Available online at www.sciencedirect.com

SciVerse ScienceDirect

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

Escherichia coli virulence genes profile of surface waters as an indicator of water quality

N. Masters [a], A. Wiegand [a], W. Ahmed [a,b], M. Katouli [a,*]

[a] Faculty of Science, Health and Education, University of the Sunshine Coast, Maroochydore DC, Queensland 4558, Australia
[b] CSIRO Land and Water, Ecosciences Precinct, 41 Boggo Road, Brisbane, Queensland 4067, Australia

ARTICLE INFO

Article history:
Received 12 April 2011
Received in revised form
17 August 2011
Accepted 5 September 2011
Available online 21 September 2011

Keywords:
Escherichia coli
Virulence factors
Surface water
Indicator of water quality

ABSTRACT

We compared the presence of 58 known virulence genes (VGs) associated with *Escherichia coli* strains causing intestinal (InPEC) and extra-intestinal (ExPEC) infections in three estuarine, four brackish and 13 freshwater sites during the dry and wet seasons. The most common VGs observed in water samples during the dry season belonged to ExPEC (*traT*; 80% and *ompA*; 70%) whilst *east1* (70%) gene was the most common among InPEC. More types of VGs were observed in water samples during wet season and included those found among InPEC (e.g. *eaeA*; 100%; *fyuA*, 90%; *paa*, 65%; *cdt*, 60%; and stx_2, 60%) and ExPEC (e.g. $iroN_{E.coli}$, 90%; *iss*, 90% and *kpsMTII*, 80%). Eight VGs were found exclusively in the wet season, of which four were found in all three water types indicating their association with storm-water run off. The number of VGs associated with ExPEC were significantly ($P < 0.05$) higher in only brackish and estuarine waters during the wet season compared to the dry season. There was no correlation between the number of *E. coli* and the presence of VGs in any of the water types in both seasons but we found similarities in VG profiles of sites with similar land uses.

1. Introduction

Human enteric pathogens such as *Escherichia coli* O157:H7 (Ibekwe and Grieve, 2003), *Salmonella* spp. (Savichtcheva et al., 2007), *Cryptosporidium* spp. (Hörman et al., 2004) and enteric viruses (Haramoto et al., 2005) have been found in environmental waters as a result of faecal pollution from point and non-point sources (O'Shea and Field, 1992; Aslan-Yilmaz et al., 2004; Ahmed et al., 2005a). Due to the complexities associated with attempting to detect all possible pathogens in a water source, faecal indicator bacteria such as enterococci and *E. coli* have long been used by water industries to assess the microbiological quality of environmental and drinking waters. However recent studies suggest a limited relationship between faecal indicator bacteria and their ability to accurately predict the presence of pathogens, especially protozoans and enteric viruses (Hörman et al., 2004; McQuaig et al., 2006). In addition, it has been reported that faecal indicator bacteria such as *E. coli* can be present in the environment in the absence of faecal pollution and regrowth is possible under suitable conditions (Desmarais et al., 2002 Power et al., 2005).

E. coli strains are normal inhabitants in the gut of warm-blooded animals including humans (Hart et al., 1993). Whilst most *E. coli* strains in the gut are non-pathogenic commensals, certain strains may carry a combination of virulence genes (VGs) 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 uraemic syndrome, urinary tract and surgical

* *Corresponding author*. Tel: +61 7 54304528, fax: +61 7 54302887.
E-mail address: mkatouli@usc.edu.au (M. Katouli).
0043-1354/$ – see front matter
doi:10.1016/j.watres.2011.09.018

site infections (Falagas and Gorbach, 1995; Johnson and Stell, 2000). Pathogenic *E. coli* strains can be classified into intestinal (InPEC) and extra-intestinal (ExPEC) on the basis of their virulence factors and clinical symptoms (Kaper et al., 2004). InPEC can be further classified into enterotoxigenic (ETEC), enteropathogenic (EPEC), enterohemorrhagic (EHEC), enteroinvasive (EIEC), and enteroaggregative (EaggEC) *E. coli* and diffusely adherent *E. coli* (DAEC) (Nataro and Kaper, 1998). InPEC can be further classified into enterotoxigenic (ETEC), enteropathogenic (EPEC), enterohemorrhagic (EHEC), enteroinvasive (EIEC), and enteroaggregative (EaggEC) *E. coli* and diffusely adherent *E. coli* (DAEC) (Nataro and Kaper, 1998; Kaper et al., 2004; Ishii et al., 2007). Common reservoirs of ETEC and EPEC include ruminants, porcine, other domesticated animals including goats, dogs, cats and humans (Levine, 1987; Nataro and Kaper, 1998; Djordjevic et al., 2004). EHEC have been isolated from various ruminants, primarily cattle (Paton and Paton, 1998). The principal reservoir for EIEC, EaggEC and DAEC are humans (Levine, 1987; Kaper et al., 2004). ExPEC can further be classified into uropathogenic *E. coli* (UPEC), meningitis neonatal *E. coli* (MNEC) and avian pathogenic *E. coli* (APEC); the two former pathotypes are commonly isolated from humans, APEC are associated with avian infections and have been isolated from poultry (Johnson and Stell, 2000; Kaper et al., 2004). It has to be noted that the possession of a single or multiple VGs does not necessarily indicate that a strain is pathogenic unless that strain has the appropriate combination of VGs to cause disease in a specific host (Gilmore and Ferretti, 2003).

Faeces from domestic and wild animals as well as humans may contain high numbers of *E. coli* strains harbouring one or more VGs (Ishii et al., 2007). Runoff from agricultural lands and sewer overflows may also contribute pathogenic *E. coli* strains containing these VGs into environmental waters. Pollution of surface waters with pathogenic strains of *E. coli* has been implicated in an increased number of disease outbreaks and consequent deaths (Feldman et al., 2002; Olsen et al., 2002). However, only a few studies have investigated the presence of *E. coli* strains carrying VGs in environmental waters (Martins et al., 1992; Lauber et al., 2003; Chern et al., 2004; Hamelin et al., 2006; Ahmed et al., 2007; Ram et al., 2007; Hamilton et al., 2010), and not all of them to the extent reported in this study.

The objectives of this study were to a) observe the prevalence of *E. coli* VGs in waters with contrasted land use and various fecal sources and b) to compare these VG profiles between sites to define the sources of fecal pollutions. Furthermore, correlations between the presence of VGs and the number of *E. coli* were also investigated.

2. Materials and methods

2.1. Sampling sites

Water samples were collected from three estuaries (EW1–EW3) with salinity ranging between 21 and 34 PSU, four rivers with brackish water (BW1–BW4) salinity range between 3 and 14 PSU and 13 creeks with fresh water (FW1–FW13) (salinity $<$ 1 PSU) in the Sunshine Coast region located approximately 100 km north of Brisbane, the capital city of Queensland, Australia. Fig. 1 shows the location of all sampling sites and their catchment area. In all, 40 water samples were collected from 20 sites during the dry season of 2009 ($n = 20$) and wet season of 2010 ($n = 20$). Table 1 shows the location, land uses and potential sources of faecal pollution of each sampling site. These sites were selected after consultation with local council and a community water quality monitoring group, who identified these sites as being suspected of having poor water quality and or previously found to have high numbers of coliforms. All samples were collected during low tide. Samples were collected during the dry season when the study areas received either no rainfall for more than 1 month or not more than 2 mm for at least 15 days prior to sampling. In contrast, samples were collected during the wet season when the sampling sites had received more than 100 mm rainfall two days prior to sampling. From each site, grab water samples were collected in 5 L sterile bottles from 30 cm below the water surface and transported on ice to the laboratory where they were processed within 6 h of collection.

2.2. Enumeration of E. coli

The membrane filtration method was used to process all water samples. Serial dilutions of samples were filtered through 0.45 μm pore size nitrocellulose membranes (Millipore, Tokyo, Japan), and placed on chromocult® coliform agar (Merck, Germany) plates and incubated at 35° C for 2 h to revive cells and then 44.5 °C for 22 h.

2.3. DNA extraction

For DNA extraction a 1 L replicate of each sample was filtered through 0.45 μm pore size membranes (Millipore), and *E. coli* were grown on chromocult® coliform agar (Merck, Germany). After incubation for 35° C for 2 h to revive cells and then 44.5° C for 22 h, filters were transferred to sterile flasks containing 50 ml of tryptic soy broth (TSB) (Oxoid, London, UK), and incubated at 44.5° C for another 24 h. This enrichment step was included to ensure low numbers of *E. coli* containing VGs genes can also be identified. DNA was extracted from 2 ml culture using DNeasy blood and tissue kit (Qiagen) according to the manufacture instructions, and the resulting DNA extracts were stored at −20 °C until use.

2.4. PCR detection of virulence genes

Using a combination of multiplex and single PCR, a total of 58 VGs associated with *E. coli* strains causing intestinal and extra-intestinal infections were selected for this study according to Chapman et al. (2006). Primers used for PCR detection of these VGs, the size of the PCR products for corresponding targets and PCR cycling conditions have been previously reported (Johnson and Stell, 2000; Chapman et al., 2006). For each PCR experiment, corresponding positive DNA, and negative control (sterile water) were included. The PCR amplification was performed using the Eppendorf Master Cycler (Eppendorf, Germany). To detect the amplified product, up to 5–8 μl aliquot of the PCR product was visualized by electrophoresis through a 2% agarose gel (Progen, Australia) in 0.6 × TBE buffer and ethidium

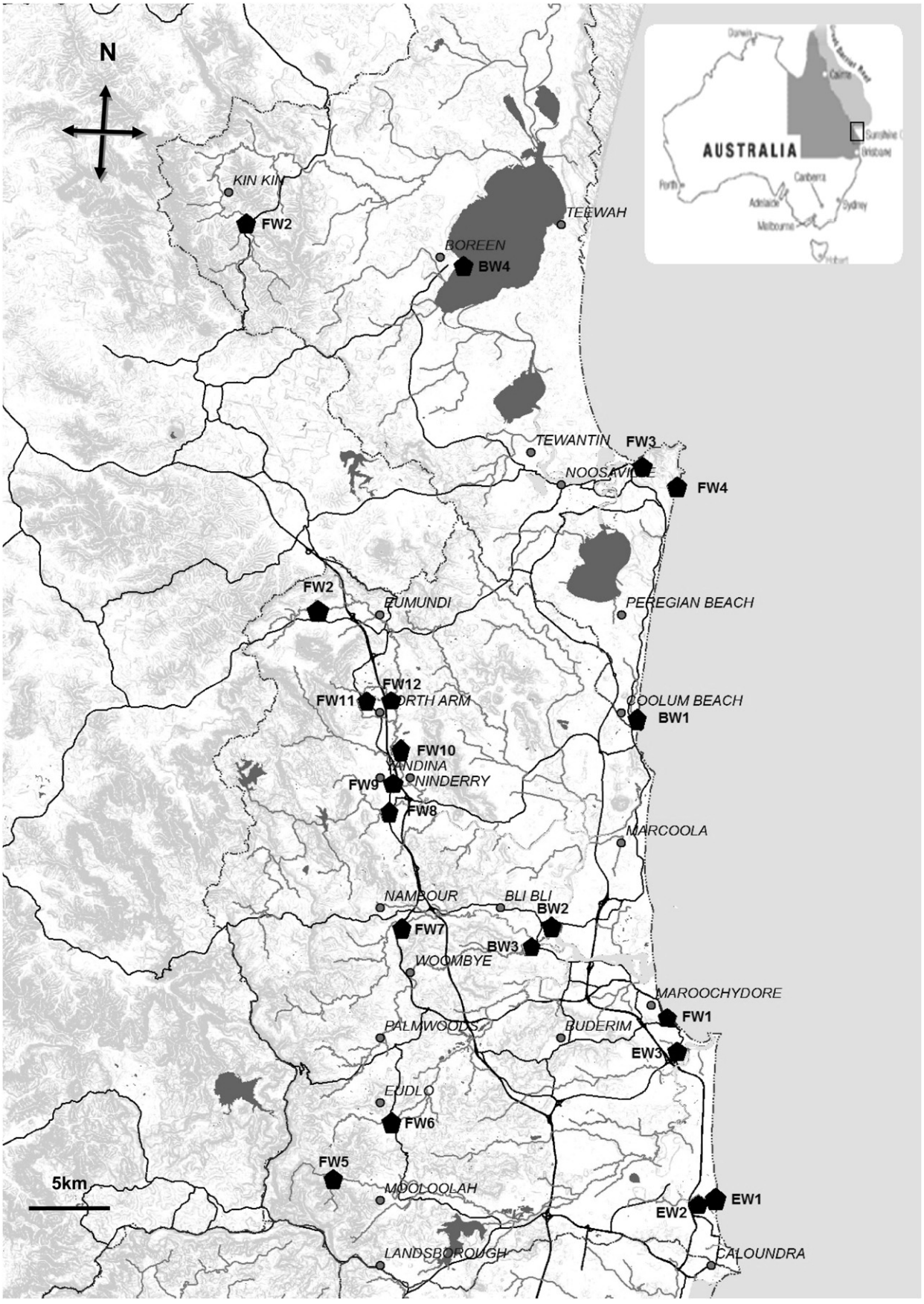

Fig. 1 – Map of sunshine coast region with location of sampling sites; EW1–EW3 indicates the estuarine water sites, BW1–BW4 indicates the brackish water sites, FW1–FW13 indicates the freshwater sites. Black lines indicate individual catchment areas.

bromide. Identification of the bands was established by comparison of the band sizes with molecular weight markers of 100 bp and 3 kb ladder (Gene-Works, Australia). Samples were considered to be positive for a specific VG when the visible band was the same size as that of the positive control DNA. To minimize PCR contamination, DNA extraction, PCR set up, and gel electrophoresis were performed in isolated rooms. To identify false positive results filtered milliQ water was used as a negative control in all experiments.

2.5. *Data analyses*

Virulence genes were initially categorized based on their functions and/or their (e.g. adhesion genes, toxin genes etc.) and/or different pathotypes (Table 2). Two separate analyses were undertaken in order to determine the impact of (i) season and (ii) water salinity on the presence or absence of genes within each gene category and also within each *E. coli* pathotype. Numbers and/or proportions of genes that tested positive in samples from comparative categories (e.g.: dry vs. wet) were compared using Z-tests.

The proportions of genes observed across sets of data were compared using a Z-test with the null hypothesis being that the proportions of positive observed genes during the dry and wet seasons were not significantly different.

Since the sampling locations were the same in both the dry and the wet seasons, a direct comparison between the presence and absence of individual genes at each location in each season could be made. The absolute difference between two corresponding sets of a particular gene category was defined as the total number of instances, in which each gene was present in a sampling location during the dry season and absent in the wet season, or vice-versa.

A linear regression analysis was applied to investigate the degree of correlation between the number of indicator bacteria and the number of VGs observed at each site.

A VG fingerprint was developed for each water sample by giving a value of 1 to each gene found in a water sample and a value of 0 for their absence. Using the PhPlate software version 4.0, similarities among VG profiles were measured after pair wise comparison of the VG fingerprint of each water sample and the obtained similarity matrix was clustered according to the un-weighted pair-group method using arithmetic averages (UPGMA) to yield a dendrogram. In the dendrogram each line represents VG profile of a water sample and they are connected to each other at the similarity level shown.

The student *t*-test was used to compare the significance of difference between the number of *E. coli* and the observed VGs numbers in samples collected during dry and wet seasons and between water types.

3. Results

The number of *E. coli* detected in water samples collected from the estuarine sites (EW1–EW3) during the dry and wet conditions ranged from 2 to 24 CFU/100 mL and 260 to 1600 CFU/100 mL, respectively (Table 3). For the brackish water sites (BW1–BW4), *E. coli* numbers ranged from 1 to 40 CFU/100 mL (dry conditions) and 100 to 5000 CFU/100 mL

Table 1 – List of the 20 sampling sites chosen for this study and their location, water characteristics and the suspected sources of faecal contamination impacting each site.

Sampling sites	Location	Land use	Suspected sources of faecal pollution
Estuarine water			
EW1	Golden Beach[b]	Urban	Urban run off, dogs, coastal birds
EW2	Currimundi[b]	Urban	Sewer overflows, waterfowls, dogs, coastal birds
EW3	Mooloolaba[b]	Urban	Urban run off, dogs, coastal birds
Brackish water			
BW1	Coolum[b]	Urban	Urban run off dogs, coastal birds
BW2	Bli Bli[b]	Peri-urban	Sewer overflows, wild animals
BW3	Nambour/Bli Bli[b]	Peri-urban	Septic systems, cattle, horses, wild animals
BW4	Boreen Point[b]	Urban	Septic systems, waterfowls, wild animals
Freshwater			
FW1[a]	Alexandra Beach[b]	Urban	Urban run off, dogs, coastal birds
FW2	Eumundi	Pasture	Septic systems, cattle, horses
FW3[a]	Noosa[b]	Urban	Urban run off, coastal birds
FW4[a]	Sunshine Beach[b]	Urban	Urban run off, dogs, coastal birds
FW5	Mooloolah[b]	Pasture	Septic systems, cattle, horses, wild animals
FW6	Eudlo	Peri-urban	Septic systems, cattle, horses, wild animals
FW7	Nambour[b]	Urban	Urban run off, cattle, horses, wild animals
FW8	Yandina	Peri-urban	Septic systems, cattle, horses, wild animals
FW9	Yandina	Pasture	Septic systems, cattle, horses, wild animals
FW10	North arm (east)	Pasture	Septic systems, cattle, horses, wild animals
FW11	North arm (west)	Pasture	Septic systems, cattle, horses, wild animals
FW12	Eumundi/North arm	Pasture	Septic systems, cattle, horses, wild animals
FW13	Kin Kin	Peri-urban	Septic systems, cattle, horses, wild animals

a Samples were collected from storm water outlets draining into the estuarine waters.
b Sampling site is used for swimming, recreational and social activities.

Table 2 – List of 58 virulence genes screened in this study, their categories based on their function and *E. coli* pathotypes carrying these genes.

E. coli pathotypes	Virulence genes categories					
	Adhesins	Toxins	Capsule synthesis	Siderophores	Invasins	Non-categorised virulence genes
ExPEC	*papAH* *fimH* *papEF* *bmaE* *sfa/focDE* *papG* allele II & III *nfaE* *papC* *focG* *papG* allele II *papG* allele III *Afa/draBC* *sfaS*	*hlyA* *cvaC* *cdtB* *cnf1* UNIV*cnf* *cdt*[a]	*kpsMTIII* *kpsMT* K1 *rfc* *kpsMT* II *kpsMT* K5 *ireA*	*fyuA* *iutA* *iroN*$_{E.coli}$	*ibeA*	PAI *traT* *ompT* *iss* *yjaA* TSPE4C2
DAEC	*aah (orfA)* *aidA* AIDA-I *(orfB)* *aidA* AIDA[c]					
EHEC	*iha* *eaeA*[a] *saa*	*exhA*[a] *stx*$_2$ *stx*$_1$				*chuA*
ETEC	*fasA* *faeG* *fanC* *fedA* F41	*eltA* *estI* *estII*				
EPEC	*paa* *eaeA*[a] *bfpA*	*exhA*[a] *cdt*[a]				
EaggEC		*east1*				
EIEC					*ipaH*	

a Indicates genes shared by more than one *E. coli* pathotype.

(wet conditions), respectively (Table 3). For the freshwater sites, these values were 2 to >5000 (dry conditions) and 28 to 5000 (wet conditions), respectively (Table 3). Generally the numbers of *E. coli* in the estuarine and brackish waters were higher during the wet conditions than those collected during dry conditions, however due to the low number of samples collected from each water type and variations within samples it was not possible to calculate the significance of these findings. This was more pronounced among fresh water samples which showed a higher variability between different sites, with 6 sites having a higher number of *E. coli* in the dry conditions compared to wet conditions (Table 3).

3.1. *Prevalence of* E. coli *VGs*

During the dry season 18 (90%) samples were positive for multiple VGs ranging from 4 to 22 genes per water sample (Table 4). Only freshwater sites were positive for >20 VGs during dry season. No VG was detected in samples collected in brackish waters designated as BW3 and BW4 during the dry season. The main source of faecal contamination of these sites was septic systems. Our previous study in this region has identified farm animals and septic systems as potential sources of VGs in these waterways. However, recent upgrading of septic systems within the regional area could be one of the reasons why no VGs were detected at these sites. Alternatively, it could be that *E. coli* strains found in these sites did not carry any VGs, this is supported by the low number of *E. coli* strains isolated from these sites (Table 3).

In total, 39 (67%) of the 58 VGs tested were detected in all water samples collected during this season (Table 4). The most commonly observed VGs were *fim*H (85%), *chu*A (85%), TSPE4C.2 (85%), *tra*T (80%), *Omp*T (70%), *east*1 (70%), *fyu*A (65%), and *ibe*A (60%) and all except *fim*H which is commonly found in all *E.coli* strains and *chu*A whch is commonly seen in EHEC pathotype, the rest are common among EXPEC strains (Fig. 2). Among the toxin genes tested, *east*1 (70%) was the most prevalent gene observed followed by *cdt* (40%), *cva*C (35%), *stx*$_1$ (25%), *stx*$_2$ (25%) and *est*I (5%). The remaining toxin genes *hly*A, *cdt*B, *cnf*1, UNIV*cnf*, *exh*A, *elt*A, *est*II could not be detected in all water samples collected during dry season.

During the wet season, all water samples were positive for multiple VGs ranging from 7 to 30 genes per water sample. Of the 58 VGs tested, 44 (76%) genes were observed in water

Table 3 – Number of *E. coli* in water samples collected during the dry and wet conditions and the corresponding number of virulence genes detected by PCR in DNA extracted from 2 mL aliquots of enriched cultures obtained from each sampling site.

Sampling sites	Number of *E. coli* (CFU/100 mL)		Number of positive observations of virulence genes	
	Dry conditions[a]	Wet conditions[b]	Dry conditions[a]	Wet conditions[b]
Estuarine water				
EW1	10	260	12	21
EW2	2	1600	14	29
EW3	24	420	6	27
Mean ± SD	12 ± 11.1	760 ± 731.8	10.7 ± 4.2[c]	25.7 ± 4.2[c]
Brackish water				
BW1	<1	100	4	26
BW2	22	>5000	17	19
BW3	40	860	0	26
BW4	38	270	0	16
Mean ± SD	33.3 ± 9.9	410 ± 398.9	5.3 ± 8.1[d]	21.8 ± 5.1[d]
Freshwater				
FW1	2	>5000	4	17
FW2	260	610	20	17
FW3	660	80	17	21
FW4	2500	160	14	19
FW5	920	2400	21	27
FW6	290	>5000	22	7
FW7	>5000	520	19	17
FW8	30	140	5	25
FW9	76	100	12	24
FW10	120	28	12	17
FW11	>5000	140	15	17
FW12	150	370	19	21
FW13	140	100	20	19
Mean ± SD	468 ± 730.6	422.5 ± 683.2	15.4 ± 5.8	19.1 ± 5

a study area received no rainfall 15 days prior sampling.
b study area received > 100 mm rainfall 2 days prior sampling.
c,d $P < 0.05$ for the number of VGs observations between water samples in dry and wet seasons.

samples collected during the wet conditions (Table 4). Except the *fim*H gene which was observed in all samples, the most commonly observed VGs belonged to both ExPEC e.g. *iroN*$_{E.coli}$, *iss*, *fyu*A, TSPE4C.2, *yja*A, *kps*MTII, PAI, *tra*T, *omp*T and IPEC e.g. *stx*$_2$, *chu*A, *saa* (found in EHEC pathotype), F41 (found in ETEC pathotype), *east1* (found in EAggEC pathotype), and *paa* and *eae*A (in EPEC pathotype). Among the toxin genes detected, *east1* (95%) was most prevalent gene followed by *cdt* (60%), *stx*$_2$ (60%), *ehx*A (45%), *cva*C (40%), *cdt*B (30%) *est*I, *elt*A and UNIV*cnf* (10%), *cnf*1 and *stx*$_1$ (5%) (Fig. 2). There was no positive observations of VGs associated with EIEC throughout the study and was therefore omitted from subsequent analysis. Comparison between the observed VGs in samples collected during the dry and wet seasons is given in Fig. 2.

Whilst the overall numbers of VGs observed in water samples collected during the wet season was significantly ($P < 0.05$) higher than those collected during the dry season, their prevalence was higher in fresh water sites FW2, FW6, FW7 and FW13 during the dry season than their corresponding sites during the wet season (Table 3). While some sites (i.e. FW7) had an exceptionally high number of bacteria during the dry season (Table 3) which would suggest an equally proportional number of VGs, other sites had lower numbers of *E. coli* and yet had a higher number of VGs (i.e. FW9) during the wet season indicating that the presence of VGs in these samples were not proportional to the number of *E. coli* strains. This could be partly explained by the diverse sources of faecal pollution, rather than hydrological conditions, within the proximal areas of those sampling sites.

VGs were grouped according to their association with different *E. coli* pathotype and further classified based on functional characteristics of the gene into toxin genes (Tox), adhesion genes (Adhes), capsule genes (Cap), iron acquisition genes (Iron), invasion genes (Inv) and non-classified (NC) VGs (Fig. 3). This enabled us to identify the prevalence of different pathotypes of *E. coli* in the water samples. We observed a significant ($P < 0.05$) difference in the presence of genes belonging to pathotypes EHEC, ETEC and EPEC pathotypes in the freshwater samples during the dry and wet seasons (Fig. 3A). The brackish (Fig. 3B) and estuarine (Fig. 3C) water samples on the other hand contained a significantly ($P < 0.05$) higher number of genes belonging to ExPEC, EHEC, EPEC and EaggEC during the wet season indicating that the brackish and estuarine sites may have similar sources of fecal pollutions. Overall, there was a significantly ($P < 0.05$) higher number of positive gene observations at all three water types and for all pathotypes (except DAEC) during the wet season compared to the dry season (Fig. 3D).

Grouping of VGs based on their functional characteristics showed striking differences in each water type during the dry

Table 4 – Virulence genes observed (indicated by ●) in at least one DNA extract of 2 mL enriched sample of fresh (FW), brackish (BW) and estuarine (EW) sampling sites in samples collected during dry and wet season. Boxed area denote virulence genes found only in the dry season samples while, grey area denotes virulence genes found only in the wet season samples.

		DRY SEASON			WET SEASON					DRY SEASON			WET SEASON		
		FW	BW	EW	FW	BW	EW			FW	BW	EW	FW	BW	EW
ExPEC	*PAI*	●	●	●	●	●	●		*ompT*	●	●	●	●	●	●
	papAH		●	●		●	●		*iss*	●	●		●	●	●
	fimH	●	●	●	●	●	●		*ireA*		●		●	●	●
	kpsMTIII	●			●	●	●		*yjaA*	●	●	●	●	●	●
	papEF				●	●	●		TSPE4C2	●	●	●	●	●	●
	ibeA	●		●	●	●	●	EHEC	*ehxA**				●	●	●
	fyuA	●	●	●	●	●	●		*iha*	●		●	●		●
	bmaE	●							*eaeA**	●			●	●	●
	sfa/focDE	●			●	●	●		*saa*	●			●	●	
	iutA	●	●		●	●	●		*stx1*	●				●	
	papG allele III	●			●	●	●		*stx2*	●			●	●	●
	kpsMTk1			●	●				*chuA*	●	●	●	●	●	●
	hlyA							ETEC	*eltA*				●		
	rfc	●							*fasA*						
	nfaE								*estII*						
	kpsMTII	●	●	●	●	●	●		*faeG*			●	●		
	papC	●	●	●	●	●	●		*fanC*			●	●		●
	cvaC	●			●	●	●		*estI*		●		●		
	cdtB				●		●		*fedA*						
	focG	●			●	●	●		*F41*	●			●		●
	traT	●	●	●	●	●	●	EPEC	*ehxA**				●	●	●
	papG allele II	●	●	●	●		●		*eaeA**	●			●	●	●
	papG allele II &III			●					*cdt**	●	●		●	●	●
	afa/draBC	●							*paa*				●	●	●
	cnf1					●			*bfpA*						
	sfaS							DAEC	*aidaA AIDAc (orfBc)*			●			
	kpsMT K5	●			●				*aidA AIDA-I (orfB)*						
	UNIVcnf					●	●		*aah (orfAa)*						
	*cdt**	●	●		●	●	●	EaggEC	*eastI*	●	●	●	●	●	●
	allele 1'							EIEC	*ipah*						
	iroN E.coli	●	●	●	●	●	●								

and wet seasons (Fig. 4). There was a significant ($P < 0.05$) difference between the number of VGs observed during the dry and the wet season for all water types (Fig. 4). However, fresh water samples shared a high number of VGs in both season as opposed to brackish and estuarine waters which contained a high number of VGs unique to the wet season (Fig. 4). This could be explained by continuous input from specific source(s) such as domestic and native animals at these sites. Brackish and estuarine waters, in contrast had higher numbers of suspected pathotypes in the wet season (Fig. 3) resulting in higher number of VGs that are normally associated with pathotypes EHEC, ExPEC, EPEC and EaggEC.

3.2. *Correlation between the numbers of* E. coli *and the presence of virulence genes*

A linear regression analysis was applied to calculate any possible correlation between the number of *E. coli* and the presence of VGs during the dry and wet season. It was found that the number of *E. coli* in water samples did not correlate to the number of positive VG observations during the dry ($P > 0.9612$) or the wet ($P > 0.2751$) seasons (Fig. 5).

3.3. *Comparison of* E. coli *virulence gene profiles*

Similarity among the VG profiles of water samples was measured and corresponded to the land use and characteristics of each sampling site to ascertain any possible correlation between the nature of the land use and its characteristics with VG contents of water at that site. The results indicated that some sites (e.g. FW5 and FW6) with similar land uses had high similarity during the both seasons (Fig. 6), whilst other sites such as FW7 and FW8 with similar land uses had a high similarity to each other only during the wet season (Fig. 6). Similar results were found with other sites during the wet seasons (e.g. FW9 and FW12 and FW11 and FW13) (Fig. 6). All sites with a similarity coefficient above 70% in the dry season and above 80% in the wet season were all fresh water with similar adjacent land uses (Fig. 6). Further analysis of available data on land use of these sites indicated that they were semi-rural locations with various agricultural uses and native animal habitats. The increased similarity during the wet season however could be due to the increased diversity of *E. coli* sources with equally diverse VGs as a result of the local run off during wet weather events.

4. Discussion

As expected the numbers of *E. coli* in the water samples collected during the dry season were much lower than those found during the wet season, this is largely due to the run off generated from heavy rainfall events transporting a high number of bacteria from various point and non-point sources to the waterways (O'Shea and Field, 1992). However, this was not the case for freshwaters where several sites showed a higher number of *E. coli* during the dry season. *E. coli* is known to have short survival rates in saline water and because of this they may not be an indicator of faecal pollution for the estuarine and brackish waters (Anderson et al., 1983). In our study, freshwater sites with a high number of *E. coli* showed to be located in pasture or peri-urban catchments

with multiple sources of faecal pollution such as defective septic systems and faecal materials from cattle, horses and wild animals. A previous study has also reported the presence of high numbers of faecal indicator bacteria originating from defective septic systems and grazing animals in these sites (Ahmed et al., 2005a, b).

The presence of strains with virulence characteristics similar to EPEC (Lauber et al., 2003), ETEC (Chern et al., 2004), EHEC (Chern et al., 2004), and ExPEC (Hamelin et al., 2006) has been previously reported in the freshwaters. A more recent study also indicates the presence of potential EPEC strains in estuarine waters (Hamilton et al., 2010). In our study a high number of *E.coli* VGs were detected in samples from the estuarine and brackish waters during both dry and wet seasons suggesting the persistence of pathogenic *E. coli* strains in these waters. Alternatively this may indicate a continuous

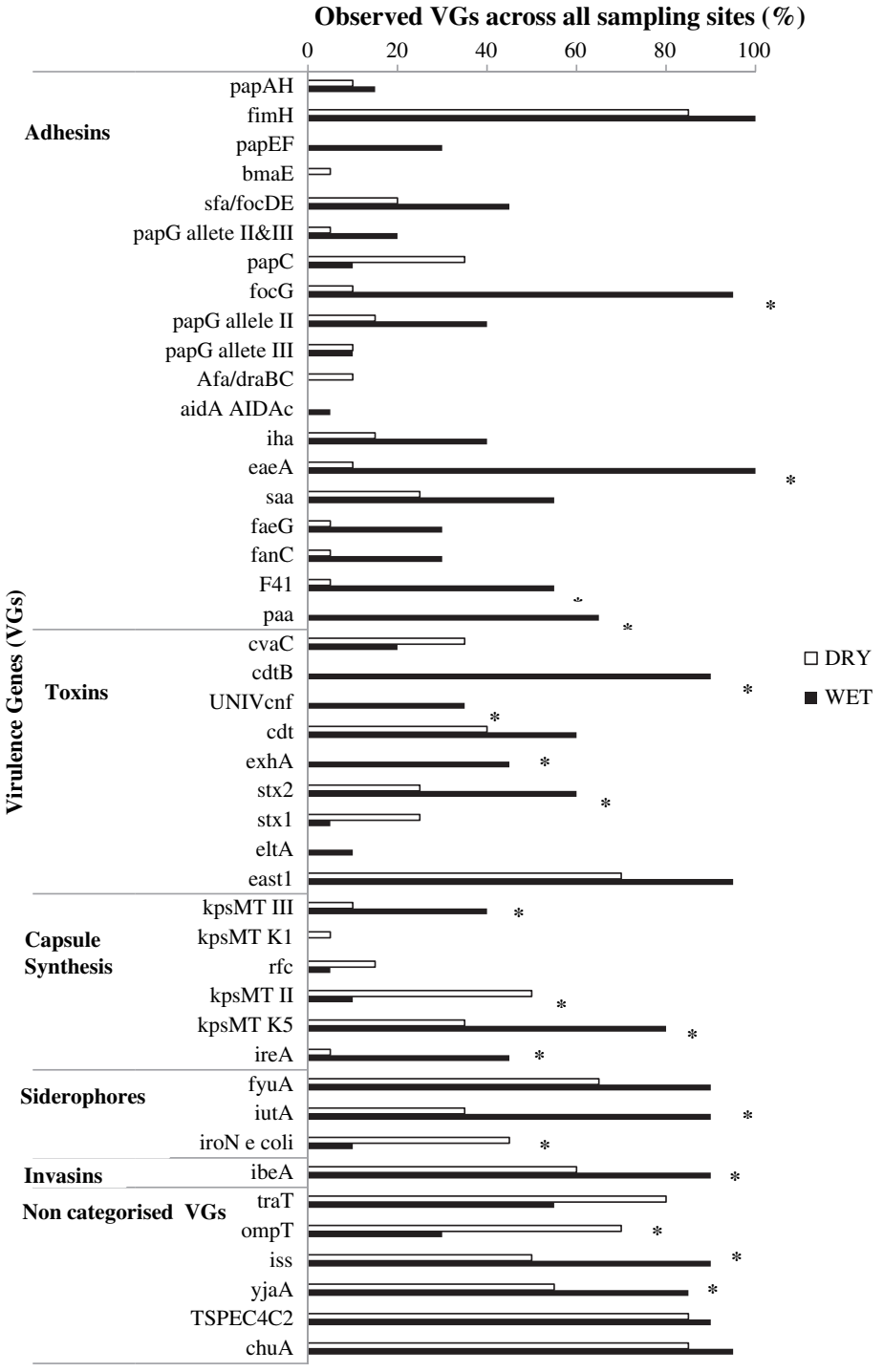

Fig. 2 – The overall percentage of positive gene observations in all water samples during wet and dry season. * indicate significance of difference ranging from $P < 0.05$ to $P < 0.001$ between the number of observations, calculated using a Z-test.

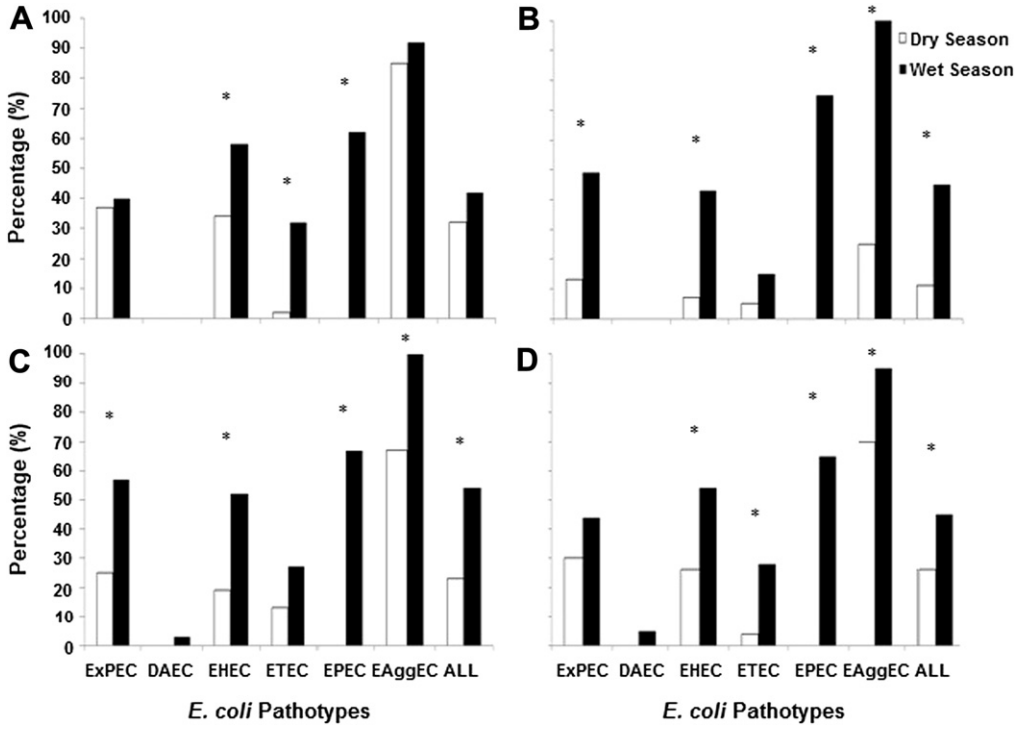

Fig. 3 – The percentage of positive gene observations for each pathotype in the different water types; fresh (A), brackish (B), estuarine (C) and overall (D). Virulence genes that belonged to more than one pathotype were grouped with all associated pathotypes. * indicate significance of difference ranging from $P < 0.05$ to $P < 0.001$ between the number of observations of VGs belonging to each pathotype, calculated using a Z-test.

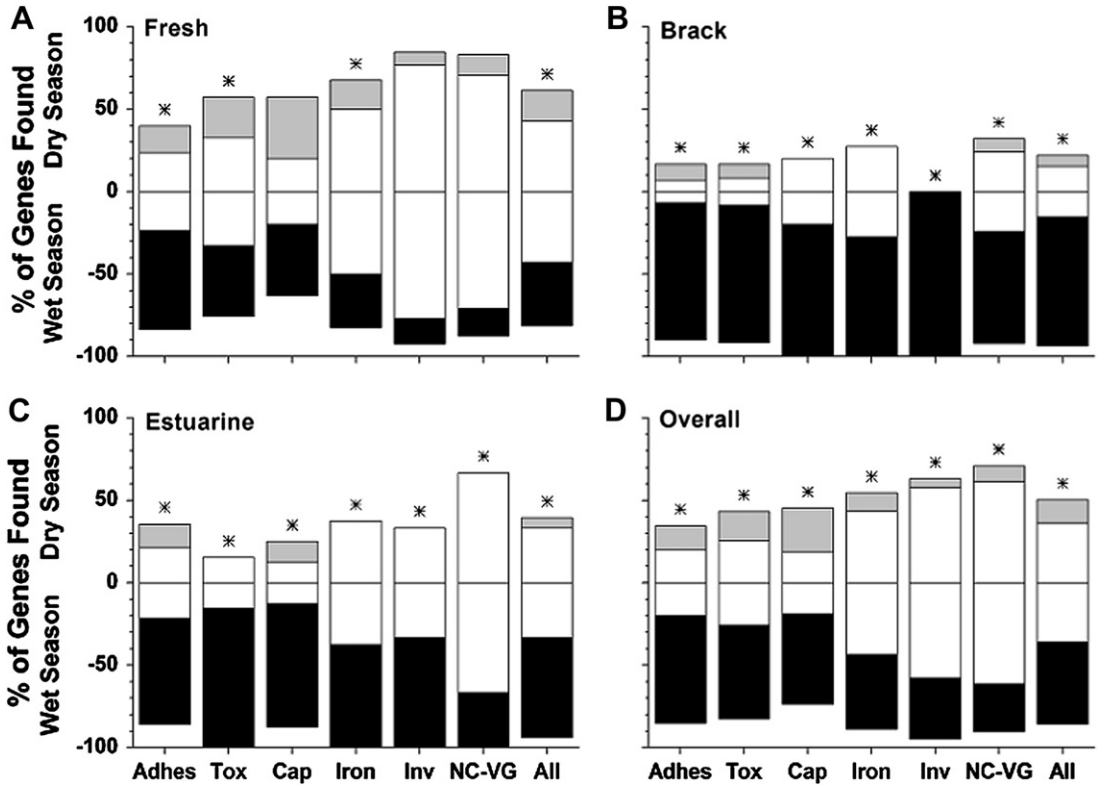

Fig. 4 – The percentage of virulence genes (VGs) classified based on functional characteristics during the dry season (top half of each graph) and the wet season (bottom half of each graph): for each water type; fresh (A), brackish (B), Estuarine (C) and Overall (D). ▭-VGs unique to the dry season, ■-VGs unique to the wet season, ▭-VGs commonly found during dry and wet season. Tox; toxin genes, Adhes; adhesion genes, Cap; capsule genes, Iron; iron acquisition genes, Inv; invasion genes, NC-VG; non-classified virulence genes. * indicate significance ($P < 0.05$) of difference for overall percentage of VGs found during the dry and wet season, calculated using a paired t-test.

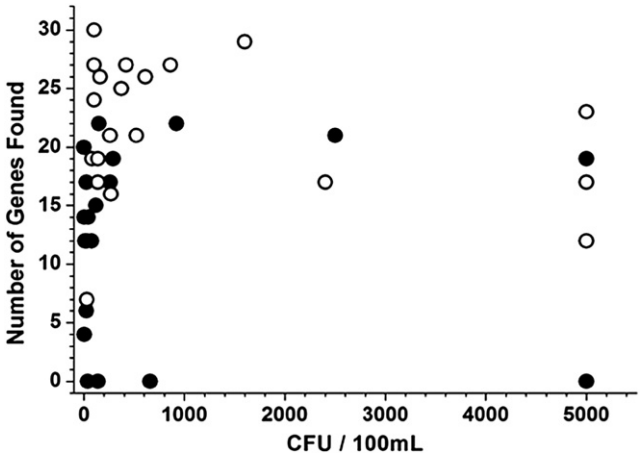

Fig. 5 – The correlation between the numbers of *E. coli* during the dry ○ and wet ● seasons and the corresponding total number of positive virulence gene observations at each sampling site.

input of these bacteria from a common source in the water or a combination of both. The fact that VGs belonging to ExPEC pathotype were more prevalent in these waters was of interest and may indicate a high potential health risk of such waters as the number of ExPEC VGs in *E. coli* is proportional to its pathogenic potential (Picard et al., 1999). A number of samples were also positive for toxin genes belonging to intestinal pathogenic bacteria. The frequency of detecting these toxin genes was higher during the wet season than the dry season especially in the freshwater sites which are normally surrounded by animal farm such as ruminants and swine that are known to harbour these VGs (Djordjevic et al., 2004; Gyles, 2007; Ishii et al., 2007). Interestingly a recent study in this region has also identified farm animals and septic systems as potential sources of these VGs in waterways (Ahmed et al., 2007). The fact that almost all water samples containing stx_1 and/or stx_2 also harboured *eaeA* gene has to be emphasised. This gene is required for the full expression of virulence by EHEC strains (Borelin et al., 1999) and is most commonly found in ruminant and avian guts (Hamelin et al., 2007).

The prevalence of *E. coli* isolates harbouring VGs in environmental waters is reported to be low ranging from 0.9% to 10% (Chern et al., 2004; Lauber et al., 2003; Martins et al., 1992).

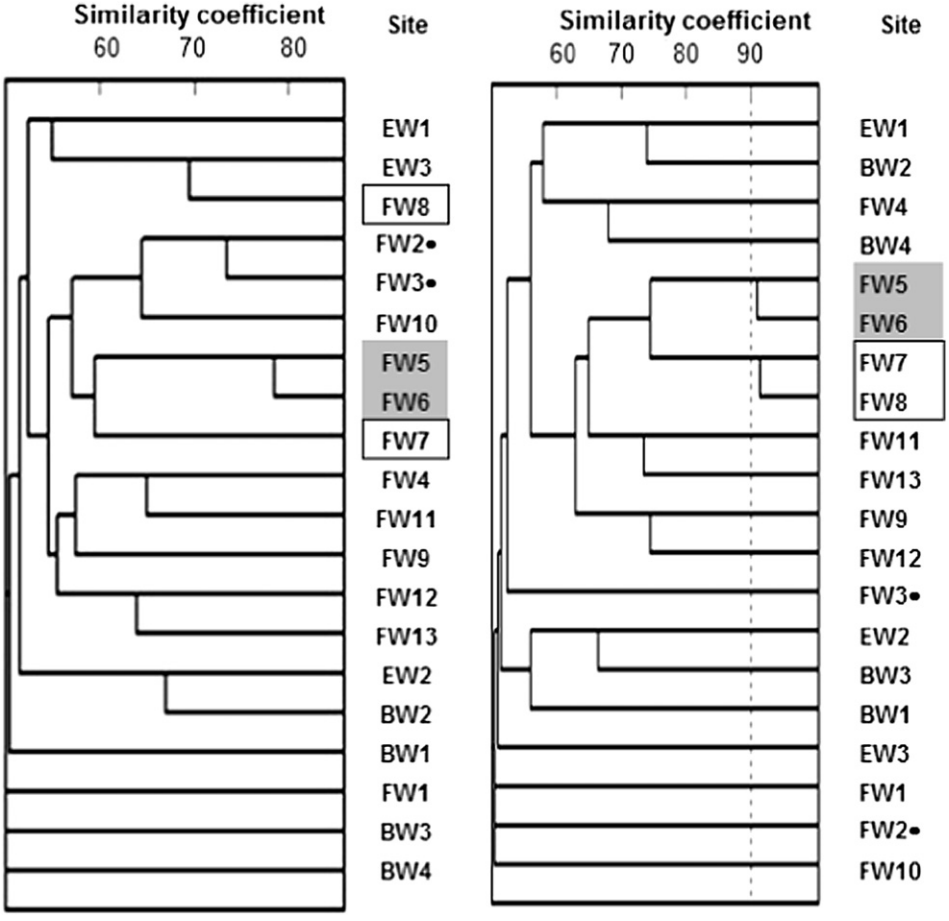

Fig. 6 – Dendrograms showing similarity among the virulence gene (VG) profiles of all water sites during the dry (left dendrogram) and wet (right dendrogram) seasons. Sites with high similarity during both seasons have been highlighted with grey. Sites with low similarity during the dry season but high similarity in their VG profiles have been identified with box. • indicates highly similar sites during dry season but no similarity during the wet season.

Therefore, a large number of isolates need to be screened in order to detect VGs. In view of this, in this study, *E. coli* isolates were concentrated from large volume (1 L) of water samples followed by an enrichment step so that PCR detection sensitivity can be increased as shown before (Ahmed et al., 2009; Savichtcheva et al., 2007; Scott et al., 2005).

Three freshwater sites (FW7, FW8, and FW12) were positive for LT1 and STa genes. The presence of ST and LT enterotoxins which are commonly associated with ETEC strains have been reported by other workers in surface waters (Obi et al., 2004; Begum et al., 2005) and is thought to be originated from swine and humans with diarrhoea.

The toxin gene *east1* was highly prevalent in water samples from the estuarine, brackish and freshwater sites during both the dry and wet season. This gene is commonly found in ruminants and swine (Yamamoto and Nakazawa, 1997) and are, in some cases, responsible for waterborne diseases in humans (Hedberg et al., 1997; Yatsuyanagi et al., 2003).

E. coli has long been used as one of the primary faecal indicator bacteria due to the previous assumption that it has limited survival ability within the environment however recent studies suggest that some lineages of *E. coli* have adapted and naturalized within tropical, subtropical and even temperate environments (Ishii and Sadowsky, 2008; Walk et al., 2007) and as such could be the pertinent reason for *E. coli* blooms especially within freshwaters.

Data from Hamilton et al. (2010), and the present study suggests *E. coli* strains harbouring clinically significant VGs may persist in the estuarine environment despite the fact that *E. coli* is not generally monitored to assess faecal pollution in the estuarine waters. In fact, there are few studies that have correlated the presence of known VGs with the number of *E. coli* in all water types. The present study supports the idea that at least some clinically significant strains of *E. coli* may persist in the waterways surviving both the dry and wet season and this may explain why some sites with very low numbers of *E. coli* showed the presence of a high number VGs. The presence of these genes could not totally be attributed to storm run off in wet seasons or fresh waters as they were also found during the dry season and in all water types.

Several sites in this study had lower numbers of *E. coli* than those required by national and international water quality guidelines for fresh and marine waters (USEPA, 1986; ANZECC, 2000) but contained a high proportion of VGs. Contrary to these, there were sites that far exceeded the existing guidelines for the number of *E. coli* for recreational waters and yet found to harbour no VGs. Whilst these data outline the limited correlation between *E. coli* numbers and VGs, they also highlight the potential of surface water with low and accepted level of *E. coli* to cause infections.

In this study we aimed to determine the viability of *E. coli* VG profiles as a potential indicator of water quality. The results suggest there is a varied distribution of VGs within the catchment. Whilst some VGs were specific to the wet season, suggesting their affiliation to storm-water run off, some genes were specific to water types and were found at specific sampling locations suggesting their association to localised faecal contamination. We also found that fresh water sites were less populated, and mainly constituted of agricultural, semi-rural areas with pockets of natural habitat as opposed to the brackish and estuarine sites. The latter's were all found to be located within medium to high density residential areas. This may indicate that the fresh water sites were probably contaminated with animal faecal *E. coli* whilst the other two water types were primarily impacted on by urban run off. The fact that Figs. 3 and 4 showed a much higher similarity between the VG profile of brackish and estuarine water samples supports this observation. In this study we used a similarity coefficient between the VG profiles of each sampling site and found a much higher similarity among sites that had common sources of contamination indicating the possible use of such method for analysis of water sites based on their VG profile. This however was more pronounced in samples collected during the wet season. There were also sites that showed to be distantly dissimilar to each other during the dry season but had a high similarity in their VG profile during the wet season suggesting that the VG profiling of *E. coli* in a water site especially after a storm run off can give a better picture of the *E. coli* flora of the site with respect to its animals and human land uses pattern.

In conclusion, a number of water samples collected during the dry and wet conditions were positive for multiple *E. coli* VGs which indicate the presence of potential pathogenic *E. coli* in these waters. However, the percentage of *E. coli* isolates harbouring these VGs is not reported in this study, and should be addressed in future studies to provide a better understanding of the potential health risk of such waters. We also suggest that VG profiling of surface waters can be used as a tool to indicate water quality and should be used in conjunction with enumeration of *E. coli* bacteria in water samples.

Acknowledgements

We are grateful to Dr Toni Chapman from Elizabeth MacArthur Agricultural Institute, Industry and Investment of NSW, Australia and Miss Nubia Ramos from the University of the Sunshine Coast for providing the positive control strains. We also thank Susie Chapman from SEQ catchments and Graham Webb from the Sunshine Regional Council for their advice on site selection.

REFERENCES

Ahmed, W., Neller, R., Katouli, M., 2005a. Evidence of septic system failure determined by a bacterial biochemical fingerprinting method. Journal of Applied Microbiology 98, 910–920.

Ahmed, W., Neller, R., Katouli, M., 2005b. Host species-specific metabolic fingerprint database for enterococci and *Escherichia coli* and its application to identify sources of fecal contamination in surface waters. Applied Environmental Microbiology 71, 4461–4468.

Ahmed, W., Tucker, J., Bettelheim, K., Neller, R., Katouli, M., 2007. Detection of virulence genes in *Escherichia coli* of an existing metabolic fingerprint database to predict the sources of pathogenic *E. coli* in surface waters. Water Research 41, 3785–3791.

Ahmed, W., Sawant, S., Huygens, F., Goonetilleke, A., Gardener, T., 2009. Prevalence and occurrence of zoonotic bacterial pathogens in surface waters determined by quantitative PCR. Water Research 43, 4919–4928.

Anderson, I.C., Rhodes, M., Kator, H.I., 1983. Seasonal variations in survival of *Escherichia coli* exposed in situ in membrane diffusion chambers containing filtered and nonfiltered estuarine water. Applied Environmental Microbiology 45, 1877–1883.

Aslan-Yilmaz, A., Okus, E., Ovez, S., 2004. Bacteriological indicators of anthropogenic impact prior to and during the recovery of water quality in an extremely polluted estuary, Golden Horn, Turkey. Marine Pollution Bulletin 49, 951–958.

Australian and New Zealand Environment and Conservations Council with the Agriculture and Resource Management Council of Australia and New Zealand, 2000. The Australian and New Zealand guidelines for fresh and marine water quality. National Water Quality Management Strategy, ANZECC and ARMCANZ, Canberra.

Begum, Y.A., Talukder, K.A., Nair, G.B., Qadri, F., Sack, R.B., Svennerholm, A.M., 2005. Enterotoxigenic *Escherichia coli* isolated from surface water in urban and rural areas of Bangladesh. Journal of Clinical Microbiology 43, 3582–3583.

Borelin, P., McEwen, S.A., Borelin-Petzold, F., Wilson, J.B., Johnson, R.O., Gyles, C.L., 1999. Associations between virulence factors of shiga toxin-producing *Escherichia coli* and disease in humans. Journal of Clinical Microbiology 37, 497–503.

Chapman, T.A., Wu, X., Barchia, I., Bettelheim, K.A., Driesen, S., Trott, D., Wilson, M., Chin, J., 2006. Comparison of virulence gene profiles of *Escherichia coli* strains isolated from healthy and diarrheic swine. Applied Environmental Microbiology 72, 4782–4795.

Chern, E.C., Tsai, Y.L., Olson, B.H., 2004. Occurrence of genes associated with enterotoxigenic and enterohemorrhagic *Escherichia coli* in agricultural waste lagoons. Applied Environmental Microbiology 70, 356–362.

Desmarais, T.R., Solo-Gabriele, H.M., Palmer, C.J., 2002. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. Applied Environmental Microbiology 68, 1165–1172.

Djordjevic, S.P., Ramachandaran, V., Bettelheim, K.A., Vanselow, B.A., Holst, P., Bailey, G., Horintzky, M.A., 2004. Serotypes and virulence gene profiles of shiga-toxin producing *Escherichia coli* strains isolated from feces of pasture-fed and lot-fed sheep. Applied Environmental Microbiology 70, 3910–3917.

Falagas, M., Gorbach, S., 1995. Practice guidelines: urinary tract infections. Infectious Diseases in Clinical Practice 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 Reports 117, 380–385.

Gilmore, M.S., Ferretti, J.J., 2003. The thin line between gut commensal and pathogen. Science 299, 1999–2002.

Gyles, C.L., 2007. Shigatoxin-producing *Escherichia coli*: an overview. Journal of Animal Science 85, E45–E62.

Hamelin, K., Braunt, G., El-Shaarawi, S., Hill, S., Edge, T.A., Bekal, S., Fairbrother, J.M., Harel, J., Maynard, C., Masson, L., Brousseau, R., 2006. A virulence and antimicrobial resistance DNA microarray detects a high frequency of virulence genes in *Escherichia coli* isolates from Great Lakes recreational waters. Applied Environmental Microbiology 72, 4200–4206.

Hamelin, K., Bruant, G., El-Shaarawi, A., Hill, S., Edge, T.A., Fairbrother, J., Harel, J., Maynard, C., Masson, L., Brousseau, R., 2007. Occurrence of virulence and antimicrobial resistance genes in *Escherichia coli* isolates from different aquatic ecosystems within St. Clair River and Detroit River areas. Applied Environmental Microbiology 73, 477–484.

Hamilton, M., Hadi, A.Z., Griffith, J.F., Sadowsky, M.J., 2010. Large scale analysis of virulence genes in *Escherichia coli* strains isolated from Avalon bay, CA. Water Research 44, 5463–5473.

Hart, C.A., Batt, R.M., Saunders, J.R., 1993. Diarrhea caused by *Escherichia coli*. Annals of Tropical Paediatrics 13, 121–131.

Haramoto, E., Katayama, H., Oguma, K., Ohgaki, S., 2005. Application of cation-coated filter method to detection of noroviruses, enteroviruses, adenoviruses, and torqueteno viruses in the Tamagawa River in Japan. Applied Environmental Microbiology 71, 2403–2411.

Hedberg, C.W., Savarino, S.J., Besser, J.M., Paulus, C.J., Thelen, V.M., Myers, L.J., Cameron, D.N., Barrett, T.J., Kaper, J.B., Osterholm, M.T., 1997. An outbreak of foodborne illness caused by *Escherichia coli* O39:NM, an agent not fitting into the existing scheme for classifying diarrheogenic *E. coli*. Journal of Infectious Disease 176, 1625–1628.

Hörman, A., Rimhannen-Finne, R., Maunula, L., von Bonsdorff, C.H., Torvela, N., Heikinheimo, A., Hänninen, M.L., 2004. *Campylobacter* spp., *Giardia* spp., *Cryptosporidum* spp., noroviruses, and indicator organisms in surfacewater in South western Finland, 2000–2001. Applied Environmental Microbiology 70, 87–95.

Ibekwe, A.M., Grieve, C.M., 2003. Detection and quantification of *Escherichia coli* O157:H7 in environmental samples by real-time PCR. Journal of Applied Microbiology 94, 421–431.

Ishii, S., Meyer, K.P., Sadowsky, M.J., 2007. Relationship between phylogenetic groups genotypic clusters and virulence genes profiles of *Escherichia coli* strains from diverse human and animal sources. Applied Environmental Microbiology 73, 5703–5710.

Ishii, S., Sadowsky, M.J., 2008. *Escherichia coli* in the environment: implications for water quality and human health. Microbes and Environments 23, 101–108.

Johnson, J.R., Stell, A.L., 2000. Extended virulence genotypes of *Escherichia coli* strains from patients with urosepsis in relation to phylogeny and host compromise. Journal of Infectious Disease 181, 261–272.

Kaper, J.B., Nataro, J.P., Mobley, H.L., 2004. Pathogenic *Escherichia coli*. Nature Reviews Microbiology 2, 123–140.

Lauber, C.L., Glatzer, L., Sinsabaugh, R.L., 2003. Prevalence of pathogenic *Escherichia coli* in recreational waters. Journal of Great Lakes Research 29, 301–306.

Levine, M.M., 1987. *Escherichia coli* that cause diarrhea: enterotoxigenic, enteropathogenic, enteroinvasive, enterohemorrhagic, and enteroadherent. Journal of Infectious Disease 155, 377–389.

Martins, M.T., Rivera, I.G., Clark, D.L., Olson, B.H., 1992. Detection of virulence factors in culturable *Escherichia coli* isolates from water samples by DNA probes and recovery of toxin-bearing strains in o-nitrophenol-β-D-galactopyranoside-4-methylumbelliferyl- β-D-glucuronide media. Applied Environmental Microbiology 58, 3095–3100.

McQuaig, S.M., Scott, T.M., Harwood, V.J., Farrah, S.R., Lukasik, J., 2006. Detection of human-derived fecal pollution in environmental waters by use of a PCR-based human polyomavirus assay. Applied Environmental Microbiology 72, 7567–7574.

Nataro, J.P., Kaper, J.B., 1998. Diarrheagenic *Escherichia coli*. Clinical Microbiology Reviews 11, 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 haemolytic uremic syndrome: implications for rural water systems. Emerging Infectious Diseases 8, 370–375.

Obi, C.L., Greene, E., Bessong, P.O., De-Villiers, B., Hoosen, A.A., Igumbor, E.O., Potgieter, N., 2004. Gene encoding virulence

markers among *Escherichia coli* isolates from diarrhoeic stool samples and river sources in rural Venda communities of South Africa. Water SA 30, 37–42.

O'Shea, M.L., Field, R., 1992. Detection and disinfection of pathogens in storm-generated flows. Canadian Journal of Microbiology 38, 267–276.

Paton, A.W., Paton, J.C., 1998. Detection and characterization of Shiga toxigenic *Escherichia coli* by using multiplex PCR assays for stx_1, stx_2, *eaeA*, enterohemorrhagic *E. coli hlyA*, rfb_{O111}, and rfb_{O157}. Journal of Clinical Microbiology 36, 598–602.

Picard, B., Garcia, J.S., Gouriou, S., Duriez, P., Brahimi, N., Bingen, E., Elion, J., Denamur, E., 1999. The link between phylogeny and virulence in *Escherichia coli* extraintestinal infection. Infection and Immunity 67, 546–553.

Power, M.L., Littlefield-Wyer, J., Gordon, D.M., Veal, D.A., Slade, M.B., 2005. Phenotypic and genotypic characterization of encapsulated *Escherichia coli* isolated from blooms in two Australian lakes. Environmental Microbiology 7, 631–640.

Ram, S., Vajpayee, P., Shanker, R., 2007. Prevalence of multi-antimicrobial agent resistant shiga toxin and enterotoxin producing Escherichia coli in surface waters of river Ganga. Environmental Science and Technology 41, 7383–7388.

Savichtcheva, O., Okayama, N., Okabe, S., 2007. Relationships between *Bacteroides* 16S rRNA genetic markers and presence of bacterial enteric pathogens and conventional fecal indicators. Water Research 41, 3615–3628.

Scott, T.M., Jenkins, T.M., Lukasik, J., Rose, J.B., 2005. Potential use of a host associated molecular marker in *Enterococcus faecium* as an index of human fecal pollution. Environmental Science and Technology 39, 283–287.

United States Environmental Protection Agency, 1986. Ambient water quality criteria for bacteria EPA-440/5-84-002. U.S. Environmental Protection Agency, Washington, D.C.

Walk, S.T., Alm, E.W., Clahoun, L.M., Mladonicky, J.M., Whittam, T.M., 2007. Genetic diversity and population structure of *Escherichia Coli* isolated from freshwater beaches. Environmental Microbiology 9, 2274–2288.

Yamamoto, T., Nakazawa, M., 1997. Detection and sequences of the enteroaggregative *Escherichia coli* heat stable enterotoxin 1 gene in enterotoxigenic *E. coli* strains isolated from piglets and calves with diarrhea. Journal of Clinical Microbiology 34, 223–227.

Yatsuyanagi, J., Saito, S., Miyajima, Y., Amano, K., Enomoto, K., 2003. Characterization of atypical enteropathogenic *Escherichia coli* strains harboring the *astA* gene that were associated with a waterborne outbreak of diarrhea in Japan. Journal of Clinical Microbiology 41, 2033–2039.