This article is downloaded from

http://researchoutput.csu.edu.au

It is the paper published as:

Author: T. Kudinha, F. Kong, J. R. Johnson, S. D. Andrew, P. Anderson and G. L. Gilbert

Title: Multiplex PCR-based reverse line blot assay to simultaneous detection of 22 virulence genes in uropathogenic Escherichia coli

Journal: Applied and Environmental Microbiology **ISSN**: 0099-2240 1098-5336

Year: 2012 **Volume:** 78 **Issue:** 4 **Pages:** 1198-1202

Abstract: Urinary tract infections (UTIs) are among the most common bacterial infections and are responsible for significant morbidity and health care costs worldwide. The main bacterial cause of uncomplicated UTI is Escherichia coli, which possesses numerous virulence factors (VFs). Many studies of the pathogenesis of E. coli UTI have centered on VF genes. Hence, the development of better molecular assays to study VF genes would facilitate these studies. We developed a highly sensitive and specific multiplex PCR-based reverse line blot (mPCR/RLB) assay to simultaneously detect 22 VF genes of uropathogenic E. coli and then used it to characterize 180 isolates from nonpregnant women of child-bearing age with cystitis and 153 fecal isolates from similar-age healthy women, in regional New South Wales, Australia. The assay accurately identified all VF genes (of the 22 under study) known to be present in 30 previously characterized control strains. The detection limits were 28 ng of DNA from E. coli isolates and 50 CFU/ml in mock-infected urine specimens containing known concentrations of E. coli. Cystitis isolates (compared to the fecal isolates) showed a significantly higher prevalence of 18 individual VF genes and contained significantly more VF genes per isolate (median number, 18.5 versus 6.5 [P = 0.001]). Discordance between paired probes for a given VF gene occurred in several clinical test isolates but no reference strains and among the test isolates was associated with fecal source (10% of VF genes versus 2% for cystitis isolates [P < 0.001]). This novel mPCR/RLB method is a potentially powerful tool for investigating the prevalence and distribution of VFs in E. coli.

URLs: http://researchoutput.csu.edu.au/R/-?func=dbin-jump-full&object_id=30986&local_base=GEN01-CSU01

Author Address:
sandrew@csu.edu.au
ptanderson@csu.edu.au

CRO Number: 30986

Multiplex PCR-based reverse line blot assay to simultaneously detect 22 virulence genes in uropathogenic *Escherichia coli*

Timothy Kudinha[1], Fanrong Kong[2], James R. Johnson[3], Scott Andrew[1], Peter Anderson[1] and Gwendolyn L. Gilbert[2,4*]

[1]Charles Sturt University, Leeds Parade, Orange, NSW 2800, Australia.

[2]Centre for Infectious Diseases and Microbiology, Institute of Clinical Pathology and Medical Research, Westmead Hospital, Westmead, NSW 2145, Australia.

[3]Veterans Affairs Medical Center and University of Minnesota, Minneapolis, Minnesota 55417, USA.

[4]Sydney Institute for Emerging Infectious Diseases and Biosecurity, University of Sydney, NSW 2006, Australia

***Corresponding author address:**

Professor Gwendolyn L. Gilbert

Centre for Infectious Diseases and Microbiology Laboratory Services (CIDMLS),

Institute of Clinical Pathology and Medical Research (ICPMR),

Westmead Hospital, Westmead,

NSW 2145, Australia.

Phone (612) 98456255, Fax (612) 98938659

Email: lyn.gilbert@sydney.edu.au

Summary

Urinary tract infections (UTIs) are among the commonest bacterial infections and are responsible for significant morbidity and healthcare costs world-wide. The main bacterial cause of uncomplicated UTI is *Escherichia coli*, which possesses numerous virulence factors (VFs). Many studies of the pathogenesis of *E. coli* UTI have centered on VF genes. Hence, the development of better molecular assays to study VF genes would facilitate these studies. We developed a highly sensitive and specific multiplex PCR based reverse line blot (mPCR/RLB) assay to simultaneously detect 22 VF genes of uropathogenic *E. coli*, then used it to characterize 180 isolates from non-pregnant women of child-bearing age with cystitis and 153 fecal isolates from similar-age healthy women, in regional New South Wales, Australia. The assay accurately identified all VF genes (of the 22 under study) known to be present in 30 previously-characterized control strains. Detection limits were 28 ng DNA from *E. coli* isolates and 50 cfu/mL in mock-infected urine specimens containing known concentrations of *E. coli.* Cystitis isolates (as compared with the fecal isolates) had significantly higher prevalence of 18 individual VF genes, and contained significantly more VF genes per isolate (median number, 18.5, vs 6.5; $p = 0.001$). Discordance between paired probes for a given VF gene occurred in several clinical test isolates but no reference strains, and among the test isolates was associated with fecal source (10% of VF genes, vs. 2% for cystitis isolates; $p < 0.001$). This novel mPCR/RLB method is a potentially powerful tool for investigation of the prevalence and distribution of VFs in *E. coli*.

Introduction

Urinary tract infections (UTIs) are among the commonest bacterial infections world-wide. They are often recurrent, may be difficult to treat, and, in some cases, lead to critical illness and death (7). Consequently, UTIs are responsible for significant morbidity, mortality and high costs world-wide (7). In the United States of America, the healthcare costs due to UTI are estimated to exceed one billion dollars annually (6, 24, 28). However, despite considerable study of this condition, its pathogenesis is still insufficiently elucidated. Therefore, studies that would improve our understanding of disease mechanisms and, potentially, improve UTI prevention and management could have significant economic and public health impact.

A single bacterial species, *Escherichia coli*, is the causative agent in 80-90% of uncomplicated UTIs in all age groups (10, 22). Uropathogenic *E. coli* (UPEC), the specialized strains of *E. coli* that cause most UTIs, possess virulence factors (VFs) that help them to colonize, invade and injure the host. The recognized VFs of UPEC include diverse adhesins, toxins, siderophores, and surface polysaccharides and proteins. In all, there are more than 40 suspected or confirmed UPEC VF genes (15).

Potential non-antimicrobial methods for UTI prevention include immunization, with VFs being logical vaccine targets (15). Therefore, knowledge of which VFs are prevalent in specific clinical syndromes and host populations is needed to inform the selection of VFs to be targeted. Because of the large number of UPEC VF genes and isolates to be studied, rapid and inexpensive molecular methods, with high throughput, are needed. Existing

multiplex PCR (mPCR) assays for studying UPEC genes detect ≤ 7 genes per PCR reaction (12, 15). Primers must be sorted into multiple pools according to compatibility and amplicon length, for resolution of PCR products by size in gel electrophoresis. Therefore, to study ≥ 30 VF genes, one must run more than 5 different mPCRs and then perform gel electrophoresis on the products, which is labor intensive, expensive and time consuming. Recently, capillary electrophoresis-based multiplex PCR assays have been described for detection of several pathogens, but none have been applied to UPEC VF genes (31).

Multiplex PCR-based reverse line blot (mPCR/RLB) assay is a well-established method for the simultaneous detection and genotyping of bacteria, fungi and viruses (19, 20, 36). Reverse line blotting involves the covalent coupling of amine-labeled oligonucleotide probes to a carboxyl-activated nylon membrane, followed by hybridization with biotin-labeled mPCR products from the test sample, then detection of the bound PCR products using an avidin-horseradish peroxidase complex and a chemiluminescent substrate, somewhat analogous to an immunoblot. The aim of the present study was to develop and evaluate a mPCR/RLB assay for the detection of 22 UPEC VF genes, using a combination of previously characterized *E. coli* control strains, urine isolates from women with cystitis, and fecal isolates from similar-age healthy women in a regional area of New South Wales (NSW) in Australia.

Materials and Methods

Setting

The study was done in the diagnostic microbiology laboratory of the Central West Pathology Service, Orange, NSW, which serves the Orange Base Hospital and other health facilities and general practices in the region.

E. coli reference strains.

The reference strains used for assay development and validation included 30 UPEC isolates for which VF gene content had been documented previously using 5 separate mPCRs (unpublished data, J.R Johnson). Here, these isolates were initially tested blindly using single PCR (sPCR), which detected all 22 VF genes included in this study in $\geq$ 1 of the isolates, in precise agreement with previous testing results (unpublished data, JR Johnson). Strains were stored in 5% glycerol in trypticase soy broth at -70°C until further use.

Cystitis isolates

E. coli isolates were selected from urine specimens submitted to the Microbiology department for culture from non-pregnant women of child-bearing age, with uncomplicated cystitis. This diagnosis was based on clinical information on the request form, as confirmed by the treating medical practitioner, and required one or more of dysuria, frequent urination, and suprapubic tenderness, without fever or loin pain, and a clean-catch, mid-stream urine culture yielding $\geq 10^8$ colony forming units (cfu) of *E. coli* per liter of urine. Quantitative urine culture was done using blood agar, MacConkey agar

and chromogenic agars, followed by conventional identification. *E. coli* isolates were stored in 5% glycerol in trypticase soy broth at -70°C. Relevant information recorded included clinical symptoms, urine microscopy results, and age. Since patient identifiers were removed, patient consent was not sought. The research ethics committees of the Sydney West Area Health Service and Charles Sturt University approved the study protocol.

Fecal isolates

Volunteers were recruited by local medical practitioners, from among non-pregnant women of childbearing age (18-45 years) who had no UTI-associated symptoms and were receiving healthcare for other conditions. Each participant provided written consent. A rectal swab was collected from each consenting volunteer and processed within 10 hours for isolation of *E. coli*, which was identified by conventional biochemical tests. One arbitrarily chosen *E. coli* colony per sample was analyzed.

DNA extraction

Approximately 5 colonies from a pure culture of *E. coli* on horse blood or MacConkey agar were suspended in 100 µL of digestion buffer (10mM Tris-HCL, pH 8.0, Triton X-100 0.45% v/v, Tween-20 0.45% v/v). This suspension was boiled for 10 min, followed by rapid cooling on ice. The tubes were then centrifuged at 16000g for 2 min and 400 µL of 0.5 TBE (1 X TBE:10mM Tris, 1 mm EDTA, pH 8.0) was added. After re-centrifugation, the supernatant was removed and stored at -20°C for future use as template DNA.

Targets

Twenty-two VF genes were chosen as PCR targets, based on their presumed importance in UTI pathogenesis (11, 14). These included: *papA*, P fimbriae structural subunit; *papC*, P fimbriae assembly; *papEF*, P fimbriae tip pilins; *papG,* P fimbriae adhesin (and alleles I, II, and III); *sfaS*, S fimbriae; *focG*, F1C fimbriae; *afa/draBC*, afimbrial adhesin (Dr antigen-specific fimbriae); *fimH*, type 1 fimbriae; *hlyA*, alpha hemolysin; *cnf1,* cytotoxic necrotizing factor type1; *fyuA*, ferric yersiniabactin receptor; *iutA*, aerobactin receptor; *iroN*, catecholate siderophore receptor; *kpsMII*, group 2 capsule; *kpsMTIII*, group 3 capsule; *traT*, serum-resistance associated; *ompT*, outer membrane protein T (protease); *bmaE*, M fimbriae; *gafD*, G fimbriae; and *usp*, uropathogenic-specific protein.

Primer and probe design

Primers used to detect VF genes were as described previously (12, 16). In addition, two sets of *E. coli*-specific primers were designed using the 16S-23S rRNA intergenic spacer region, which harbors both conserved and variable domains within species (Supplemental Table S1). Pairs of probes (sense and antisense) for each VF and *E. coli* primer were then designed based on the sequences immediately downstream or upstream of the sense and antisense primers (18). Probes were designed to have similar physical characteristics (namely: melting temperature 58-65°C; length 18-30 bases; moderate, weak, or no secondary structure; and no dimer formation), to allow simultaneous hybridization without loss of sensitivity (20). Primers and probes sequences were checked for specificity against all sequences in Genbank using SeqSearch in the Australian National Genomic Information Services (ANGIS) programs (www.angis.org.au).

Primers were 5'-labeled with biotin to enable detection by hybridization with a streptavidin-peroxidase substrate. Probes were 5'-labeled with an amine group to allow covalent binding to nylon membranes and to enable membrane stripping and reuse of the membrane without probe loss. Primers and probes were synthesized by AuGCT Biotechnology Synthesis Laboratory, Beijing, China.

PCR amplification

sPCR and mPCR amplifications used a 30 μL reaction mixture comprising 2 μL template DNA, 0.075 μL each of 24 primer pairs (100 pmol μL^{-1}), 1.25 μL deoxynucleoside triphosphate (0.125 mM of each), 2.5 μL 10 x buffer (Qiagen), 3.0 μL 25mM $MgCl_2$, 0.2 μL HotStar *Taq* polymerase (5 U μL^{-1}) (Qiagen), and molecular grade water (Eppendorf) to 30 μL. Amplification was performed on a Mastercycler gradient thermocycler under cycling conditions of 95°C for 15 min, followed by 35 cycles of 94°C for 30s, 55°C for 30s, then 72°C for 1 min, with a final 10 min extension at 72°C and a hold at 22°C. The amplified PCR products were processed immediately or within 1 week of storage at 4°C. For sPCR, amplified products were stained with SYBR DNA gel stain, separated electrophoretically in 2% agarose gels, and imaged using an ultraviolet transilluminator and digital capture system. To confirm the presence of previously identified VF genes, the 30 reference strains were processed in duplicate using DNA lysates prepared from separate colonies. Only one isolate from the reference strains produced a discrepant result (for only 1 VF gene) between the duplicate lysates. This was

resolved by reculturing the isolate from one colony and retesting. Reference strains and test isolates were tested in parallel by both sPCR and mPCR/RLB.

RLB hybridization assay

RLB hybridization assay was performed as previously described (18). Briefly, 5'-labeled probes were covalently bound to a nylon membrane. For this, a Biodyne C membrane was cut to 15^2 cm^2 size, labeled and placed in a 45-lane (slot) Miniblotter™ (MN45, Immunetics) and 150 µL of each diluted oligonucleotide probe was added to a separate lane. After room temperature incubation and several washing steps, the membrane was ready for hybridization.

Amplified PCR products were denatured by boiling for 10 min, then cooled on ice. The probe labeled membrane was placed in the blotter with the slots perpendicular to the probe lanes. Denatured PCR products (150µL) were added to the slots, one sample per slot, such that each sample contacted each of the probes. After 1 hour hybridization at 60^oC, the membrane was washed twice at 60^oC for 10 min and incubated in peroxidase-labeled streptavidin conjugate (Roche Diagnostics, Germany) at 42^oC for 1 hour. It was then washed at 42^oC and again at 25^oC. Detection involved 2 min incubation at 21^oC in a chemiluminescence blotting substrate (Amersham, GE Healthcare) for 2 min, and then covered with X-ray chemiluminescence film (Amersham Hyperfilm ECL, GE Healthcare) and exposed for 5 min. A clearly visible black hybridization signal was considered a positive result.

Sensitivity of the mPCR/RLB assay

The sensitivity of the mPCR/RLB assay was assessed by, first, preparing a series of 5x dilutions containing from 5-100 ng (total DNA) of genomic DNA in TE buffer from the 30 reference *E. coli* strains and testing them in the mPCR/RLB assay. Second, mock-infected urine specimens were prepared by diluting suspensions of each of the 30 reference strains by from 10^{-1} to 10^{-9} in fresh urine from healthy people. Bacterial concentrations were determined by quantitative plating. Template DNA was extracted from each dilution using the Roche COBAS Amplicor extraction kit (23). Additionally, to simulate field conditions, 10 clinical urine specimens that contained $\geq 10^8$ cfu/L of *E. coli* were tested in parallel by sPCR and mPCR/RLB. Template DNA was extracted as for the mock-infected urines. Finally, to assess the species specificity of the mPCR/RLB assay, single isolates of each of three bacterial species closely related to *E. coli*, (i.e., *Klebsiella pneumoniae*, *Enterobacter aerogenes,* and *Shigella sonnei*) were tested in duplicate.

Quality control of mPCR /RLB results

The positive control for each membrane was the mPCR product pool from a reference strain template DNA sample containing all 22 VF targeted VF genes. The negative control contained only master mix. Isolates with weak hybridization signals, or discordant signals between the probes for a given target, were rested. A test isolate's hybridization signal was considered positive if of similar or stronger intensity than the corresponding positive control. Results for a given probe pair were considered positive if either probe yielded a positive hybridization signal.

Data analysis

Comparisons of proportions between cystitis and fecal isolates were tested using a Chi-squared test or Fisher's exact test. P values < 0.05 were considered statistically significant. An isolate's VF score was the sum of *pap* (counted only once regardless of the number of *pap* genes detected) plus all other VF genes detected in the isolate. VF scores were compared using the Mann-Whitney U test.

Results

sPCR vs. mPCR/RLB

Overall, in comparison with sPCR, the new mPCR/RLB assay was extremely accurate in detecting specific UPEC VF genes. The 30 reference UPEC strains were tested in duplicate (using separate colonies) by both sPCR and mPCR/RLB. The mPCR/RLB assay correlated precisely with sPCR, detecting 100% of the VFs known to be present in the reference strains, with no false detections (i.e., sensitivity, specificity, and accuracy, all 100%). For each target, both probes gave concordantly positive or negative signals. Additionally, with representatives of 3 non-*E. coli* enteric species, i.e., *Klebsiella pneumoniae*, *Enterobacter aerogenes*, and *Shigella sonnei*, no cross reactions were detected by sPCR or mPCR/RLB for any of the gene targets.

Among the 333 test *E. coli* isolates, sPCR and mPCR/RLB again exhibited 100% correspondence for gene presence/absence. However, in the mPCR/RLB, 7 (4%) of 180 cystitis isolates and 8 (5%) of 153 fecal isolates showed within-probe-pair discordance for ≥ 1 VF genes. Within-probe-pair discordance varied in frequency by both gene and

source group, i.e., fecal vs. cystitis (Table 3). That is, such discrepancies occurred in up to 4% of isolates each for the adhesin genes *papAH*, *papEF*, *papC*, *papGII-III*, *afa/dra*, *fimH*, *sfaS*, and *focG*. For each VF gene with such discrepancies, the proportion of isolates with discordant results was generally higher among fecal isolates (≥ 9% for all genes except *fimH*) than cystitis isolates (≤ 4% for all genes except *cnf1*). This was most evident with the *pap* operon genes, each of which exhibited within-probe-pair discordance significantly more frequently among the fecal isolates. Using detected VF genes as the unit of analysis, for all VF genes combined, discordance was significantly more common among fecal isolates (10% of VF genes) than cystitis isolates (2% of VF genes) ($p < 0.001$).

For each isolate exhibiting within-probe-pair discordance, repeat testing by both mPCR/RLB and sPCR was done for all 22 VF genes. All previously noted mPCR/RLB within-probe-pair discrepancies were confirmed. Additionally for isolates with a single probe-positive result for a given VF gene, sPCR uniformly yielded an amplicon of the expected size for that VF gene, confirming presence of an intact VF gene despite the mPCR/RLB within-probe-pair discordance.

Analytical sensitivity of the mPCR/RLB assay

Based on testing of serial dilutions of total DNA from the 30 reference UPEC strains, the mPCR/RLB assay's limit of detection for each targeted VF gene was 28 ng of DNA (total amount). In mock-infected urine specimens containing $\geq 50 \times 10^3$ cfu/L *E. coli,* and in 10 naturally infected urine specimens containing $\geq 10^8$ cfu/L *E. coli*, the assay also detected

all targeted VF genes present in the sample, as defined by sPCR testing of the respective urine isolates.

Molecular epidemiology of cystitis and fecal isolates

To assess the mPCR/RLB assay's functionality in a molecular epidemiological application, 180 *E. coli* urine isolates from 180 women with cystitis, and 153 *E. coli* rectal isolates from 153 healthy women volunteers were tested. Subjects' ages ranged from 18 to 45 years (median per group, 32 years). All subjects resided in the Central West region of NSW (population ~ 180, 000; area, 63, 262 km^2).

All VF genes sought were found in ≥ 1 cystitis isolate, ranging in prevalence from 1% (*gafD* and *bmaE)* to 96% (*fimH)* (Table 2) whereas 4 VF genes were not found in any fecal isolate. Compared with fecal isolates, cystitis isolates had a higher prevalence of 18 of the 22 studied genes, including 8 adhesin genes, 2 toxin genes, 3 siderophore genes, and 5 genes for protectins, plus a numerically higher prevalence of the remaining 4 genes. Accordingly, cystitis isolates had significantly higher VF scores than did fecal isolates (median 18.5, vs. 6.5: $p = 0.001$).

Discussion

Current PCR-based assays for detection of UPEC VF genes are limited by the number of genes detected per PCR reaction (12, 15), thereby creating a need for on-going improvement in these assays. Here, we successfully developed a novel mPCR/RLB assay that detects simultaneously 22 UPEC VF genes using a single mPCR reaction. The assay exhibited 100% sensitivity and specificity in comparison with sPCR in detecting the 22 targeted VF genes among 30 previously characterized reference strains and 333 clinical test isolates. It also exhibited no cross-reactivity with representatives of three bacterial species closely related to *E. coli*.

We tested the new assay's practical utility by comparing the distribution of VF genes among cystitis and fecal isolates from non-pregnant women of child-bearing age in a region of NSW in Australia. We found that cystitis isolates contain significantly more VF genes than fecal isolates, which in turn tend to contain sequence variants of VF genes. In addition, we documented the mPCR/RLB assay's ability to detect VF genes directly, in naturally or mock-infected urine samples containing $\geq 10^8$ cfu/mL of *E. coli.*

Interestingly, in the mPCR/RLB 4% of test isolates yielded reproducible within-probe-pair signal discordance (one probe positive, one negative) for some VF genes. This phenomenon not only was limited to test isolates, but among the test isolates segregated by clinical source. We performed sPCRs, using the same primers as in the mPCR, to ascertain whether such discordance resulted from modified probe target sequences (sPCR amplicon present) or non-specific cross-reactions (sPCR amplicon absent). sPCR

uniformly yielded an amplicon of the expected size for all VF genes with mPCR/RLB probe-pair discrepancies. This, plus previous work by others (18, 19, 23, 25, 32) and DNA sequencing of selected sPCR products (data not shown), indicated that within-probe-pair discordance resulted from sequence variation within the target region of the signal-negative probes.

In previous studies discordant signals between paired probes occured rarely and represented at least a 2-bp change in the probe target region (35). Many UPEC VF genes are polymorphic (3, 20, 21, 30), due to point mutations, deletions, or other types of mutation, as reported for *cnf1* (12, 27). From a statistical and epidemiological perspective, infrequent discordance between paired probes should not significantly impair the overall interpretation of the data. On the contrary, such discordance may actually help identify isolates, or individual VF genes and their flanking regions, to analyze for sequence diversity in relation to phenotype and clinical source.

In that regard, we observed highly statistically significant prevalence differences for probe-pair discordance among *pap* operon genes between fecal and cystitis isolates, with discordance being confined to the fecal isolates ($p < 0.005$). This agrees with previous ecidence that fecal and asymptomatic bacteriuria isolates more commonly have partially deleted *pap* operons than do cystitis and pyelonephritis isolates (10, 26).

The present cystitis isolates harbored many more VF genes than did fecal isolates (median, 18.5 vs. 6.5 per isolate). Although this agrees with several previous studies (10,

15, 25), other studies found no such differences between cystitis and fecal isolates, but differences between both these groups and pyelonephritis isolates (14, 15, 25). Such between-study differences could reflect differences in VF gene distribution according to geographical location, host population characteristics, and/or sampling methods, including case definitions and inclusion criteria (5, 29, 35). They demonstrate the need for additional studies of this sort, which our novel mPCR/RLB assay should facilitate.

From a technical perspective, the mPCR/RLB assay allows the reliable, simultaneous detection of 22 known UPEC VF genes, with the potential to screen up to 43 samples for up to 43 genes simultaneously (given the slot-blot apparatus's 43 lanes), using one mPCR reaction per isolate. Once pure cultures are obtained, the assay's turnaround time is ≤ 24 h, including DNA extraction, mPCR set up and running, and RLB hybridization. The RLB membrane can be prepared in ≤ 2h and can be re-used at least 20 times, with interval stripping (18). Although several DNA microarrays and probe hybridization techniques have been described for studying *E. coli* VF genes (1, 4, 9), the mPCR/RLB assay is cheaper, simpler to perform, and more flexible than DNA microarrays (8, 33), and can be used to evaluate candidate primers and probes for microarray use.

This assay is sufficiently sensitive to detect VF genes in actual urine specimens, as demonstrated by its performance with 10 artificially or naturally infected urines. However, its main usefulness likely will be for high-throughput analysis of already-isolated strains in research laboratories. It likely can be further developed to detect most

of the currently recognized VF genes of UPEC in one mPCR reaction, thereby facilitating the study of UPEC.

Acknowledgements

We would like to thank staff of the Department of Microbiology at Central West Pathology services for help with the culturing and identification of *E. coli* isolates.

References

1. **Bekal, S., R. Brousseau, L. Masson, G. Profontaine, J. Fairbrother, and J. Harel.** 2003. Rapid identification of *Escherichia coli* pathotypes by virulence gene detection with DNA microarrays. J. Clin. Microbiol. **41(5):** 2113-2125.

2. **Bijlsma, I.G., L. van Dijk, J.G. Kusters, and W. Gaastra.** 1995. Nucleotide sequences of two fimbrial major subunit genes, *pmpA* and *ucaA*, from canine-uropathogenic *Proteus mirabilis* strains. Microbiol. **141:**1349-57.

3. **Boyd, E.F., and D.L Hartl.** 1998. Diversifying selection governs sequence polymorphisms in the major adhesin proteins *FimA, PapA,* and *sfa* of *Escherichia coli*. J. Mol. Evol. **47:**258-67.

4. **Bruant, G., C. Maynard, S. Bekal, S. Gaucher, L. Masson, R. Brousseau, and J. Harel.** 2006. Development and validation of an oligonucleotide microarray for detection of multiple virulence and antimicrobial resistance genes in *Escherichia coli*. Appl. Environ. Microbiol. **72 (5):**3780-3784.

5. **Duriez, P., O. Clermont, S. Bonarcorsi, E. Bingen, A. Chaventre, J. Elion, B. Picard, and E. Denamur.** 2001. Commensal *Escherichia coli* isolates are phylogenetically distributed among geographically distinct human populations. Microbiology. **147:**1671-1676.

6. **Foxman, B.** 2002. Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. Am. J. Med. **113 (1A):**5S-13S.

7. **Foxman, B., and B.P. Brown.** 2003. Epidemiology of urinary tract infections: transmission and risk factors, incidence and costs. Infect. Dis. Clin North. Am. **17:**227-241.

8. **Goguet de la Salmoniere, Y.O., C.C. Kim, A.G. Tsolaki, A.S. Pym, M.S. Siegrist, and P.M. Small.** 2004. High-throughput method for detecting genomic-deletion polymorphisms. J. Clin. Microbiol. **42:**2913-2918.

9. **Hamelin, K., G. Bruant, A. El-Shaarawi, S. Hill, T.A. Edge, S. Bekal, J.M. Fairbrother, J. Harel, C. Maynard, L. Masson, and R. Brousseau.** 2006. A virulence and antimicrobial resistance DNA microarray detects a high frequency of virulence genes in *Escherichia coli* isolates from great lakes recreational waters. Appl. Environ. Microbiol. **72 (6):**4200-4206.

10. **Hooton, T.M., and W.E. Stamm.** 1997. Diagnosis and treatment of uncomplicated urinary tract infection and treatment of uncomplicated urinary tract infection. Infect. Dis. Clin. North. Am. **11(3):**551-81.

11. **Johnson, J.R.** 2003. Microbial virulence determinants and the pathogenesis of urinary tract infection. Infect. Dis. Clin. North. Am. **17:**261-278.

12. **Johnson, J.R., and A.L. Stell.** 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.

13. **Johnson, J.R., A.L. Stell, N. Kaster, C. Fasching, and T.T. O'Bryan.** 2001. Novel molecular variants of allele 1 of the *Escherichia coli* P fimbrial adhesin gene *papG*. Infect. Immun. **69:**2318-2327.

14. **Johnson, J.R., K. Owens, A. Gajewski and M.A. Kuskowski.** 2005. Bacterial characteristics in relation to clinical source of *Escherichia coli* isolates from women with acute cystitis or pyelonephritis and uninfected women. *J. Clin. Microbiol*. **43:**6064-72.

15. **Johnson, J.R., and T.A. Russo.** 2005. Molecular epidemiology of extraintestinal pathogenic (uropathogenic) *Escherichia coli*. Int. J. Med. Micro. **295:**383-404.

16. **Johnson, J.R., T.A Russo, P.I Tarr, U, Carlino, S.S. Bilge, J.CJ. Vary, and A.L. Stell.** 2000. Molecular epidemiological and phylogenetic associations of two novel putative virulence genes, *iha* and $iroN_{E.coli}$, among *Escherichia coli* isolates from patients with urosepsis. Infect. Immun. **68:**3040-3047.

17. **Karkkainen, U-M, J. Kauppinin, R. Ikaheimo, M.-L.Katila, and A.Siitonen.**1998. Rapid and specific detection of three different G adhesin classes of P-fimbriae in uropathogenic *Escherichia coli* by polymerase chain reaction. J. Microbiol. Methods **34:** 23-29.

18. **Kong, F., and G.L. Gilbert.** 2006. Multiplex PCR-based reverse line blot hybridization assay (mPCR/RLB)-a practical epidemiological and diagnostic tool. Nature Protocols **1(6):**2668-2680.

19. **Kong, F., L.Ma, and G.L. Gilbert.** 2005. Simultaneous detection and serotype identification of *Streptococcus agalactiae* using multiplex PCR and reverse line blot hybridization. J. Med. Microbiol. **54:**1133-1138.

20. **Marc, D., and M. Dho-Moulin.** 1996. Analysis of the *fim* cluster of an avian O2 strains of *Escherichia coli*: serogroup-specific sites within *fimA* and nucleotide sequence of *fimI*. J. Med. Microbiol. **44:** 444-52.

21. **Marklund, B.I., J.M. Tennent, E. Garcia, *et al.*** 1992. Horizontal gene transfer of the *Escherichia coli pap* and *prs* pili operons as a mechanism for the development of tissue-specific adhesive properties. Mol. Microbiol. **6:**2225-42.

22. **Martinez, J.J., and S.J. Hultgren**. 2002. Requirement of Rho-family GTPases in the invasion of type 1-piliated uropathogenic *Escherichia coli*. Cell. Microbiol. **4(1):**19-28.

23. **Mckechnie, M.L., R. Hillman, D. Couldwell, F. Kong, E. Freedman, H. Wang and G.L. Gilbert.** 2009. Simultaneous identification of 14 genital microorganisms in urine by use of a multiplex PCR-based reverse line blot assay. J. Clin. Microbiol. **47(6):**1871-1877.

24. **Mehnert-Kay, S.A.** 2005. Diagnosis and management of uncomplicated urinary tract infections. American Family Physician **72 (3):** 451-456.

25. **Plos, K., T. Carter, S. Hull, R. Hull, and C. Svanborg.** 1990. Frequency and organization of pap homologous DNA in relation to clinical origin of uropathogenic *Escherichia coli*. J. Infect. Dis. **161:** 528-524.

26. **Moreno, E., A. Andreu, C. Pigrau, M.A. Kuskowski, J.R. Johnson, and G. Prats.** 2008. Relationship between *Escherichia coli* strains causing acute cystitis in women and the fecal *E. coli* population of the host. J. Clin. Microbiol. **46(8):**2529-2534.

27. **Rippere-Lampe, K.E., D. O'Brien, R. Conran, and H.A. Lockman.** 2001. Mutation of the gene encoding cytotoxic necrotizing factor type 1 (*cnf1*) attenuates the virulence of uropathogenic *Escherichia coli*. Infect. Immun. **69(6):**3954-3964.

28. **Russo, T.A., and J.R. Johnson.** 2003. Medical and economic impact of extraintestinal infections due to *Escherichia coli*: focus on an increasingly important endemic problem. Microbes Infect. **5:**449-456.

29. **Sannes, M.R., M.A. Kuskowski, K. Owens, A. Gajewski, and J.R. Johnson.** 2004. Virulence factor profiles and phylogenetic background of *Escherichia coli* isolates from veterans with bacteremia versus uninfected control patients. J. Infect. Dis. **190:**2121-2128.

30. **Sokurenko, E.V., V. Chesnokova, D.E. Dykhuizen *et al.*** 1998. Pathogenic adaptation of *Escherichia coli* by natural variation of the *fimH* adhesin. Proc. Natl. Acad. Sci. USA **95:**8922-6.

31. Stevenson, J.B., W.C. Hymas, and D.R. Hillyard. 2011.A novel capillary electrophoresis-based multiplex PCR assay for detection of respiratory patjogens. Annals. Clin. Lab. Sci. **41 (1)**: 33-38

32. **Wang, O., F. Kong, P. Jelfs, and G.L.Gilbert.** 2008. Extended phage locus typing of *Salmonella enterica serovar Typhimurium,* using multiplex PCR-based reverse line blot hybridization. J. Med. Microbiol. **57:**827-838.

33. **Wen, L., Q. Wang, Y. Li, F. Kong, G.L. Gilbert, B. Cao, L. Wang and L. Feng.** 2006. Use of a serotype-specific DNA microarray for identification of group B s*treptococcus* (*Streptococcus agalactiae*). J. Clin. Microbiol. **44:**1447-1452.

34. **Zeng, X., F. Kong, H. Wang, A. Darbar, and G.L.Gilbert.** 2006. Simultaneous detection of nine antibiotic resistance-related gene in *Streptococcus agalactiae* using multiplex PCR and reverse line blot hybridization assay. Antimicrob. Agents Chemother. **50:**204-209.

35. **Zhang, L., B. Foxman, and C. Marrs.** 2002. Both urinary and rectal *Escherichia coli* are dominated by strains of phylogenetic group B2. J. Clin. Microbiol. **40:**11-14.

36. **Zhao, Z., F. Kong, and G.L Gilbert.** 2006. Reverse line blot assay for direct identification of seven *Streptococcus agalactiae* major surface protein antigen genes. Clin. Vaccine Immunol. **13:**145-149.

Supplemental Table S1: Novel oligonucleotide probes and primers developed for the mPCR/RLB assay used in this study

Probe/Primer name[a]	Length (bp)	Tm (°C)	GenBank Accession no.	Sequence (5'to 3') and coordinates within GenBank entry
papC-AP	20	63	X61239	4817 GTAGCCGGCCATATTCACAT 4798
papC-SP	20	64	X61239	4924 TAGTCCGCTGGCAAATTTGT 4944
fimH-AP	25	61	AJ225176	1864 CACATCATTATTGGCGTAAATATTC 1840
fimH-SP	25	60	AJ225176	2231 GTTTATCAATAAAGAAATCA CAGGG 2255
cnf1-AP	28	61	X70670	1700 TTTTCTATTGCATGTAGAACAGAATTTA 1673
cnf-1-SP	26	65	X70670	2097 CCAG GAGGTACTTAGCAGCGTTATAA 2122
hlyA-AP	21	62	M10133	2570 AAACATTGCCTGTTTTGAAGC 2550
hlyA-SP	23	60	M10133	3549 GGTGATGACCATATAGAAGGAAA 3571
papG I-III-AP	23	65	M20181	1244 AAAGCTGGGAACCATTTTTTCAT 1222
papG I-III-SP	21	65	M20181	2209 GCAACGCTGCTCATGATATTG 2229
papG I-AP	22	64	X61239	8913 CCTGAAAATACACTCCACTCGC 8892
papG I-SP	26	60	X61239	9285 GCCACTAATACATTGATGTTATCATT 9310

papG II-AP	22	60	M20181	1645 GTTTCTGACACAAATTACCTGC 1624
papG II-SP	20	65	M20181	1755 CGGCATACAGCGTCATTTCG 1774
papGIII-AP	23	61	X61238	1464 TCAGACCAGTAAAGACCATGAGT 1442
papGIII-SP	24	61	X61238	1632 GTTTCAATATCGGAGAAAAAAGAA 1655
afa/draBC-AP	23	65	X76688	4632 CGGCGTCGGGTTAACCCCCTTCA 4610
afa/draBC-SP	21	65	X76688	5139 CTGAAGACCTGTCTGACCCGT 5159
focG-AP	23	62	S68237	678 TACCCTCCCTGTAACAGTAATCG 656
focG-AP	26	62	S68237	947 TTGGTTCAACAAAA GTTGTTACAGTG 972
sfaS-AP	23	64	S53210	718 GTTCTTTGCAAAACATTACCCGT 697
sfaS-SP	21	63	S53210	875 GGATGGGCAGACATACTATGC 895
fyuA-AP	20	62	Z38064	814 GTGCCGCCTAAGTCATCGCT 795
fyuA-SP	20	62	Z38064	1519 AGGGATATAAACCTTCCGGG 1538
iutA-AP	20	60	X05874	892 ATTCATCGATGTTCAGCGTA 873
iutA-SP	21	64	X05874	1111 TACTACCGCGATGAGTCGTTG 1131
traT-AP	23	63	J01769	505 CTCAAGGTTACGCTTCTTGATTG 482
traT-SP	22	65	J01769	706 GATAAGATGGATCTGCGGGAGT 727

papAH-AP	25	62	X61239	1842 GGAATAGTTGGAGCAGCATTATTTA 1818
papAH-SP	20	64	X61239	2475 CCCTTCCTGAATACTGGGGA 2494
papEF-AP	24	63	X61239	8104 CATTTTATTTTACGAGATATAAAATTAACG 8075
papEF-SP	24	62	X61239	8317 AAAAACCATAAGCATATCCTGTCC 8340
bmaE-AP	23	62	M15677	123 TGTATGAGTGGCTGTTACTGTCA 101
bmaE-SP	23	62	M15677	539 GACGGAAAATTAACCTGATGAAA 561
gafD-AP	21	61	L33969	153 ATTGCATGAGTGCTGGAATAA 132
gafD-AP	27	60	L33969	1013 GAATATCAACATTTACGTTTTCCTATC 1039
iroN-AP	20	63	AF135597	1770 GGTTTGGGGCTTTAAAGGTT 1751
iroN-SP	20	61	AF135597	2341 TGGGGACGAACTTCAATTAC 2360
ompT-AP	20	64	X06903	648 GTCGAGTTGACTGACTTTTCG 628
ompT-SP	21	61	X06903	1126 TGGTGGAATCATCTGATAAC 1146
usp-AP	20	65	AB056434	6863 CTTCACCCGTATGAACACCAT 6843
usp-SP	20	63	AB056434	7201 GCTGCCTGGTGTTGTAACAG 7220
kpsII-AP	23	64	X53819	340 TGCATAATGTAACCAAAAATGCC 318
kpsII-SP	29	64	X53819	516 ACGCTGATTTACGTTGCTGTTTATATATT 544

kpsIII-AP	24	62	AF007777	4095 CCATAAAAACCAATGCTGATATTG 4072
kpsIII-SP	25	60	AF007777	4393 CAGGAAAA TCTACTTTGTTACGTTT 4417
E.coli 16S-AP	20	20	J01859	1210 AGGGCCATGATGACTTGACG 1191
E.coli 16S-SP	23	23	J01859	1458 GCGCTTACCACTTTGTGATTCAT 1480
E.coli 16S-AB	19	60	J01859	1539 AGGAGGTGATCCAACCGCA 1521
E.coli 16S-SB	20	62	J01859	1169 AACTGGAGGAAGGTGGGGAT 1188

[a]The suffix B indicates a biotin-labeled primer, and P indicates an amine-labeled probe. “A”, antisense; “S”, sense.

Table 2. Distribution of virulence-associated genes by source among 333 *Escherichia coli* cystitis and fecal isolates from non-pregnant women of child-bearing age

Category	Virulence gene	Prevalence, no. (column %)		
		Cystitis (n = 180)	Fecal (n = 153)	P value[a]
Adhesins	*afa/draBC*	27 (15)	3 (2)	<0.001
	bmaE	2 (1)	0 (0)	>0.05
	sfaS	47 (26)	17 (11)	0.017
	fimH	173 (96)	136 (89)	>0.05
	focG	90 (50)	20 (13)	<0.001
	papG1	4 (2)	0 (0)	>0.05
	papG11	59 (33)	17 (11)	<0.001
	papG111	47 (26)	12 (8)	<0.001
	papAH	108 (59)	54 (35)	<0.001
	papC	144 (80)	47 (31)	<0.001
	papEF	139 (77)	57 (37)	<0.001
	gafD	2 (1)	0 (0)	>0.05
Toxins	*cnf1*	68 (38)	23 (15)	< 0.001
	hlyA	122 (68)	29 (19)	< 0.001
Siderophores	*iutA*	121 (67)	18 (12)	< 0.001
	fyuA	140 (78)	24 (16)	< 0.001
	iroN	122 (68)	44 (29)	<0.01

Protectins	*kpsMII*	108 (60)	23 (15)	<0.001
	kpsMTIII	13 (7)	0 (0)	<0.001
	traT	139 (77)	47 (31)	<0.001
	ompT	120 (70)	73 (48)	<0.01
	usp	131 (73)	38 (25)	<0.01

[a]P values (by Fisher's exact test) are shown where $P < 0.05$ comparing cystitis with fecal isolates. The 22 virulence factors analyzed were; *papA*, P fimbriae structural subunit; *papC*, P fimbriae assembly; *papEF*, fimbriae tip pilins; *papG,* P fimbriae adhesin (and alleles 1, 11 and III); *sfaS*, S fimbriae; *focG*, F1C fimbriae; *afa/draBC*, Afimbrial adhesin (Dr-binding adhesin); *fimH*, type 1 fimbriae; *hlyA*, hemolysin; *cnf1,* cytotoxic necrotizing factor type1; *fyuA*, ferric yersiniabactin receptor; *iutA*, aerobactin receptor; *iroN*, catecholate siderophore receptor; *kpsMII* group 2 capsule (with K1 and K2 variants); *kpsMTIII*, group 3 capsule; *traT*, serum-resistance associated; *ompT*, outer membrane protein T (protease); *bmaE*, M fimbriae; *gafD*, (G) fimbriae adhesin.

Table 3. Distribution of discordant multiplex PCR results among 180 and 153 *Escherichia coli* cystitis and fecal isolates, respectively, from non-pregnant women of childbearing age.

Category	Virulence gene	Discordance, proportion of detected genes (%)		
		Cystitis isolates (n = 180)	Fecal isolates (n = 153)	P value
Adhesins	*afa/draBC*	1 /27 (4)	2/3 (67)	> 0.05
	sfaS	2 /47 (4)	3/ 17 (18)	> 0.05
	fimH	4/173 (2)	6/136 (4)	> 0.05
	focG	2 /90 (2)	3 /20 (15)	> 0.05
	papGII	0 /59 (0)	3/17 (18)	< 0.025
	papGIII	0 /47 (0)	5 /12 (42)	< 0.0015
	papAH	0 /108 (0)	5 /54 (9)	< 0.025
	papC	0 /144 (0)	4/47 (9)	< 0.005
	papEF	0 139 (0)	5 /57 (9)	< 0.005
Toxins	*cnf1*	7 /68 (10)	2/23 (9)	> 0.05
Others	(varied)	0/ 1521(0)	0/ 296 (0)	
Total	any of above	16/882 (1.8)	38/385 (9.9)	<0.001

[a]P values (by Fisher's exact test) are shown where P < 0.05 comparing cystitis with fecal isolates. The 22 virulence factors analyzed were; *papA*, P fimbriae structural subunit;

papC, P fimbriae assembly; *papEF*, fimbriae tip pilins; *papG,* P fimbriae adhesin (and alleles 1, 11 and III); *sfaS*, S fimbriae; *focG*, F1C fimbriae; *afa/draBC*, Afimbrial adhesin (Dr-binding adhesin); *fimH*, type 1 fimbriae; *hlyA*, hemolysin; *cnf1,* cytotoxic necrotizing factor type1; *fyuA*, ferric yersiniabactin receptor; *iutA*, aerobactin receptor; *iroN*, catecholate siderophore receptor; *kpsMII* group 2 capsule (with K1 and K2 variants); *kpsMTIII*, group 3 capsule; *traT*, serum-resistance associated; *ompT*, outer membrane protein T (protease); *bmaE*, M fimbriae; *gafD*, (G) fimbriae adhesin.