WATER RESEARCH 41 (2007) 3785–3791

Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/watres

Detection of virulence genes in *Escherichia coli* of an existing metabolic fingerprint database to predict the sources of pathogenic *E. coli* in surface waters

W. Ahmed[a,*], J. Tucker[a], K.A. Bettelheim[b], R. Neller[a], M. Katouli[a]

[a]Faculty of Science, Health and Education, University of the Sunshine Coast, Maroochydore DC 4558, Qld, Australia
[b]Microbiological Diagnostic Unit, Public Health Laboratory, Department of Microbiology and Immunology, University of Melbourne, Parkville, Vic, Australia

ARTICLE INFO

Article history:
Received 10 August 2006
Received in revised form
12 December 2006
Accepted 17 December 2006
Available online 7 February 2007

Keywords:
Faecal contamination
Escherichia coli
Virulence genes
Biochemical fingerprinting
Faecal source tracking

ABSTRACT

A collection of 366 *Escherichia coli* strains from 10 host groups and surface waters were tested for the presence of 15 virulence genes associated with strains causing intestinal and extra-intestinal infections. The virulence genes included *eaeA*, VT1, 2 and 2e, LT1, ST1 and 2, Einv gene, EAgg gene, CNF1 and 2, *papC*, O111 and O157 side chain LPS. Of the 262 strains obtained from nine different hosts, 39 (15%) carried one or more of these virulence genes. These included six strains from humans, two from horses, eight from dogs, two from ducks, five from cattle, seven from chickens, four from pigs, two from sheep and three from deer. Of the remaining 104 strains obtained from water samples, 10 (10%) also carried one or more of the tested virulence genes. Of these, six had identical biochemical phenotypes (BPTs) to strains isolated from humans (two strains), dogs (two strains), chickens (one strain) and sheep (one strain) with 4 BPTs also carrying same virulence genes. Our results indicate that the sources of clinically important *E. coli* strains found in surface waters due to faecal contamination can be predicted by using a combination of biochemical fingerprinting method and the detection of virulence genes. From the public health point of view this information will be of great importance for evaluating the risk associated with public use of the catchment.

1. Introduction

Escherichia coli strains are normal inhabitants of the gut of warm-blooded animals including humans. Whilst most *E. coli* strains in the gut are non-pathogenic, certain strains may carry virulence genes which enable them to cause intestinal infections such as diarrhoea or haemolytic colitis, or to cause extra-intestinal infections such as neonatal meningitis, nosocomial septicaemia, haemolytic uremic syndrome, urinary tract and surgical site infections (Falagas and Gorbach, 1995). Enterovirulent strains can be categorized as enteropathogenic *E. coli* (EPEC), enterotoxigenic *E. coli* (ETEC), enteroinvasive *E. coli* (EIEC), enterohemorrhagic *E. coli* (EHEC), enteroaggregative *E. coli* (EAEC), and diffusely adherent *E. coli* (DAEC) (Nataro and Kaper, 1998). Whilst pathogenic strains may cause disease in both humans and animals, it is also known that some strains may be host specific. For instance, heat labile toxin 1 (LT1) and heat stable toxin 1 (ST1) producing *E. coli* have been found in humans and/or pigs while the production of heat stable toxin 2 (ST2) has only been reported in humans (Echeverria et al., 1984; Guth et al., 1986).

*Corresponding author. Tel.: +61 07 5479 0875; fax: +61 07 5430 2887.
E-mail address: shuhat@yahoo.com (W. Ahmed).
0043-1354/$ - see front matter
doi:10.1016/j.watres.2006.12.026

Warm-blooded animals release such bacteria into the environment where they may be transported to surface waters via surface and sub-surface run-off (Fujioka, 2002). Contamination of surface waters with pathogenic strains of *E. coli* has been implicated in an increase in the number of disease outbreaks and consequent deaths (Feldman et al., 2002; Olsen et al., 2002). The presence or absence of *E. coli* and other faecal indicator bacteria in surface waters can only be used to predict the quality of water with regards to faecal contamination and does not provide sufficient information with respect to their possible source(s) or the degree of risk posed by such contamination (Gilpin et al., 2002). In response to this, researchers have developed microbial source tracking (MST) methods to predict and correct the faecal contamination from point and non-point source (Carson et al., 2001; Hagedorn et al., 2003). However, MST methods cannot be used to determine the potential public health risks associated with faecal contamination. It has been reported that there are over 30 virulence genes present in *E. coli* strains with the potential to cause diseases in humans and animals (Gordon, 2001). The advance in molecular technique such as polymerase chain reaction (PCR) and quantitative-PCR has made it easy to detect and quantify the virulence genes in *E. coli* from animals and environmental samples.

In this study, we postulated that fingerprinting of *E. coli* strains from host groups and environmental samples carrying clinically significant virulence genes can be compared to predict the sources of such strains in environmental samples. In view of the above, we used PCR to detect the presence of 15 virulence genes in *E. coli* strains of an existing metabolic fingerprint database of 10 host groups to predict the sources of pathogenic *E. coli* in surface waters.

2. Materials and methods

2.1. Sources of isolates

A total of 3728 *E. coli* isolates from 10 host groups were tested with a biochemical fingerprinting method in a previous MST study (Ahmed et al., 2005a). The host groups included humans, cattle, horses, dogs, ducks, sheep, kangaroos, pigs, deer and chickens. In all, 530 biochemical phenotypes (BPTs) were identified among all isolates. One representative strain of each BPT was saved in vials containing 1.5 ml tryptic soy broth (Oxoid, UK) with 15% glycerol at −80 °C. From this collection, 262 (49%) strains were randomly selected from all host groups (Table 2) for this study and tested for the presence of 15 virulence genes associated with *E. coli* strains causing intestinal and extra-intestinal infections (Table 1).

From the same study, a total of 550 *E. coli* isolates were also collected from five sampling sites of a creek and typed for their BPTs. The creek is approximately 8 km in length and is located within Eudlo Catchment with total area of 7980 ha, mostly comprised of rural areas with a high density of septic systems and farm/domestic animals (Ahmed et al., 2005b). In all, 282 BPTs were found among all isolates (Ahmed et al., 2005a). Of these, 104 strains were randomly selected to represent five sampling sites (i.e. up to 23 strains from each site) and tested for the presence of virulence genes. Strains from water samples which carried one or more of the 15 virulence genes were typed with the biochemical fingerprinting method to be compared with the BPTs of strains carrying same virulence genes from different host groups.

2.2. DNA extraction

All *E. coli* strains were streaked twice on McConkey's agar (Oxoid, UK) for purity. A single colony was inoculated into 50 ml flask containing 10 ml Luria Bertani (LB) broth made of 10 g (w/v) tryptone (Oxoid, UK), 5 g (w/v) NaCl, 5 g (w/v) yeast extract (Oxoid, UK) and 1 L distilled water. The flasks were kept in an incubator shaker at 110 rpm for overnight. DNA was extracted from bacteria grown in pure culture using a previously described method (Tillett and Neilan, 2000).

2.3. PCR amplification

Fifteen sets of PCR primers were used in this study. Primer sets were diluted according to the manufacturer's (Gene Works, Australia) instructions. Primer sets for the attachment and effacement (*eaeA*) gene, verotoxin (VT) 1, 2 and 2e, heat-labile toxin (LT) 1, heat-stable toxins (ST) 1 and 2, enteroinvasive (Einv) gene, enteroaggregative (EAgg) gene, cytotoxic necrotizing factors (CNF) 1 and 2 gene (Pass et al., 2000) were diluted to 25 nM concentrations. However, primers for hemolysin A (*hlyA*) (Boyd and Hartl, 1998), pyelonephritis-associated pilli (*papC*) (Lai et al., 1999), LPS O111 and O157side chain (Paton and Paton, 1998), were diluted to a concentration of 50 nM. A standard master mix of 11.8 μl sterile MilliQ water, 2.4 μl $MgCl_2$, 2.4 μl reaction buffer, 2 μl deoxynucleoside triphosphates (dNTPs), 0.16 μl of *Taq* polymerase (Fisher-biotech, Australia), 0.4 μl forward and 0.4 μl reverse primer was used per reaction. This resulted in a total volume of 19.56 μl per tube. A 2 μl of DNA template was added in each tube. For *hlyA* and *papC*, $MgCl_2$ concentration gradient was performed for optimization (i.e. 2.5 μl). The standard PCR reaction mixture consisted of 11.9 μl sterile Milli-Q water, 2.5 μl $MgCl_2$, 2.4 μl reaction buffer, 2 μl dNTPs, 0.16 μl of *Taq* polymerase (Fisher-biotech), 0.3 μl forward and 0.3 μl reverse primer per reaction. This resulted in a total volume of 19.56 μl per tube. Table 1 shows the primer sequence and the amplicon size of the target genes.

PCR (Eppendorf, Mastercycler gradient, Germany) for *eaeA*, VT1, VT2, VT2e, LT1, ST1, ST2, Einv, Eagg, CNF1 and CNF2 was performed under optimized conditions and consisted of 5 cycles of 95 °C for 30 s and 72 °C for 1 min followed by 25 cycles at 95 °C for 30 s; 63 °C for 30 s, 72 °C for 30 s and 1 cycle of 72 °C for 5 min. PCR amplification of *hlyA* consisted of 1 cycle of 94 °C for 30 s; 30 cycles of 94 °C for 30 s, 55 °C for 1 min, 68 °C for 6 min and 1 cycle of 72 °C for 10 min. PCR for O157 and O111 LPS side-chain consisted of 35 cycles of 95 °C for 1 min, 65 °C for 2 min for the first 10 cycles, decrementing to 60 °C by cycle 15 and 72 °C for 1.5 min incrementing to 2.5 min from cycles 25–35. Ladders (i.e. 100 base pairs; GeneWorks) used to assess the PCR for all of the primers except *hlyA* for which 1 kb ladders (GeneWorks) were used. To detect the amplified product, 3 μl aliquot of the PCR product was visualized by electrophoresis through a 1.5% agarose gel (Progen, Australia) in 1 × TAE buffer (50 × TAE: 242 g Tris base, 57.1 ml glacial

Table 1 – The primer sequence and the amplicon size of 15 virulence genes found among *Escherichia coli* strains associated with intestinal and extra-intestinal sites

Target gene	Pathogenic factor	Primer Sequences (5′′–3′′)	Amplicon size (base pairs)	Control *E. coli* strains
EaeA	Attaching and effacing (EPEC)	F: 5′-TGAGCGGCTGGCATGAGTCATAC-3′ R: 5′-TCGATCCCCATCGTCACCAGAGG-3′	241	12079
EAgg	Enteroaggregative adhesion (EAEC)	F: 5′-AGACTCTGGCGAAAGACTGTATC-3′ R: 5′-ATGGCTGTCTGTAATAGATGAGAAC-3′	194	Hague
Einv	Invasion (EIEC)	F: 5′-TGGAAAAACTCAGTGCCTCTGCGG-3′ R: 5′-TTCTGATGCCTGATGGACCAGGAG-3′	140	D434
VT1	Verotoxin (EHEC)	F: 5′-ACGTTACAGCGTGTTGCTGGGATC-3′ R: 5′-TTGCCACAGACTGCGTCAGTTAGG-3′	121	12079
VT2	Verotoxin (EHEC)	F: 5′-TGTGGCTGGGTTCGTTAATACGGC-3′ R: 5′-TTGCCACAGACTGCGTCAGTTAGG-3′	102	12079
VT2e	Verotoxin (EHEC)	F: 5′-CCAGAATGTCAGATAACTGGCGAC-3′ R: 5′-GCTGAGCACTTTGTAACAATGGCTG-3′	322	E40883
O111	Side-chain LPS (EHEC)	F: 5′-TAGAGAAATTATCAAGTTAGTTCC-3′ R: 5′-ATAGTTATGAACATCTTGTTTAGC-3′	406	97m 2716
O157	O157:H7 Side-chain LPS (EHEC)	F: 5′-CGGACATCCATGTGATATGG-3′ R: 5′-TTGCCTATGTACAGCTAATCC-3′	259	9602-5069
HlyA	A-haemolysin (UPEC)	F: 5′-GACAAAGCACGAAAGATG-3′ R: 5′-CAACTGCAATAAAGAAGC-3′	2930	J96
CNF1	Cytotoxic necrotizing factor 1 (UPEC)	F: 5′-GGCGACAAATGCAGTATTGCTTGG-3′ R: 5′-GACGTTGGTTGCGGTAATTTTGGG-3′	552	7/6/96, MAP
CNF2	Cytotoxic necrotizing factor 2 (UPEC)	F: 5′-GTGAGGCTCAACGAGATTATGCACTG-3′ R: 5′-CCACGCTTCTTCTTCAGTTGTTCCTC-3′	839	7/6/96, MAP
PapC	P fimbria (UPEC)	F: 5′-GACGGCTGTACTGCAGGGTGTGGCG-3′ R: 5′-ATATCCTTTCTGCAGGGATGCAATA-3′	328	J96
LT1	Heat labile toxin 1 (ETEC)	F: 5′-TGGATTCATCATGCACCACAAGG-3′ R: 5′-CCATTTCTCTTTTGCCTGCCATC-3′	360	0147:K89
STI	Heat-stable toxin 1 (ETEC)	F: 5′-TTTCCCCTCTTTTAGTCAGTCAACTG-3′ R: 5′-GGCAGGATTACAACAAAGTTCACAG-3′	160	11602
ST2	Heat-stable toxin 2 (ETEC)	F: 5′-CCCCCTCTCTTTTGCACTTCTTTCC-3′ R: 5′-TGCTCCAGCAGTACCATCTCTAACCC-3′	423	0149:K+K88

acetic acid, made up to 1 L with H_2O. Identification of the bands was established by comparison of the band sizes with molecular weight markers of 100 bp and 1 kb ladder (GeneWorks) after staining with ethidium bromide and exposure to UV light.

2.4. *Biochemical fingerprinting*

The principle of the biochemical fingerprinting with the PhPlate system (PhPlate AB, Stockholm, Sweden) has been previously described (Möllby et al., 1993). In this study, we used PhP-RE plates for typing of *E. coli* bacterial isolates. Reagents used in the PhP-RE plates have been described before (Kühn et al., 1995). The growth medium for PhP-RE contained 0.1% (w/v) proteose peptone (Oxoid) and 0.011% (w/v) bromothymol blue (Merck Pty Ltd, Australia) according to the manufacturer's instructions (PhPlate AB, Stockholm, Sweden). Sterile toothpick was used to pick the *E. coli* strains from the primary isolating agar plates and suspended into the first well of each row (containing 350 μl of growth medium). Aliquots of 25 μl of bacterial suspension were transferred into each of the other 11 wells (containing 150 μl growth medium) with the aid of a multi-channel pipette. Plates were then incubated at 37 °C and absorbance (A_{620}) was measured at 7, 24 and 48 h using a micro-plate reader (Lab-systems Multiskan, Helsinki, Finland). After the final reading the mean value for all three readings was calculated for each isolate (biochemical fingerprint) (Kühn et al., 1991).

The isolates were compared pair-wise and the obtained similarity matrix was clustered according to the unweighted-pair group method (UPGMA) method (Sneath and Sokal, 1973). An identity (ID) level of 0.965 was established for the system based on testing 10 isolates in duplicate. Isolates showing similarity to each other above the ID-level were regarded as identical and were assigned to a BPT. All data handling, including optical readings, calculations of similarities among fingerprints of strains as well as clustering and printing dendrograms, was performed using the PhPlate software version 4001 (PhPlate system, Stockholm, Sweden).

3. Results

3.1. *Prevalence of virulence genes*

In all, 39 (15%) out of the 262 representative strains from nine host groups carried one or more virulence genes. These included six strains from humans, two strains from horses,

Table 2 – Number of strains tested from different host groups and the number of strains carrying virulence genes

Host groups	No. of strains in collection	No. of strains tested	No. of strains carrying virulence genes (%)
Cattle	53	32	5 (15.6)
Chickens	59	30	7 (22.6)
Deer	31	12	3 (25.0)
Dogs	64	33	8 (24.2)
Ducks	69	32	2 (6.30)
Horses	60	30	2 (6.70)
Humans	92	36	6 (16.6)
Kangaroos	19	13	0 (0.00)
Pigs	53	31	4 (12.9)
Sheep	30	13	2 (15.4)
Total	530	262	39 (14.9)

eight from dogs, two from ducks, five from cattle, seven from chickens, four from pigs, two from sheep and three from deer. No virulence genes were identified among strains from kangaroos. Ten (10%) out of the 104 strains tested from water samples, also carried one or more virulence genes.

3.2. *Distribution of virulence genes*

Among the 39 strains carrying one or more virulence genes, 22 (56%) carried *eaeA*, 11 (28%) carried *papC*, 7 (18%) carried *hlyA*, 5 (13%) carried CNF1 and 2 (5%) carried CNF2, 8 (21%) carried O157 side-chain LPS, 6 (15%) carried VT1, 1 (3%) carried VT2 and 2 (5%) carried VT2e genes. Of these, 14 (36%) strains carried multiple virulence genes. Among the 10 strains carrying one or more virulence genes from water samples, six carried *eaeA* genes, three carried *papC*, three carried *hlyA*, two carried CNF1, four carried CNF2 and one carried O157 LPS. Eight strains carried multiple virulence genes (Table 3). Virulence genes EAgg, Einv, LT1, O111 LPS, ST1 and ST2 were not present among strains either from host groups or water samples.

3.3. *Comparison of strains*

Strains from host groups and water samples carrying single or multiple virulence genes were typed with the biochemical fingerprinting method. Cluster analysis was used to compare the BPTs of the 10 strains from water samples with those (i.e. 39 strains) found in the host groups to identify the likely source of these strains. Of these, 6 BPTs were identical to those of humans (two strains), dogs (two strains), chickens (one strain) and sheep (one strain) with 4 BPTs also carrying same virulence genes (i.e. CH241 vs. W94; HUM 63 vs. W 134; HUM 57 vs. W 208 and DO 90 vs. W 19) (Table 3). The other 4 BPTs carried virulence genes coding for CNF2 (i.e. W 72), *eaeA* and CNF2 (i.e. W 176), *papC*, *hlyA* and CNF1 (i.e. W 185) and O157 side-chain LPS (i.e. W 210) and were not identical to any of the BPTs from host groups.

4. Discussion

In this study, we postulated that fingerprints of *E. coli* strains from host groups and environmental samples carrying clinically significant virulence genes can be compared with each other to identify the likely sources of such strains in environmental samples. One important feature of such approach is that the analysis could be focused on strains carrying virulence genes rather than commensal populations of little significance. This also may reduce the overlapping effects of commensal *E. coli* among different host groups as reported previously (Ahmed et al., 2005a; Hartel et al., 2002). In addition, identical fingerprints and as well as the presence of single or multiple genes in two or more compared strains can boost the confidence level that the sources have been correctly assigned.

The strains used in this study were selected from a larger collection of strains used to identify the sources of faecal contamination in a previous MST study where 67% of the BPTs from water samples had identical BPTs to those found in different host groups (Ahmed et al., 2005a). However, the strains were randomly chosen to eliminate the chance of selecting identical BPTs between host groups and water samples which may bias the results. In all, 39 strains from host groups carried one or more virulence genes tested. Of interest was the positive PCR result for the *eaeA* gene, or a gene with some homology to *eaeA* among these animals. This gene has been associated with enteropathogenic *E. coli* (EPEC)/enterohaemorrhagic *E. coli* (EHEC) and is responsible for the attaching and effacing lesions in human enterocytes (Jerse et al., 1990). However, in the absence of any *in vivo* study it was not possible to determine whether strains positive for this gene were in fact capable of expressing themselves. Of the six strains from septic tanks (i.e. human origin) showing the presence of virulence genes, three carried either CNF1 or CNF2 genes. While CNF1 has shown to be associated with strains causing diarrhoea in cattle (Orden et al., 1999), these genes are frequently found among strains causing urinary tract infection and therefore, easier to interpret their presence in septic tanks, which receives wastewater from humans. Of interest also was the presence of genes coding for O157 side-chain LPS and VT1 in one strain from septic tank. EHEC can cause acute bloody diarrhoea, haemorrhagic colitis and the life-threatening haemolytic uremic syndrome in humans (Griffin and Tauxe, 1991). Among the animals, these genes were only found in cattle. A number of strains also carried *papC* and *hlyA*. These genes are normally found among *E. coli* strains causing urinary tract infections (Foxman et al., 1995). Indeed, data obtained in this study showed those *papC* genes were more distributed among strains from septic tanks and dog faeces than other animals. There is a possibility that dogs receive these strains from humans, as they are companion animals. In contrast, none of the strains carried ST1, ST2 and LT1 genes. *E. coli* strains carrying these genes are commonly found among cases of human and animal diarrhoea worldwide (Hart et al., 1993).

Of the 10 strains from water samples, six had identical BPTs to those found in different host groups with four strains also having similar virulence genes. Two strains from the water

Table 3 – *Escherichia coli* virulence genes detected in host groups and water samples

Strains source	Strains code	Virulence genes								
		CNF1	CNF2	*eaeA*	*hlyA*	O157 LPS	*papC*	VT1	VT2	VT2e
Cattle	C 29	—	—	—	—	—	—	VT1	—	VT2e
	C 65	—	—	*eaeA*	—	—	—	—	—	—
	C 67	—	—	*eaeA*	—	—	—	—	—	—
	C 74	—	—	*eaeA*	—	—	—	—	VT2	—
	C 172	—	—	*eaeA*	—	—	—	—	—	—
Chickens	CH 179	—	—	*eaeA*	—	—	—	—	—	—
	CH 192	—	—	—	—	O157 LPS	—	—	—	—
	CH 239	—	—	—	—	O157 LPS	—	—	—	—
	CH 240	—	—	*eaeA*	—	—	—	—	—	—
	CH 241[a]	—	—	*eaeA*	—	—	—	—	—	—
	CH 242	—	—	*eaeA*	—	—	—	—	—	—
	CH 245	—	—	—	—	—	*papC*	—	—	—
Deer	DE 160	—	—	—	—	—	—	VT1	—	—
	DE 162	—	—	—	—	—	—	VT1	—	—
	DE 163	—	—	—	—	—	—	VT1	—	—
Dogs	DO 90[b]	CNF1	—	*eaeA*	*hlyA*	—	*papC*	—	—	—
	DO 96	—	—	—	—	—	*papC*	—	—	—
	DO 105[c]	—	CNF2	—	*hlyA*	—	—	—	—	—
	DO 112	—	—	—	—	—	*papC*	—	—	—
	DO 113	—	—	*eaeA*	—	—	—	—	—	—
	DO 120	CNF1	—	—	—	—	—	—	—	—
	DO 133	—	—	—	—	O157 LPS	—	—	—	—
	DO 210	—	—	—	—	—	*papC*	—	—	—
Ducks	D77	—	—	*eaeA*	—	—	—	—	—	—
	D 255	—	—	*eaeA*	—	—	—	—	—	—
Horses	H 37	—	—	—	—	—	*papC*	—	—	—
	H 62	CNF1	—	*eaeA*	*hlyA*	—	*papC*	—	—	—
Humans	HUM 3	—	—	*eaeA*	—	—	*papC*	—	—	—
	HUM 19	CNF1	—	*eaeA*	*HlyA*	O157 LPS	—	—	—	—
	HUM 33	—	—	*eaeA*	—	—	—	—	—	—
	HUM 53	CNF1	—	*eaeA*	—	—	*papC*	—	—	—
	HUM 57[d]	CNF1	—	*eaeA*	—	—	*papC*	—	—	—
	HUM 63[e]	—	—	*eaeA*	—	—	*papC*	—	—	—
Pigs	P 183	—	—	*eaeA*	—	—	—	—	—	—
	P 271	—	—	*eaeA*	—	—	—	—	—	—
	P 273	—	—	—	*hlyA*	O157 LPS	—	—	—	VT2e
	P 279	—	—	—	*hlyA*	—	—	—	—	—
			—	—	—	—	—	—	—	—
Sheep	SH 184	—	—	*eaeA*	—	O157 LPS	—	VT1	—	–
	SH 195[f]	—	—	*eaeA*	—	—	—	—	—	—
Water samples	W 8[c]	—	CNF2	—	*hlyA*	—	—	—	—	—
	W 19[b]	—	—	*eaeA*	*HlyA*	—	—	—	—	—
	W 72	—	CNF2	—	—	—	—	—	—	—
	W 94[a]	—	—	*eaeA*	—	—	—	—	—	—
	W 121[f]	CNF1	—	*eaeA*	—	—	—	—	—	—
	W 176	—	CNF2	*eaeA*	—	—	—	—	—	—
	W 134[e]	—	—	*eaeA*	—	—	*papC*	—	—	—
	W 185	CNF1	—	—	*HlyA*	—	*papC*	—	—	—
	W 208[d]	—	CNF2	*eaeA*	—	—	*papC*	—	—	—
	W 210	—	—	—	—	O157 LPS	—	—	—	—

CNF1: cytotoxic necrotizing factor 1 gene; CNF2: cytotoxic necrotizing factor 2 gene; *eaeA*: attachment and effacement gene,; *hlyA*: hemolysin A gene; O157 LPS: O157 lipopolysaccharides; *papC*: pyelonephritis–associated pilli gene; VT: verotoxin. a vs. a, b vs. b, c vs. c, d vs. d, e vs. e and f vs. f: Identical biochemical phenotypes (BPTs).

samples were identical to the strains from septic tanks. These septic tanks were defective and identified as contributors of faecal contamination in a previous study (Ahmed et al., 2005b). To determine the stability, biochemical fingerprints of these 39 strains from this study were compared with the fingerprints of same strains in a previous study (Ahmed et al., 2005a). The fingerprints were found to be stable (data not shown).

We also compared the BPTs of all strains (i.e. 262 strains) from all host groups with the BPTs found in water samples alone, it was found that 11 (10.6%) BPTs were categorized as humans and 58 (55.8%) BPTs were categorized as animals and the remaining 55 (52.9%) BPTs could not be identified (data not shown). When we considered only virulence genes, it was found that 10 strains from water samples carried similar virulence gene(s) to those that were found in host groups (i.e. two from septic tanks and eight from animals). When we considered the combination of BPTs and the presence of virulence genes, it was found that only six strains (i.e. two from humans and four from animals) had the same BPTs and as well as same virulence genes. However, it has to be noted that virulence gene(s) could be found simultaneously in strains from different host groups. For example, strains such as CH 241 (i.e. chicken), C 74 (i.e. cattle) and P 274 (i.e. Pig) all carried *eaeA* gene. Therefore, presence of this gene in strain from water sample (i.e. W 94) could not be used alone to discriminate among above sources. However, biochemical fingerprint of CH 241 was identical to W 94 indicating chicken may be the likely source.

One disadvantage of such method is that the number of strains carrying *E. coli* virulence genes can be low such as O157:H7 is found at low frequencies in certain host groups such as cattle (Wells et al., 1991). Therefore, successful application of this method may require knowledge of the occurrence of these genes in host groups as well as testing a large number of isolates from surface waters. In addition, this method could be quite laborious due to detection of each virulence gene separately as performed in this study. Multiplex PCR or gene array can be used for the rapid detection of these genes among strains from host groups and water samples. The biochemical fingerprinting method used in this study offers a high resolution and stability of the BPTs as shown in other studies (Kühn et al., 1991, 1995) with the possibility of testing a larger number of isolates within a short time (Ahmed et al., 2005a, b; Möllby et al., 1993).

5. Conclusions

Our data suggest that biochemical fingerprinting of *E. coli* strains carrying virulence genes has the potential to identify the likely sources of pathogenic strains of *E. coli* in environmental samples. However, the number of strains carrying virulence genes could be low in certain host groups and therefore, the successful application of this method will require testing large number of isolates from host groups and surface waters, and using multiplex PCR or gene array technology for the rapid detection of virulence genes. Nonetheless, *E. coli* strains with identical BPTs and virulence genes in host groups and water samples indicate that the combination of biochemical fingerprinting and PCR detection of virulence genes can be used to predict the sources of pathogenic strains in surface waters. The information generated from such studies can be of great importance in assisting public health authorities and environmental protection agencies to implement proper management strategy for improving the quality of surface waters.

Acknowledgments

We thank Mr. Chris Gaham for his assistance in processing samples. We would also like to thank Mr. Alexander Kuzevski for assistance in laboratory work. This study was partially funded by an internal grant, SRG03/5, from the University of the Sunshine Coast.

REFERENCES

Ahmed, W., Neller, R., Katouli, M., 2005a. Host species-specific metabolic fingerprint database for enterococci and *Escherichia coli* and its application to identify sources of fecal contamination in surface waters. Appl. Environ. Microbiol. 71 (8), 4461–4468.

Ahmed, W., Neller, R., Katouli, M., 2005b. Evidence of septic systems failure determined by a bacterial biochemical fingerprinting method. J. Appl. Microbiol. 98 (4), 910–920.

Boyd, E.F., Hartl, D.L., 1998. Chromosomal regions specific to pathogenic isolates of *Escherichia coli* have a phylogenetically clustered distribution. J. Bacteriol. 180 (5), 1159–1165.

Carson, S.A., Shear, B.L., Ellersieck, M.R., Afsaw, A.M.H.A., 2001. Identification of *Escherichia coli* from humans and animals by Ribotyping. Appl. Environ. Microbiol. 67 (4), 1503–1507.

Echeverria, P., Seriwatana, P.J., Patmaroj, U., Mosley, S.L., McFarland, A., Chityothin, O., Chaicumpa, W., 1984. Prevalence of heat stable II enterotoxigenic *Escherichia coli* in pigs, water and people at farms in Thailand determined by DNA hybridization. J. Clin. Microbiol. 19 (4), 489–491.

Falagas, M., Gorbach, S., 1995. Practice guidelines: urinary tract infections. Infect. Dis. Clin. Pract. 4, 241–257.

Feldman, K.A., Mohle-Boetani, J.C., Ward, J., Furst, K., Abbott, S.L., Ferrero, D.V., Olsen, A., Werner, S.B., 2002. A cluster of *Escherichia coli* O157: non-motile infections associated with recreational exposure to lake water. Public Health Rep. 117 (4), 380–385.

Foxman, B., Zhang, L., Palin, K., Tallman, P., Marrs, S., 1995. Bacterial virulence characteristics of *Escherichia coli* isolates from first-time urinary tract infection. J. Infect. Dis. 171 (6), 1514–1521.

Fujioka, R.S., 2002. Microbial indicators of water quality. In: Hurst, C.J., Crawford, R.L., Knudsen, G.R., McInereny, K.J., Stetzenbach, L.D. (Eds.), Manual of Environmental Microbiology, second ed. ASM Press, Washington, DC, pp. 234–243.

Gilpin, B.J., Gregor, J.E., Savill, M.G., 2002. Identification of the source of faecal pollution in contaminated rivers. Wat. Sci. Technol. 46 (3), 9–15.

Gordon, D.M., 2001. Geographical structure and host specificity in bacteria and the implications for tracing the source of coliform contamination. Microbiology 147 (pt5), 1079–1085.

Griffin, P.M., Tauxe, R.V., 1991. The epidemiology of infections caused by *Escherichia coli* O157:H7, other Enterohemorrhagic *E. coli*, and the associated haemolytic uremic syndrome. Epidemiol. Rev. 13, 60–98.

Guth, B.E., Pickett, C.L., Twiddy, E.M., Holmes, R.K., Gomes, T.A., Lima, A.A., Guerrant, R.L., Franco, B.D., Trabulsi, L.R., 1986.

Production of type II heat labile enterotoxin by *Escherichia coli* isolated from food and human faeces. Infect. Immun. 54 (2), 587–589.

Hagedorn, C., Crozier, J.B., Mentz, K.A., Booth, A.M., Graves, A.K., Nelson, A.J., Reneau Jr., R.B., 2003. Carbon source utilization profiles as a method to identify sources of faecal pollution in water. J. Appl. Microbiol. 94 (5), 792–799.

Hart, C.A., Batt, R.M., Saunders, J.R., 1993. Diarrhoea caused by *Escherichia coli* Ann. Trop. Paediat. 13 (2), 121–131.

Hartel, P.G., Summer, J.D., Hill, J.L., Collins, J.C., Entry, J.A., Segers, W.I., 2002. Geographic variability of *Escherichia coli* isolates from animals in Idaho and Georgia. J. Environ. Qual. 31 (4), 1273–1278.

Jerse, A.E., Yu, J., Tall, B.D., Kaper, J.P., 1990. A genetic locus of enteropathogenic *Escherichia coli* necessary for the production of attaching and effacing lesions on tissue culture cells. Proc. Natl. Acad. Sci. USA. 87 (20), 7839–7843.

Kühn, I., Allestam, G., Strenström, T.A., Möllby, R., 1991. Biochemical fingerprinting of water coliform bacteria—a new method for measuring the phenotypic diversity and for comparing different bacterial populations. Appl. Environ. Microbiol. 57 (11), 3171–3177.

Kühn, I., Katouli, M., Wallgren, P., Söderlind, O., Möllby, R., 1995. Biochemical fingerprinting as a tool to study the diversity and stability of intestinal microflora. Microecol. Therap. 23, 40–148.

Lai, X., Wang, S.Y., Uhlin, B.E., 1999. Expression of cytotoxicity by potential pathogens in the standard *Escherichia coli* collection of reference (EROR) strains. Microbiology 145 (pt11), 3295–3303.

Möllby, R., Kühn, I., Katouli, M., 1993. Computerized biochemical fingerprinting—a new tool for typing of bacteria. Rev. Med. Microbiol. 14, 231–241.

Nataro, J.P., Kaper, J.B., 1998. Diarrhoeagenic *Escherichia coli*. Clin. Microbiol. Rev. 11 (1), 142–201.

Olsen, S.J., Miller, G., Breuer, T., Kennedy, M., Higgins, C., Walford, J., McKee, G., Fox, K., Bibb, W., Mead, P.A., 2002. Waterborne outbreak of *Escherichia coli* O157 infections and hemolytic uremic syndrome: implications for rural water systems. Emerg. Infect. Dis. 8 (4), 370–375.

Orden, J.A., Ruiz-Santa-Quiteria, J.A., Cid, D., Garcia, S., Lafuenter, D.E., 1999. Prevalence and characteristics of necrotoxigenic *Escherichia coli* (NTEC) strains isolated from diarrhoeic calves. Vet. Microbiol. 66 (4), 265–273.

Pass, M.A., Odedra, R., Batt, R.M., 2000. Multiplex PCRs for identification of *Escherichia coli* virulence genes. J. Clin. Microbiol. 38 (5), 2001–2004.

Paton, A.W., Paton, J.C., 1998. Detection and characterization of shiga toxigenic *Escherichia coli* by using multiplex PCR assay for stx_1, stx_2, *eae*A, Enterohemorrhagic *E. coli hly*A, rfb_{O111}, and rfb_{O157}. J. Clin. Microbiol. 36 (2), 598–602.

Sneath, P.H.A., Sokal, R.R., 1973. Numerical Taxonomy: The Principles and Practice of Numerical Classification. W.H. Freeman & Co., San Francisco.

Tillett, D., Neilan, B.A., 2000. Xanthogenate nucleic acid isolation from cultured and environmental cyanobacteria. J. Phycol. 35 (1), 251–258.

Wells, J.G., Shipman, L.D., Greene, K.D., Sowers, E.G., Green, J.H., Cameron, D.M., Downes, F.P., Martin, M.L., Griffin, P.M., Ostroff, S.M., Potter, M.E., Tauxe, R.V., Wachsmuth, I.W., 1991. Isolation of *Escherichia coli* serotype O157:H7 and other Shiga-like-toxin-producing *E. coli* from dairy cattle. J. Clin. Microbiol. 129 (5), 985–989.