentiating uropathogenic *E. coli* from nonuropathogenic *E. coli*. To identify *E. coli* genes associated with UTIs in humans, we combined microarray-based and PCR-based analyses to investigate different *E. coli* source groups derived from feces of healthy humans and from patients with cystitis, pyelonephritis, or urosepsis. The *cjrABC-senB* gene cluster, *sivH*, *sisA*, *sisB*, *eco274*, and *fbpB*, were identified to be associated with UTIs. Of these, *cjrABC-senB*, *sisA*, *sisB*, and *fbpB* are known to be involved in urovirulence in the mouse model of ascending UTI. Our results provide evidence to support their roles as urovirulence factors in human UTIs. In addition, the newly identified UTI-associated genes were mainly found in members of phylogenetic groups B2 and/or D.**

Escherichia coli is the most common cause of urinary tract infections (UTIs), including acute cystitis, pyelonephritis, and urosepsis, three common and clinically distinct UTI syndromes. It is widely accepted that uropathogenic *E. coli* (UPEC) originates from the distal gut microbiota (8, 13, 15). To cause ascending UTI, UPEC needs to overcome and adapt to different distinct host environments, such as the bladder, the kidneys, and even the bloodstream. Accordingly, UPEC tends to be distinct from the commensal *E. coli* strains in the intestinal tract in having extra virulence genes, allowing their successful transition from the intestinal tract to the urinary tract.

An epidemiological association between an *E. coli* gene and UTIs may suggest that the gene itself encodes a factor contributing to urovirulence or has a genetic linkage to such a gene. Therefore, the genes associated with UTIs are potentially valuable in differentiating UPEC from nonuropathogenic *E. coli* and in the development of strategies for managing and preventing this particular type of disease.

Palaniappan et al. have developed an oligonucleotide-spotted microarray containing probes representing 342 *E. coli* genes to differentiate *E. coli* pathotypes (27). A majority of the genes are derived from the UPEC strain CFT073, the enterohemorrhagic *E. coli* strain EDL933, and the commensal *E. coli* K-12 strain MG1655. The remaining genes are derived from other *E. coli* strains capable of causing intestinal infections. The association of the majority of the 342 genes in the array with UTIs has not yet been investigated, except for 36 uropathogenic genes included among them (27).

To identify *E. coli* genes associated with UTIs, we used the microarray developed by Palaniappan et al. to screen for the genes potentially associated with UTIs and then performed a PCR-based analysis with a larger bacterial sample size to confirm these genes' epidemiologic associations with UTIs. One gene cluster and five individual genes (here the gene cluster and five individual genes are referred to as MIGs [microarray-identified genes]) were associated with UTIs. Of these, the gene cluster and three of the individual genes have recently been shown to be involved in urovirulence in the mouse model of UTI (6, 20, 22). In addition, we analyzed the phylogenetic distribution of the MIGs and assessed the correlations between these MIGs, as well as between these genes and 15 known virulence genes.

MATERIALS AND METHODS

***E. coli* isolates and patients.** The UTI-associated isolates in this study, cystitis, pyelonephritis, and urosepsis isolates, were collected from two hospitals in Taiwan: the China Medical University Hospital (CMUH) at Taichung City in central Taiwan and National Cheng Kung University Hospital (NCKUH) at Tainan City in southern Taiwan. The UTI-associated isolates used in this study were a subset of the 2,206 *E. coli* strains isolated from the urine specimens submitted to the diagnostic laboratories of the two hospitals between June 2006 and April 2007 (1,619 isolates from CMUH and 587 isolates from NCKUH). According to the diagnostic criteria mentioned below, among the 1,619 isolates from CMUH, 696 isolates were from cystitis patients, 421 isolates from pyelonephritis patients, and 372 from urosepsis patients. Among the 587 isolates from NCKUH, 229 isolates were from cystitis patients, 141 isolates from pyelonephritis patients, and 94 isolates from urosepsis patients. To-

Received 31 March 2011 Returned for modification 15 May 2011
Accepted 30 October 2011
Published ahead of print 9 November 2011
Address correspondence to Ching-Hao Teng, chteng@mail.ncku.edu.tw.
Supplemental material for this article may be found at http://jcm.asm.org/.

doi:10.1128/JCM.00640-11

gether, there were 925 cystitis isolates, 562 pyelonephritis isolates, and 466 urosepsis isolates. Out of each type of UTI isolates, we randomly selected 67 (7.2%) cystitis isolates, 72 (12.8%) pyelonephritis isolates, and 64 (13.8%) urosepsis isolates for this study. The biliary tract infection (BTI)-associated bacteremia *E. coli* strains ($n = 24$) were obtained from the blood specimens of BTI patients with bacteremia at NCKUH between September 2004 and November 2007. In addition, 115 commensal fecal isolates were collected from the feces of healthy donors between June 2006 and April 2007. Each bacterial isolate in this study was derived from a different patient or healthy donor.

According to the diagnostic criteria of UTIs previously described (35), the prerequisite for patients with UTIs was that their fresh urine samples contained bacterial counts of $\geq 10^5$ CFU/ml. Cystitis was defined by the presence of dysuria, urinary frequency, and/or lower abdominal pain. Pyelonephritis was based on the presence of body temperature ≥38.3°C, flank pain, and/or costovertebral angle tenderness, with or without the syndrome of cystitis. Urosepsis was defined by the presence of bacteremia in addition to UTI syndromes.

The diagnostic criteria for BTI were fever, abdominal pain in the right upper quadrant, and/or jaundice, with imaging demonstrating the presence of acute cholecystitis or acute cholangitis (37).

DNA microarray analysis. For genomic DNA preparation, a single colony of bacteria was inoculated into LB broth and incubated at 37°C for 12 h to a concentration of around 3×10^9 CFU/ml. Four milliliters of the overnight bacterial culture was then subjected to genomic DNA extraction with a Qiagen DNeasy kit (Qiagen, Valencia, CA) according to manufacturer's instructions. The integrity of the genomic DNA was verified by 1% agarose gel electrophoresis stained with ethidium bromide. Two micrograms of the genomic DNA was used for the enzyme digestion and labeling processes and then concentrated to a volume of 12 μl for the following hybridization with the microarray. Microarray printing, bacterial genomic DNA labeling, microarray hybridization, and data acquisition were performed as previously described (27). Each slide had triplicate spots of each gene. Similarly, the derived data were analyzed as previously described but with a modification in the cutoff criteria, determining the positive signals. For each gene, the background-subtracted median fluorescence intensities of triplicate spots were averaged and $\log_2$ transformed, designated *L* values. A gene was considered positive when its *L* value was greater than the value of the mean (*M*) subtracted by the standard deviation (SD) of the *L* values derived from 48 genes commonly present in the *E. coli* strains MG1655, CFT073, and EDL933 (27).

Fifteen of the 342 *E. coli* gene probes were selected as positive controls for the microarray experiments, because the target genes they hybridize with were identified in all the current available *E. coli* strains with complete genome sequences (some of them have been demonstrated as *E. coli* essential genes) (see Table S1 in the supplemental material). Twenty *Salmonella enterica* serovar Typhimurium LT2-specific gene probes and the autoblanks (spots of dimethyl sulfoxide without any probes) were used as the negative controls (see Table S1) (27).

PCR-based genotyping and phylogenetic typing. The frequencies of the genes screened in by the microarray analysis were determined by PCR-based analysis. The primers were designed to target the conserved regions of the MIGs (Table 1). The PCRs were heated to 95°C in an automated thermal cycler for 5 min, followed by 30 cycles of denaturation (95°C, 45 s), annealing (59°C, 45 s), and extension (72°C, 50 s). *Taq* polymerase was used in the reactions. The phylogenetic groups of the 342 *E. coli* isolates were determined based on the PCR-based method described previously (5). The frequencies of the 15 selected known virulence genes of extraintestinal pathogenic *E. coli* were determined by PCR-based assays, using primers and PCR conditions as described previously (3, 16).

The following *E. coli* strains were used as controls for the PCR-based analysis of the MIGs and the known virulence genes. The *E. coli* strains which served as positive controls included CFT073 (*sivH*, *shiA*, *sisA*, *sisB*, *fbpB*, *papGII*, *chuA*, *ompT*, *sat*, *iha*, *usp*, *ireA*, *iroN*, and *hlyA*), UTI89 (*cjrA*, *cjrB*, *cjrC*, *senB*, *cnf1*, *sfaS*, and *ibeA*), EDL933 (*eco274*), J96 (*papGI* and *papGIII*), and one of the UTI-associated clinical isolates, A53, which was identified to harbor *afa/dra* by sequencing the PCR product amplified by the *afa/dra* specific primers (*afa/dra*). MG1655 was used as a negative control for all the genes except *ompT*. An *ompT* deletion mutant of *E. coli* strain RS218 was constructed and served as the negative control of *ompT*.

TABLE 1 Primer sequences used in this study

Gene	Sequence (5′–3′)	Amplicon size (bp)
cjrA	AAAGGGTGGTCCTGGGAGAT ACGTCAGTTGCTGGCTTTCA	223
cjrB	CGAAGTTCAGCCCGCTATGT GCTTTCCCAAGATGCCTCAG	397
cjrC	AAACCTCAGCGCAAAATCGT AGGCTTCAGGAATGGGTTCA	518
senB	CCGTTGAAAGATCCGAGACC GTTTGGGTAGACCGGCATGT	312
sivH	TACAGCACGCGTAAACCGTA TGGCAGTACAGTTCCGATCA	866
shiA	TCACCTTACTGGTATGAACTC TCCAGGGCCAGACATATTCA	451
sisA	TTGCCCGACAGGAGAATGAC GCAGTATATGGCGTGCCTGT	360
sisB	GAACGATAGATTATGCTTTG TCAGTACACTGAAGGCTCGC	518
eco274	TTGACAAAGCCTGCCTGACC CCTCCAACCCGTGTTTTTGC	207
fbpB	GCAAATCGCGCAGGATAAAG ACGCACAAGGAGGTGCGTAT	821

The PCR amplification was done in a 25-μl reaction mixture. Amplifications were carried out in Eppendorf Mastercycler gradient thermal cycler (Eppendorf, Hamburg, Germany). The PCR products were electrophoresed in 1.5% agarose gels, stained with ethidium bromide, and photographed using an AlphaImager HP system (Cell Biosciences, Inc., Santa Clara, CA). The sizes of the products were determined by comparing them with a 100-bp DNA ladder (Fermentas Inc., Glen Burnie, MD) run on the same gel. All PCR tests were performed 3 times with independently prepared boiled lysates. Additional investigations were further conducted, if discrepancies between the independent assays occurred.

Statistical analysis. Comparisons involving the frequencies of a given gene in different groups were measured by using two-tailed Fisher's exact test. A *P* value of <0.05 was arbitrarily set as the threshold for statistical significance. Correlations between genes were measured by using Fisher's exact test (two-tailed). Because of multiple comparisons, a *P* value of <0.01 was arbitrarily set as the threshold for statistical significance, with a *P* value of <0.05 as the borderline statistical significance (12, 14).

RESULTS

***E. coli* genes potentially associated with UTIs.** To screen for *E. coli* genes that are potentially associated with UTIs, a microarray-based pilot study was conducted by using 40 *E. coli* isolates which were divided into 4 source groups according to the clinical syndromes (or conditions) with which they were associated: fecal isolates from healthy humans ($n = 8$), cystitis isolates ($n = 12$), pyelonephritis isolates ($n = 10$), and urosepsis isolates ($n = 10$)

TABLE 2 The microarray-analysis-derived frequencies of the genes which were potentially associated with UTIs

Gene	No. (%) of *E. coli* isolates			
	Fecal isolates ($n = 8$)	Cystitis isolates ($n = 12$)	Pyelonephritis isolates ($n = 10$)	Urosepsis isolates ($n = 10$)
cjrA	2 (25)	5 (42)	6 (60)	7 (70)
cjrB	2 (25)	5 (42)	6 (60)	7 (70)
cjrC	2 (25)	5 (42)	6 (60)	7 (70)
senB	2 (25)	5 (42)	6 (60)	7 (70)
sivH	2 (25)	5 (42)	8 (80)[a]	7 (70)
shiA	4 (50)	8 (67)	9 (90)	9 (90)
eco274	2 (25)	8 (67)	7 (70)	6 (60)
fbpB	4 (50)	8 (67)	8 (80)	8 (80)

[a] $P < 0.05$, pairwise comparisons between the indicated UTI-associated source group with the fecal source group.

(Table 2). The bacterial isolates were subjected to gene profiling using the DNA microarray developed by Palaniappan et al. (27). With the gene profiles derived from the microarray data, we determined the frequencies of each microarray-detectable gene in the 4 source groups of *E. coli* (see Table S1 in the supplemental material). Based on the results of this pilot study, eight genes (*cjrA*, *cjrB*, *cjrC*, *SenB*, *sivH*, *shiA*, *eco274*, and *fbpB*) were selected for further study, since they exhibited higher frequencies in the UTI-associated source groups than those in the fecal source group. While these differences in frequencies were consistently observed between UTI-associated and fecal source groups, only the *sivH* in the pyelonephritis isolates reached statistical significance (Table 2). Based on BLAST and literature searches, their potential functions were predicted (Table 3). Among them, *cjrA*, *cjrB*, *and cjrC* are located in the operon *cjrABC*, and *senB* is located downstream of the operon with its 5′ end partially overlapping with *cjrC* (32). Thus, *cjrABC* and *senB* are referred to as *cjrABC-senB* in this study. According to recent studies using the mouse model of ascending UTI, *cjrABC-senB*, *shiA*, and *fbpB* are involved in the virulence of UPEC (6, 20, 22). However, this is the first time their association with human UTIs has been studied.

PCR confirmed associations between the MIGs and UTIs. To further confirm whether the genes identified from the microarray-based pilot study are associated with UTIs, a PCR-based analysis with a larger sample size was performed. Here, another source group containing *E. coli* isolates causing BTI-associated bacteremia (BTI-associated isolates) was included in addition to the original 4 source groups. The 5 source groups contained a total of 342 *E. coli* isolates (Table 4). We compared the distributions of 6 known virulence genes (*papGII*, *cnfI*, *hlyA*, *chuA*, *iroN*, and *usp*) in the UTI-associated source groups with those in the fecal isolates, all the virulence genes exhibited significantly higher frequencies in the UTI-associated source groups than the fecal isolates (see Table S2 in the supplemental material).

According to the PCR analysis with all the *E. coli* isolates, 107 isolates had at least 1 gene of *cjrABC-senB*. A total of 104 out of the 107 isolates (97%) contained all 4 genes, suggesting that their coexistence is common in *E. coli*. Therefore, we investigated the distribution of the intact gene cluster, *cjrABC-senB*, by detecting the 4 genes separately with different primer pairs specific to each gene (Table 1). Two *shiA* homologs, *sisA* and *sisB*, are identified in *E. coli* (22), but the microarray used in this study could not differentiate between them. SisA and SisB share 86% identity at the level of amino acid sequences, with their N termini being the most divergent parts. The distributions of *shiA*, *sisA*, and *sisB* were investigated separately by using *shiA* primers, able to detect the sequence common to *sisA* and *sisB*, and primers specific to *sisA* and *sisB*, in the PCR-based analysis (Table 1).

Overall, the results confirmed that these MIGs are associated with UTIs. The frequencies of most of the MIGs in each of the UTI-associated source groups (cystitis, pyelonephritis, and urosepsis) were significantly higher than those in the fecal source group (Table 4). Although the frequencies of *eco274* in the cystitis and fecal isolates were not significantly different, its frequencies in the pyelonephritis and urosepsis isolates were significantly higher than those in the fecal isolates.

When the three UTI-associated source groups were compared, *shiA* and *sisA* showed significantly higher frequencies in the pyelonephritis isolates than in the cystitis isolates (Table 4). When the UTI-associated groups were compared with the BTI-associated group, the distributions of *cjrABC-senB*, *sivH*, *shiA*, *sisA*, and *fbpB* markedly favored the UTI-associated bacterial isolates (Table 4). *sisB* tended to exhibit higher frequencies in the UTI-associated isolates than in the BTI-associated isolates, although the differences did not reach statistical significance. In addition, when the

TABLE 3 The potential functions of the MIGs

Gene	Designation in microarray[a]	Potential function of the gene product[b]	Accession no. in the representative UPEC strains[c]	Completely genome-sequenced UPEC strains harboring the genes
cjrA	*cjrA*	Putative inner membrane protein	YP_538626	UTI89, UMN026
cjrB	*cjrB*	TonB-like protein	YP_538627	UTI89, UMN026
cjrC	*cjrC*	Putative TonB-dependent receptor	YP_538628	UTI89, UMN026
senB	*senB*	Enterotoxin Tie protein	YP_538629	UTI89, UMN026
sivH[d]	*eco293*	Putative intimin or invasin protein	NP_754913	UTI89, CFT073, UMN026, 536, IAI39
shiA	*eco294*	Potential suppressor of innate immune response	NP_755432 and NP_756354	CFT073, UMN026, IAI39,
eco274	*eco274*	Potential transcriptional regulator	YP_002410135	UMN026, IAI39
fbpB	*eco288*	Potential iron-chelating protein	NP_752239	UTI89, CFT073, 536

[a] The gene's designation used in the microarray described previously (27).
[b] The potential functions of all the genes are based on the BLAST search, except for that of *shiA*, which is based on the finding of Lloyd et al. (22).
[c] The accession no. indicates UTI89-derived *cjrA*, *cjrB*, *cjrC*, and *senB*; CFT073-derived *shiA*, *sivH*, and *fbpB*; IAI39-derived *eco274*, respectively.
[d] *sivH* is also named *sinH*.

TABLE 4 Distributions of the MIGs among 342 *E. coli* isolates in different source groups

Gene	No. (%) of *E. coli* isolates					*P* value[a]								
	Fecal isolates ($n = 115$)	Cystitis isolates ($n = 67$)	Pyelonephritis isolates ($n = 72$)	Urosepsis isolates ($n = 64$)	BTI isolates[b] ($n = 24$)	Fecal vs cystitis	Fecal vs pyelonephritis	Fecal vs urosepsis	Fecal vs BTI	Cystitis vs pyelonephritis	Cystitis vs urosepsis	Pyelonephritis vs urosepsis	BTI vs cystitis	BTI vs pyelonephritis	BTI vs urosepsis
cjrABC-senB	23 (20)	24 (36)	32 (44)	23 (36)	2 (9)	0.023	0.001	0.031	–	–	–	–	0.016	0.001	0.015
sivH	21 (18)	36 (54)	37 (51)	30 (47)	4 (13)	<0.001	<0.001	<0.001	–	–	–	–	0.002	0.004	0.013
shiA	40 (35)	47 (70)	62 (86)	54 (84)	6 (25)	<0.001	<0.001	<0.001	–	0.025	–	–	<0.001	<0.001	<0.001
sisA	35 (30)	45 (67)	61 (85)	51 (80)	5 (22)	<0.001	<0.001	<0.001	–	0.017	–	–	<0.001	<0.001	<0.001
sisB	10 (9)	19(28)	16(22)	20 (31)	3 (13)	0.001	0.016	<0.001	–	–	–	–	–	–	–
eco274	31 (27)	22 (33)	31 (43)	30 (47)	7(38)	–	0.026	0.009	–	–	–	–	–	–	–
fbpB	18 (16)	30 (45)	32 (44)	32 (50)	3 (9)	<0.001	<0.001	<0.001	–	–	–	–	0.006	0.006	0.001

[a] Only *P* values of <0.05 (by Fisher's exact test) are shown. –, *P* value of ≥0.05.

[b] "BTI isolates" indicates BTI-associated bacteremia isolates.

BTI-associated isolates were compared with the fecal isolates, the frequencies of the MIGs were not significantly different (data not shown).

Phylogenetic distribution of the MIGs. The MIGs were mainly concentrated within phylogenetic groups B2 and/or D (Table 5). *cjrABC-senB* and *eco274* showed significantly greater frequencies in group D than in group B2, while *sivH* and *fbpB* were present with significantly higher frequencies in group B2 than in group D. *sivH* and *fbpB* were almost entirely confined to group B2 and were not found in group A or B1. The frequencies of *shiA*, *sisA*, and *sisB* in group B2 and in group D isolates were not significantly different.

Stratification of the MIGs by phylogeny. To determine whether the associations of the MIGs with UTIs were still present in individual phylogenetic groups, we further evaluated the distributions of the genes in the fecal and UTI-associated isolates with stratification of individual phylogenetic groups. Since the genes were mainly concentrated in groups B2 and/or D, and were relatively rare in groups A and B1 (Table 6), only the group B2 and group D strains were assessed.

In group B2, the distributions of *shiA*, *sisA*, and *sisB* still markedly favored the UTI-associated isolates over the fecal isolates (Table 6). The frequencies of *sivH* in the cystitis and pyelonephritis isolates of B2 bacteria were higher than those in the B2 fecal isolates, while the frequencies of *fbpB* in the cystitis and urosepsis isolates of B2 bacteria were higher than those in the B2 fecal isolates (Table 6).

In group D, when the UTI-associated source groups were compared with the fecal source group, a significant difference in frequencies favoring the cystitis and pyelonephritis isolates was still detected for *cjrABC-senB*, *shiA*, and *sisA*.

Correlations of the MIGs with one another and with known virulence genes. We performed pairwise comparisons of the MIGs with one another and with the 15 selected known virulence genes among the 203 UTI-associated *E. coli* isolates (the total isolates in the source groups of pyelonephritis, cystitis, and urosepsis) (Table 7).

When the MIGs were compared with one another, both positive and negative associations were detected. The positive associations seemed to be related to the phylogenetic distribution of the *E. coli* strains which carried these genes. *sivH* and *fbpB*, which were mainly concentrated in group B2 (Table 5), were positively associated with each other. *cjrABC-senB* and *eco274*, whose frequencies in group D were significantly higher than those in group B2 (Table 5), demonstrated a positive association. *sisA* and *sisB*, whose frequencies in groups B2 and D were not significantly different (Table 5), exhibited a marginal positive statistical association ($P = 0.014$) (data not shown). In addition, negative associations were detected when *sivH* was compared with *cjrABC-senB* and *eco274*.

As to the associations between MIGs with the 15 known virulence genes, positive and negative associations were also detected (Table 7). Of note, the association patterns of *sivH* and *eco274* with some of the known virulence genes were opposite. *sivH* was positively associated with *cnf1*, *usp*, *ireA*, *iroN*, *sfaS*, while *eco274* was negatively associated with these genes. In addition, *sivH* was negatively associated with *sat* and *iha*, while *eco274* was positively associated with these genes (Table 7).

TABLE 5 Phylogenetic distribution of the MIGs in the 342 *E. coli* isolates from all the source groups

Gene	No. (%) of *E. coli* isolates				*P* value[a]				
	Group A ($n = 61$)	Group B1 ($n = 15$)	Group B2 ($n = 185$)	Group D ($n = 81$)	A vs B2	A vs D	B1 vs B2	B1 vs D	B2 vs D
cjrABC-senB	6 (10)	0 (0)	60 (32)	38 (47)	<0.001	<0.001	0.006	<0.001	0.028
sivH	0 (0)	0 (0)	126 (68)	2 (2)	<0.001	–	<0.001	–	<0.001
shiA	14 (23)	1 (7)	138 (75)	56 (69)	<0.001	<0.001	<0.001	<0.001	–
sisA	7 (11)	0 (0)	137 (74)	53 (65)	<0.001	<0.001	<0.001	<0.001	–
sisB	8 (13)	1 (7)	40 (22)	19 (23)	–	–	–	–	–
eco274	2 (3)	0 (0)	53 (29)	66 (81)	<0.001	<0.001	0.013	<0.001	<0.001
fbpB	0 (0)	0 (0)	111 (60)	4 (5)	<0.001	–	<0.001	–	<0.001

[a] Only *P* values of <0.05 (by Fisher's exact test) are shown. –, *P* value of ≥0.05.

DISCUSSION

This study is the first to identify the epidemiological associations of the *E. coli* genes *cjrABC-senB*, *sivH*, *sisA*, *sisB*, *eco274*, and *fbpB* with UTIs by utilizing the microarray and PCR-based analyses on fecal isolates and three distinct UTI syndrome-associated isolates. Most of the MIGs were associated with all three distinct UTI syndromes, cystitis, pyelonephritis, and urosepsis, while *eco274* was associated only with pyelonephritis and urosepsis. *cjrABC-senB*, *sisA*, *sisB*, and *fbpB* have recently been found to be involved in the virulence of UPEC in the mouse model of ascending UTI (6, 20, 22). Consistently, our results show these genes' associations with human UTIs, while supporting their roles as *E. coli* virulence factors in human UTIs. As for *eco274* and *sivH*, their roles in urovirulence have not yet been assessed in the mouse model. However, they are likely to be urovirulence genes themselves or have genetic linkage with such genes, based on their associations with UTIs. Thus, these MIGs could be potential targets for developing preventive and/or therapeutic strategies for UTIs as well as potential markers of UPEC.

cjrABC-senB has been shown to be present on the plasmids of enteroinvasive *E. coli* (EIEC) and two UPEC strains, UTI89 and UMN026 (26, 32, 34). Interruption of this gene cluster in UTI89 decreases the bacterium's ability for bladder colonization in the early stage of the mouse model of UTI (6). The *cjrABC* operon is predicted to be involved in iron acquisition, which may contribute to urovirulence (6). *senB* encodes the TieB protein, which may have some role in enterotoxicity of EIEC (26). However, its role in UTIs is not yet clear. *senB* and the gene encoding the ShET-2 toxin, named *senA* (or *sen*), are located on a plasmid of enteroinvasive *E. coli* strain EI37 (26). It has been proposed that the polar effect caused by interruption of *senB* may affect expression of *senA* (26). However, Soto et al. showed that *senA* was not present in all of the 170 UPEC clinical isolates they examined (33); therefore, it seems unlikely that the association of *senB* with UTIs is strictly based on its effect on *senA* expression.

shiA is primarily identified in the SHI-2 pathogenicity island (PAI) of *Shigella flexneri* (25, 36). This gene has been shown to be involved in the downregulation of inflammatory responses in both the rabbit ileal loop and mouse lung models of shigellosis (10, 11). The *shiA* homologs, *sisA* and *sisB*, in the UPEC strain CFT073 have been shown to be involved in suppressing the host immune response, facilitating bacterial colonization

TABLE 6 Prevalence of the MIGs stratified by phylogeny

Gene	No. (%) of *E. coli* isolates[a]				*P* value[b]		
	Fecal isolates ($n = 42$, $n = 25$)	Cystitis isolates ($n = 45$, $n = 13$)	Pyelonephritis isolates ($n = 48$, $n = 20$)	Urosepsis isolates ($n = 45$, $n = 16$)	Cystitis vs fecal	Pyelonephritis vs fecal	Urosepsis vs fecal
Group B2							
cjrABC-senB	10 (24)	14 (31)	18 (38)	71 (38)	–	–	–
sivH	21 (50)	35 (78)	36 (75)	30 (67)	0.008	0.017	–
shiA	19 (45)	34 (76)	45 (94)	39 (87)	0.005	<0.001	<0.001
sisA	19 (45)	33 (73)	45 (94)	39 (87)	0.009	<0.001	<0.001
sisB	1(2)	14 (31)	12 (25)	13 (29)	<0.001	0.002	0.001
eco274	15(36)	10 (22)	12 (25)	15 (33)	–	–	–
fbpB	17 (40)	29 (64)	30 (63)	32 (71)	0.032	–	0.005
Group D							
cjrABC-senB	8 (32)	10 (77)	14 (70)	6 (38)	0.016	0.017	–
sivH	0 (0)	1 (8)	1 (5)	0 (0)	–	–	–
shiA	12 (48)	12 (92)	17 (85)	13 (81)	0.012	0.013	–
sisA	11 (44)	12 (92)	16 (80)	12 (75)	0.005	0.018	–
sisB	4 (16)	4 (31)	4 (20)	5 (31)	–	–	–
eco274	16 (64)	12 (92)	17 (85)	15 (94)	–	–	–
fbpB	1(4)	1 (8)	2 (10)	0 (0)	–	–	–

[a] *n*, total no. of isolates for group B2 and group D, respectively.
[b] Only *P* values of <0.05 (by Fisher's exact test) are shown. –, *P* value of ≥0.05.

TABLE 7 Associations of the MIGs with one another and with the selected known virulence genes in the 231 UTI-associated *E. coli* isolates[a]

Gene	*P* value					
	cjrABC-senB	*eco274*	*sivH*	*fbpB*	*sisA*	*sisB*
Identified						
cjrABC-senB	NA	++	(++)	−	+	−
eco274	++	NA	(++)	−	−	−
sivH	(++)	(++)	NA	++	++	−
fbpB	−	−	++	NA	+	+
sisA	+	−	++	+	NA	−
sisB	−	−	−	+	−	NA
Known						
papG I	−	−	−	−	−	−
papG II	−	−	−	−	++	−
papG III	−	−	−	+	−	−
chuA	++	−	++	++	++	−
ompT	++	++	−	++	−	+
afa/draBC	−	−	−	−	−	−
sat	++	++	(++)	−	++	++
iha	++	++	(++)	−	++	++
cnf1	−	(++)	++	++	−	++
usp	−	(++)	++	++	++	−
ireA	(++)	(++)	++	−	++	−
iroN	(++)	(++)	++	++	−	−
sfaS	−	(+)	++	++	−	−
ibeA	−	−	+	+	(+)	−
hlyA	−	−	+	++	−	++

[a] −, $P \geq 0.01$; +, $P < 0.01$; ++, $P < 0.001$. Because of multiple comparisons, a *P* value of <0.01 was arbitrarily set as the threshold for statistical significance, with a *P* value of <0.05 as borderline statistical significance (12, 14). Parentheses indicate negative associations. "Identified" and "Known" indicate the MIGs and selected known virulence genes in this study, respectively.

of the bladder and kidney during the initial stage of UTI in mice (22).

fbpB is located in the *fbp* locus, which contains the genes *fbpABCD* and is potentially involved in iron acquisition (28). *fbpB* is predicted to encode a periplasmic siderophore-binding protein (28). CFT073 contains two identical copies of *fbpABCD* located in distinct genomic islands, PAI-CFT073-aspV and GI-CFT073-cobU (21). The CFT073 mutant with deletion of the two *fbp* loci is significantly outcompeted by the wild-type strain in the bladders and kidneys of mice (20).

sivH was originally identified in the CS54 island of *Salmonella enterica* serotype Typhimurium (24). The *sivH* of *Salmonella* is known to be involved in the colonization of the Peyer's patches in mice (17). Deletion of the genomic island RDI 13, which contains *sivH*, in the meningitis-associated *E. coli* strain RS218 decreases the bacterium's ability to adhere to and invade human brain microvascular endothelial cells (39). Whether SivH contributes to the urovirulence of UPEC is unknown.

eco274 is classified as an EDL933-specific gene in the DNA microarray used in our initial screening (27). This gene is located in the O island number 148 of EDL933 (29). Its role in urovirulence is unknown. When the three UTI-associated groups were compared with the fecal isolates, it was noted that *eco274* was not associated with cystitis but rather only with pyelonephritis and urosepsis. Thus, *eco274* is likely to encode a virulence factor involved only in pyelonephritis and urosepsis or have a genetic linkage with a gene encoding such a virulence factor.

The distributions of several MIGs in the UPEC strains revealed in this study are consistent with results from other related studies. The frequencies of *cjrABC-senB* in the three UTI-associated groups ranged from 36% to 43% in the present study, which is similar to the findings of Cusumano et al. that *senB* exists in 8 of 18 UPEC isolates (44%) (6). Similarly, the higher frequencies of *sisA* in the UTI-associated groups compared to those of *sisB* in our study (67% to 80% versus 28% to 31%) support the hypothesis of Lloyd et al. (20) that *sisA* is more prevalent than *sisB* in UPEC isolates. Their hypothesis is based on the findings that in all sequenced bacterial species, *sisA* is mainly distributed in extraintestinal pathogenic *E. coli* strains, while *sisB* is mainly distributed in enteric strains (22). Also, our finding that no significant difference in the distributions of *fbpB* between the cystitis and pyelonephritis isolates is consistent with the findings of Parham et al. that the distributions of the *fbp* locus in these two types of isolates are not significantly different (28). However, the *fbpB* frequencies in the cystitis and pyelonephritis isolates in our study (45% and 44%, respectively) are lower than those of the *fbp* locus in the same types of isolates in the study of Parham et al. (58% and 59%, respectively). In addition, Lloyd et al. classified *fbpB* as UPEC specific based on an investigation of 11 UPEC and 4 fecal or commensal *E. coli* strains, showing that *fbpB* is present in all the UPEC strains but not in the fecal/commensal strains (21). However, in our study, *fbpB* was detected in 18 out of the 115 fecal isolates (Table 4), suggesting that *fbpB* is not a UPEC specific gene, although its frequencies in fecal isolates were significantly lower than those in the UTI-associated isolates.

The significantly higher frequency of *shiA* in the pyelonephritis group than that in the cystitis group may be due to the distribution of *sisA*. This is because the frequency of the *shiA* distribution was the composite of the *sisA* and *sisB* distributions, and only *sisA* but not *sisB* exhibited higher frequencies in the pyelonephritis group than in the cystitis group. In addition, the significantly higher frequencies of *sisA* in the pyelonephritis group may suggest that *sisA* plays a more important role in pyelonephritis than in cystitis.

The distributions of the MIGs, which were mainly concentrated in groups B2 and/or D, are similar to those of most extraintestinal virulence genes, concentrated in groups B2 and/or D as well (12, 15). Such accordance is supportive to our assertion that the MIGs are potential virulence genes or have genetic linkage to such genes.

The associations of the MIGs with extraintestinal infections may be syndrome dependent (i.e., BTI versus UTI), because these genes were correlated with the UTI-associated isolates, but not with the BTI-associated isolates, compared with the fecal isolates. Wang et al. have shown that *E. coli* strains responsible for BTI mainly belong to the phylogenetic groups A and B1 (37), unlike the other extraintestinal pathogenic strains, including UPEC, mainly belonging to phylogenetic group B2 and, to a lesser extent, group D (7, 16, 30). Similarly, 50% (12/24) of the BTI-associated isolates in this study belong to phylogenetic group A (data not shown). Given our finding that these MIGs were mainly concentrated in phylogenetic groups B2 and/or D, such distinct phylogenetic distribution of BTI-associated isolates might be responsible for the syndrome-dependent associations. However, the possibility that these genes are specifically associated with UTIs, but not other types of extraintestinal infections, cannot be excluded. A detailed study to assess more types of extraintestinal infections caused by *E. coli* may be required.

The associations of *cjrABC-senB*, *sivH*, and *fbpB* with the UTI syndromes may not necessarily be phylogenetic group dependent, although these associations were observed exclusively in one of the investigated phylogenetic groups, B2 or D (Table 6). These observations may be due to the decreased sample size after stratification by phylogeny.

Two genes having genetic linkage, such as their colocalization in the same plasmid or genomic island, may result in a positive association. However, the MIGs positively associated with each other were not found in the same genomic islands or plasmids, according to the BLAST search on the completely sequenced *E. coli* strains. These observations suggest that such associations are not due to genetic linkages but due to a process of coselection, which may facilitate the pathogenesis in UTIs. As an example, *cjrABC-senB* is located in plasmids of *E. coli* strains (26, 32, 34), while its positively associated genes, *eco274* and *sisA*, are located in the chromosome (34, 38). Also, *sivH* is located in a three-gene genomic island (named RDI 13 in RS218, as mentioned) which is always inserted between *yfgJ* and *xseA* in the chromosome (2, 4, 34, 38), while *sisA* and *fbpB*, which are positively associated with *sivH*, are located in other PAIs (2, 34, 38).

The positive association of *sisA* with the known virulence genes, *papGII*, *iha*, and *sat*, may be due to genetic linkages among them (Table 7), because *sisA* and these known virulence genes are located in a PAI, PAI-CFT073-pheV, in CFT073 (21). However, the associations are not absolute. In addition, *sisA* was not associated with *hly*, which is also located in PAI-CFT073-pheV (21). These observations demonstrate that the genetic linkages between virulence factors in a PAI are not constant, supporting the suggestion of Johnson et al. that virulence genes may be transferred horizontally, independent of the PAIs where they were originally located, in addition to being transferred with the entire PAI (14).

It is known that UPEC requires a combination of multiple virulence genes to cause infection (23). The virulence gene combination of a UPEC strain may determine the pathogenesis process employed by this strain to cause infection. Johnson et al. have identified two groups of urovirulence genes. The member genes in the same group exhibit positive associations but, in general, negatively associate with genes in the other group (14). Also, we found that *eco274* and *sivH*, which were negatively associated with each other, exhibited opposite association patterns with a portion of the known virulence genes (as mentioned in Results). These findings imply that sets of virulence genes, with members that are for the most part discrete, may exist among UPEC strains to direct bacteria through distinct pathways to cause UTIs. However, a further study with more virulence factors to determine their distributions and cooccurrence is necessary to test this inference.

The expression of a bacterial gene in an environment may reflect the role of this gene in bacterial adaptation to this particular environment (31). Thus, transcriptome analyses of UPEC genes during UTIs may provide clues to whether or not and how a bacterial gene is involved in pathogenesis of UTIs. According to a recent microarray analysis of *E. coli* global gene expression in 8 urine samples from different women with UTIs, expression of *sisA* is detected in most of the urine samples (9), consistent with the notion that this gene is a virulence factor in human UTIs. However, the expression of *sisB* and *fbpA*, which is located upstream of *fbpB* in the *fbp* locus, was not detected in these urine samples (9). This may be because *sisB* and the genes in the *fbp* locus are only transiently required for pathogenesis of human UTIs and the urine samples only represent a stage of the infection that these genes are not involved. Alternatively, since most urovirulence genes only exist in a portion of UPEC strains, the *E. coli* strains in these urine samples may not have harbored these potential virulence genes.

In conclusion, the MIGs are potential targets for developing preventive and/or therapeutic strategies to manage UTIs as well as potential markers for differentiating UPEC from nonuropathogenic *E. coli*. Virulence factors of UPEC are good targets for prevention and treatment of UTIs. For example, the FimH adhesin of type 1 fimbriae is responsible for colonization of UPEC on the uroepithelium of the bladder. FimH antagonists have been developed as an anti-adhesive drug for oral treatment of UTIs (18). Also, FimH and iron receptors, such as IreA, Hma, and IutA, are able to induce a protective immune response against UPEC infections (1, 19). Thus, *cjrABC-senB*, *sisA*, *sisB*, and *fbpB*, which are involved in urovirulence of UPEC, are potential therapeutic and/or preventive targets. In addition, all the MIGs are potential markers for UPEC. Such markers may be valuable in public health for monitoring biological threats, such as outbreaks of UTIs caused by *E. coli* and the emergence of new virulent *E. coli* strains, and also in basic microbiology research, such as studies in evolution and classification of pathogenic *E. coli*. However, so far, none of the known urovirulence factors alone is sufficient to account for the virulence properties of UPEC and most of the urovirulence genes only exist in a portion of UPEC strains. Therefore, to develop effective and widely usable preventive and/or therapeutic strategies to manage *E. coli*-caused UTIs, a combination of multiple urovirulence genes to serve as the targets may be necessary. Accordingly, the properties of the MIGs revealed in this study, including their prevalence, phylogenetic distribution, and correlation patterns with other known virulence genes, may be beneficial for designing such a gene combination for controlling these *E. coli*-caused diseases.

ACKNOWLEDGMENTS

We thank Shainn-Wei Wang and Jyh-Mirn Lai for providing valuable comments on the manuscript.

This work was supported by the National Science Council, Taiwan (grant no. NSC 99-2320-B-006-005-MY3), National Cheng Kung University Hospital, Taiwan (grant no. NCKUH 96-10006001), China Medical University, Taiwan (grant no. CMU 99-S-04), and the Multidisciplinary Center of Excellence for Clinical Trial and Research, Department of Health, Executive Yuan, Taiwan (grant no. DOH100-TD-B-111-002).

REFERENCES

1. **Alteri CJ, Hagan EC, Sivick KE, Smith SN, Mobley HL.** 2009. Mucosal immunization with iron receptor antigens protects against urinary tract infection. PLoS Pathog. **5**:e1000586.
2. **Brzuszkiewicz E, et al.** 2006. How to become a uropathogen: comparative genomic analysis of extraintestinal pathogenic *Escherichia coli* strains. Proc. Natl. Acad. Sci. U. S. A. **103**:12879–12884.
3. **Chapman TA, et al.** 2006. Comparison of virulence gene profiles of *Escherichia coli* strains isolated from healthy and diarrheic swine. Appl. Environ. Microbiol. **72**:4782–4795.
4. **Chen SL, et al.** 2006. Identification of genes subject to positive selection in uropathogenic strains of *Escherichia coli*: a comparative genomics approach. Proc. Natl. Acad. Sci. U. S. A. **103**:5977–5982.
5. **Clermont O, Bonacorsi S, Bingen E.** 2000. Rapid and simple determination of the *Escherichia coli* phylogenetic group. Appl. Environ. Microbiol. **66**:4555–4558.
6. **Cusumano CK, Hung CS, Chen SL, Hultgren SJ.** 2010. Virulence plasmid harbored by uropathogenic *Escherichia coli* functions in acute stages of pathogenesis. Infect. Immun. **78**:1457–1467.

7. **Goullet P, Picard B.** 1986. Highly pathogenic strains of *Escherichia coli* revealed by the distinct electrophoretic patterns of carboxylesterase B. J. Gen. Microbiol. **132**:1853–1858.
8. **Grüneberg RN.** 1969. Relationship of infecting urinary organism to the faecal flora in patients with symptomatic urinary infection. Lancet **ii**:766–768.
9. **Hagan EC, Lloyd AL, Rasko DA, Faerber GJ, Mobley HL.** 2010. *Escherichia coli* global gene expression in urine from women with urinary tract infection. PLoS Pathog. **6**:e1001187.
10. **Ingersoll MA, et al.** 2003. The ShiA protein encoded by the *Shigella flexneri* SHI-2 pathogenicity island attenuates inflammation. Cell. Microbiol. **5**:797–807.
11. **Ingersoll MA, Zychlinsky A.** 2006. ShiA abrogates the innate T-cell response to *Shigella flexneri* infection. Infect. Immun. **74**:2317–2327.
12. **Johnson JR, Delavari P, Kuskowski M, Stell AL.** 2001. Phylogenetic distribution of extraintestinal virulence-associated traits in *Escherichia coli*. J. Infect. Dis. **183**:78–88.
13. **Johnson JR, Kaster N, Kuskowski MA, Ling GV.** 2003. Identification of urovirulence traits in *Escherichia coli* by comparison of urinary and rectal *E. coli* isolates from dogs with urinary tract infection. J. Clin. Microbiol. **41**:337–345.
14. **Johnson JR, O'Bryan TT, Kuskowski M, Maslow JN.** 2001. Ongoing horizontal and vertical transmission of virulence genes and *papA* alleles among *Escherichia coli* blood isolates from patients with diverse-source bacteremia. Infect. Immun. **69**:5363–5374.
15. **Johnson JR, Russo TA.** 2005. Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. Int. J. Med. Microbiol. **295**: 383–404.
16. **Johnson JR, Stell AL.** 2000. Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. J. Infect. Dis. **181**:261–272.
17. **Kingsley RA, et al.** 2003. Molecular and phenotypic analysis of the CS54 island of *Salmonella enterica* serotype Typhimurium: identification of intestinal colonization and persistence determinants. Infect. Immun. **71**: 629–640.
18. **Klein T, et al.** 2010. FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. J. Med. Chem. **53**:8627–8641.
19. **Langermann S, et al.** 1997. Prevention of mucosal *Escherichia coli* infection by FimH-adhesin-based systemic vaccination. Science **276**:607–611.
20. **Lloyd AL, Henderson TA, Vigil PD, Mobley HL.** 2009. Genomic islands of uropathogenic *Escherichia coli* contribute to virulence. J. Bacteriol. **191**: 3469–3481.
21. **Lloyd AL, Rasko DA, Mobley HL.** 2007. Defining genomic islands and uropathogen-specific genes in uropathogenic *Escherichia coli*. J. Bacteriol. **189**:3532–3546.
22. **Lloyd AL, Smith SN, Eaton KA, Mobley HL.** 2009. Uropathogenic *Escherichia coli* suppresses the host inflammatory response via pathogenicity island genes *sisA* and *sisB*. Infect. Immun. **77**:5322–5333.
23. **Marrs CF, Zhang L, Foxman B.** 2005. *Escherichia coli* mediated urinary tract infections: are there distinct uropathogenic *E. coli* (UPEC) pathotypes? FEMS Microbiol. Lett. **252**:183–190.
24. **McClelland M, et al.** 2001. Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. Nature **413**:852–856.
25. **Moss JE, Cardozo TJ, Zychlinsky A, Groisman EA.** 1999. The *selC*-associated SHI-2 pathogenicity island of *Shigella flexneri*. Mol. Microbiol. **33**:74–83.
26. **Nataro JP, et al.** 1995. Identification and cloning of a novel plasmid-encoded enterotoxin of enteroinvasive *Escherichia coli* and *Shigella* strains. Infect. Immun. **63**:4721–4728.
27. **Palaniappan RU, et al.** 2006. Differentiation of *Escherichia coli* pathotypes by oligonucleotide spotted array. J. Clin. Microbiol. **44**:1495–1501.
28. **Parham NJ, et al.** 2005. Prevalence of pathogenicity island IICFT073 genes among extraintestinal clinical isolates of *Escherichia coli*. J. Clin. Microbiol. **43**:2425–2434.
29. **Perna NT, et al.** 2001. Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. Nature **409**:529–533.
30. **Picard B, et al.** 1999. The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. Infect. Immun. **67**:546–553.
31. **Seed PC, Hultgren SJ.** 2005. Blueprinting the regulatory response of *Escherichia coli* to the urinary tract. Trends Microbiol. **13**:246–248.
32. **Smajs D, Weinstock GM.** 2001. The iron- and temperature-regulated *cjrBC* genes of *Shigella* and enteroinvasive *Escherichia coli* strains code for colicin Js uptake. J. Bacteriol. **183**:3958–3966.
33. **Soto SM, Guiral E, Bosch J, Vila J.** 2009. Prevalence of the *set-1B* and *astA* genes encoding enterotoxins in uropathogenic *Escherichia coli* clinical isolates. Microb. Pathog. **47**:305–307.
34. **Touchon M, et al.** 2009. Organised genome dynamics in the *Escherichia coli* species results in highly diverse adaptive paths. PLoS Genet. **5**:e1000344.
35. **Tseng CC, Huang JJ, Ko WC, Yan JJ, Wu JJ.** 2001. Decreased predominance of *papG* class II allele in *Escherichia coli* strains isolated from adults with acute pyelonephritis and urinary tract abnormalities. J. Urol. **166**: 1643–1646.
36. **Vokes SA, Reeves SA, Torres AG, Payne SM.** 1999. The aerobactin iron transport system genes in *Shigella flexneri* are present within a pathogenicity island. Mol. Microbiol. **33**:63–73.
37. **Wang MC, et al.** 2009. Different roles of host and bacterial factors in *Escherichia coli* extra-intestinal infections. Clin. Microbiol. Infect. **15**:372–379.
38. **Welch RA, et al.** 2002. Extensive mosaic structure revealed by the complete genome sequence of uropathogenic *Escherichia coli*. Proc. Natl. Acad. Sci. U. S. A. **99**:17020–17024.
39. **Xie Y, et al.** 2006. Identification and characterization of *Escherichia coli* RS218-derived islands in the pathogenesis of *E. coli* meningitis. J. Infect. Dis. **194**:358–364.