Marine Pollution Bulletin 56 (2008) 1469–1475

Contents lists available at ScienceDirect

Marine Pollution Bulletin

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

Population similarity analysis of indicator bacteria for source prediction of faecal pollution in a coastal lake

W. Ahmed [a,b,c,*], M. Hargreaves [d], A. Goonetilleke [c], M. Katouli [b]

[a] Department of Natural Resources and Water, Brisbane, Queensland 4068, Australia
[b] Faculty of Science, Health and Education, University of the Sunshine Coast, Maroochydore, Queensland, Australia
[c] School of Urban Development, Queensland University of Technology, Brisbane, Queensland, Australia
[d] School of Life Sciences, Queensland University of Technology, Brisbane, Queensland, Australia

ARTICLE INFO

Keywords:
Faecal pollution
Microbial source tracking
Biochemical fingerprinting
Antibiotic resistance analysis
Population similarity analysis

ABSTRACT

Biochemical fingerprinting (BF) databases of 524 enterococci and 571 *Escherichia coli* isolates and an antibiotic resistance analysis (ARA) database comprising of 380 *E. coli* isolates from four suspected sources (i.e. dogs, chickens, waterfowls, and human sewage) were developed to predict the sources of faecal pollution in a recreational coastal lake. Twenty water samples representing four sampling episodes were collected from five sites and the enterococci and *E. coli* population from each site were compared with those of the databases. The degree of similarity between bacterial populations was measured as population similarity (Sp) coefficient. Using the BF-database, bacterial populations of waterfowls showed the highest similarity with the water samples followed by a sewage treatment plant (STP). Higher population similarities were found between samples from STP and water samples especially at two sites (T2 and T3) which were located near the sewerage pipes collecting wastewater from the study area. When using the ARA-database, the highest similarity was found between *E. coli* populations from STP and water samples at sites T2 and T4. Both faecal indicators and as well as methods predicted human faecal pollution, possibly through leakage from submerged sewerage pipes. The results indicated that the Sp-analysis of faecal indicator bacterial populations from suspected sources and water samples can be used as a simple tool to predict the source(s) of faecal pollution in surface waters.

1. Introduction

Faecal pollution from warm-blooded animals including humans poses public health risks due to possible exposure to a wide array of pathogenic bacteria, viruses and protozoa (Baker and Herson, 1999; Fong et al., 2005). Non-point sources (NPS) such as domestic and wild animals, urban run-off and malfunctioning septic systems and/or point sources (PS) such as industrial outlets, combined sewer overflows and wastewater effluents are known to be potential sources of such pollution (Ahmed et al., 2005; McLellan, 2004; Parveen et al., 1997). Faecal indicator bacteria such as coliforms, *Escherichia coli* and enterococci have long been used by the water quality monitoring authorities to detect faecal pollution and the presence of potential pathogenic microorganisms. However, the presence of these indicator bacteria in aquatic environments does not provide information regarding their possible sources (Harwood et al., 2000; McLellan, 2004).

The identification of major polluting source(s) is vitally important in order to implement appropriate mitigation strategies to minimize faecal pollution and consequent public health risks. In view of this, various microbial source tracking (MST) techniques have been developed or are under development to distinguish different sources of animal and/or human faecal pollution (Field and Samadpour, 2007; Stoeckel and Harwood, 2007). These methods can be classified as database-dependent (DD) and database-independent (DI) methods. Database-dependent methods require the development of a host-origin database of phenotypic or genotypic attributes of the selected indicator bacteria (e.g. enterococci or *E. coli*) from suspected sources. Unknown environmental bacterial strains are compared to the database to identify their most likely sources (Harwood et al., 2003). In contrast, DI methods are mainly polymerase chain reaction (PCR) based and do not require the development of a database. Instead, these methods detect the presence of specific gene(s) associated with certain groups of bacteria from humans and animals (Field and Samadpour, 2007).

The most commonly used DD methods include antibiotic resistance analysis (ARA) (Harwood et al., 2000; Moore et al., 2005), carbon source utilization (CSU) (Hagedorn et al., 2003), biochemical fingerprinting (BF) (Ahmed et al., 2005; Blanch et al., 2006),

* Corresponding author. Address: Department of Natural Resources and Water, Brisbane, Queensland 4068, Australia.
E-mail address: Warish.Ahmed@nrw.qld.gov.au (W. Ahmed).

0025-326X/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.marpolbul.2008.04.043

rep-PCR (Johnson et al., 2004; McLellan, 2004), and ribotyping (Scott et al., 2004). On the other hand, DI methods include host-specific markers such as human and cattle specific *Bacteroides* (Bernhard and Field, 2003), human-specific enterococci surface protein (*esp*) marker found in *E. faecium* (Scott et al., 2005), and host-specific viruses (Fong et al., 2005; Love and Sobsey, 2007). The performance of these methods has not been fully evaluated or some are under evaluation in environmental studies. However, to date none of these methods are considered as being superior to others in terms of identifying the all possible sources of pollution. Two recent review articles have discussed the advantages and disadvantages of various MST methods (Field and Samadpour, 2007; Stoeckel and Harwood, 2007). Recently, it has been reported that a combination of two or more methods could reliably be used to identify the sources of faecal pollution in environmental waters (Ahmed et al., 2007; Edge and Hill, 2007).

In this study, we used two phenotypic DD methods: BF and ARA to identify the sources of enterococci and *E. coli* in a small recreational coastal lake. A population similarity (Sp) analysis was used to compare the overall similarity between bacterial populations from suspected sources with those found in the water samples. Based on the Sp-analysis, we also report the performance of each database (BF or ARA) alone or in combination (BF-ARA) in identifying the dominant source(s) of faecal pollution.

2. Materials and methods

2.1. Study lake

The study lake (i.e. Tooway Lake) is located within the Caloundra City, in southeast Queensland, Australia. The city has a population of approximately 70,000, and is one of the fastest growing cities in Australia and is characterized by numerous beaches, rainforest, and national parks. The lake is extensively used for swimming and recreational activities, and is surrounded by sewered residential areas. The lake is linear and approximately 4 km in length (Fig. 1). The lake mouth is connected to the sea and is periodically closed by the formation of sand plugs. The mouth reopens only after significant storm events. Water quality monitoring by Caloundra City Council reported non-conforming faecal coliforms exceeding the Australian and New Zealand Environment and Conservation Council (ANZECC) recreational water quality guidelines for fresh and marine waters of 150 faecal coliforms 100 ml^{-1} for primary contact (ANZECC, 2000).

2.2. Host groups sampling

Dogs, chickens, and waterfowls were identified as potential NPS faecal pollution during a thorough sanitary survey. In addition, sewerage pipes collecting raw sewage were also suspected as possible PS of human faecal pollution. For this reason composite raw sewage samples ($n = 3$) were collected from the primary influent of a sewage treatment plant (STP) servicing the area. Faecal samples from waterfowls ($n = 24$) were collected from a golf-course located at the upper reaches of the lake. Dog faecal samples ($n = 17$) were collected from a dog kennel located 250–300 m below the golf-course and finally chicken faecal samples ($n = 21$) were collected from various properties adjacent to the lake. All faecal samples from animals were collected from fresh defecation of individual animal using aseptic technique, and transported on ice to the laboratory and processed within 6 h.

2.3. Isolation of faecal indicators

Animal faecal samples were streaked on m-Enterococcus (Difco, London, UK) and RAPID' *E. coli* 2 (REC 2) agar plates (Bio-rad, USA). Plates were incubated at 45 °C for 48 h (for enterococci) and 24 h (for *E. coli*). All isolates from m-Enterococcus plates were also tested for esculin hydrolysis using bile esculin agar (Oxoid, London, UK), and incubated at 45 °C for 1 h to confirm their identification as enterococci (Manero and Blanch, 1999). Single purple colonies from REC 2 agar plates were streaked on McConkey agar (Oxoid, London, UK) for purity and tested for indole production and citrate cleavage. Indole positive and citrate negative isolates were confirmed as *E. coli*. Raw sewage samples were suspended in 100 ml of buffered water (0.042 g l^{-1} KH_2PO_4 and 0.4055 g l^{-1} $MgCl_2$) and vortexed for 3 min. Serial dilutions were made and filtered through 0.45 μm pore size (47 mm diameter) nitrocellulose membranes (Advantec, Tokyo, Japan) and placed on m-Enterococcus (for enterococci) and REC 2 (for *E. coli*) agar plates. After incubation, the confirmatory tests were performed as described above.

2.4. Biochemical fingerprinting

Biochemical fingerprinting method was used to type enterococci and *E. coli* from the faecal samples of the four suspected sources. Two types of micro plates specifically developed for typing of enterococci strains (PhP-RF plates) and *E. coli* (PhP-RE plates) (PhPlate system, PhPlate AB, Stockholm, Sweden) were used. The

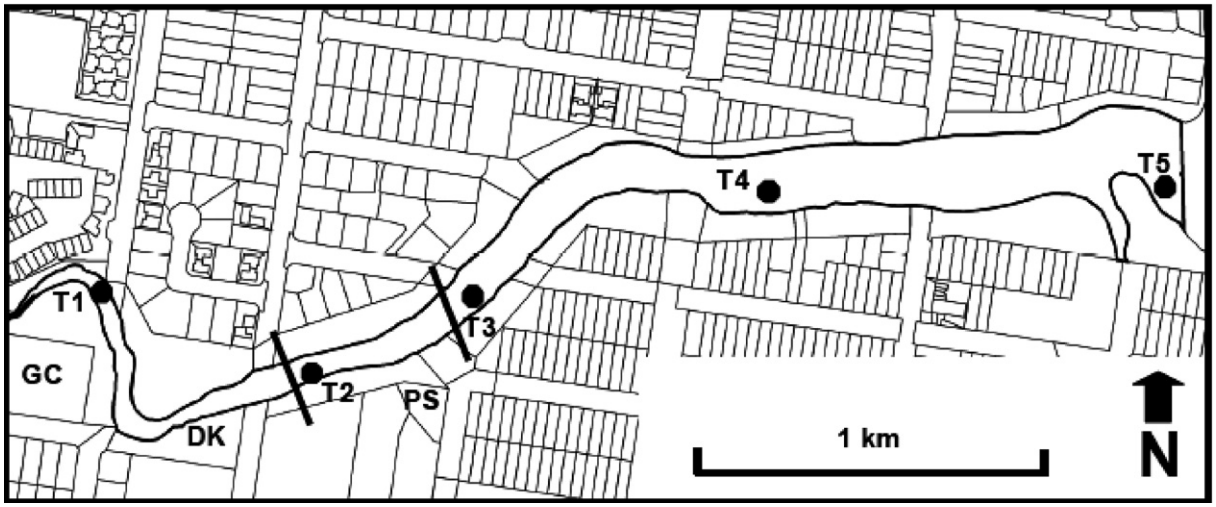

Fig. 1. Sampling sites (T1–T5) on Tooway Lake. GC, DK, and PS showing the location of the golf-course, dog kennel, and pumping station, respectively.

11 substrates used for enterococci and *E. coli* have been described before (Iversen et al., 2002; Kühn et al., 1991). The fingerprinting procedures were performed according to the manufacturer's instruction. In brief, enterococci and *E. coli* colonies were suspended in the first well of each row containing 350 μl of growth medium. Aliquots of 25 μl of bacterial suspensions were transferred into each of the other 11 wells containing 150 μl growth medium. Plates were then incubated at 37 °C and A_{620} was measured at 7, 24, and 48 h for *E. coli* and at 16, 40 and 64 h for enterococci using a micro plate reader (Lab-system 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 with arithmetic averages (UPGMA) (Sneath and Sokal, 1973). An identity level 0.96 was established based on the reproducibility of the system as described before (Ahmed et al., 2005). Isolates showing similarity to each other above the ID-level were regarded as identical and were assigned to the same biochemical phenotype (BPT). Enterococci ($n = 524$) and *E. coli* ($n = 571$) BPTs from the host groups/sources were used to develop BF-databases.

2.5. Antibiotic resistance analysis

Escherichia coli isolates from the sources were also tested for their antibiotic resistance patterns (ARPs) according to previously published method (Carroll et al., 2005; Harwood et al., 2000). In brief, media was prepared by amending sterile trypticase soy agar (TSA) (Oxoid, London, UK) with antibiotics and poured into petridishes. The ARP of each isolate comprised of 32 observations of eight antibiotics (Sigma Chemical Co., St. Louis, Mo, USA) at four concentrations and included: amoxicillin, 5, 10, 15, and 20 μg ml^{-1}; cephalothin, 10, 25, 50, and 100 μg ml^{-1}; erythromycin, 20, 50, 80, and 100 μg ml^{-1}; gentamicin, 20, 40, 60, and 80 μg ml^{-1}; ofloxacin, 5, 10, 15, and 20 μg ml^{-1}; chlortetracycline, 20, 40, 60, and 80 μg ml^{-1}; tetracycline, 20, 40, 60, and 80 μg ml^{-1}; moxalactam, 5, 10, 15, and 20 μg ml^{-1}. Isolates were inoculated into nutrient broth and incubated at 37 °C for 18 h, and then transferred into sterile multipoint inoculator cups and inoculated onto pre-made TSA agar plates with antibiotics. One TSA agar plate (Oxoid) was left blank and served as the control. All plates were incubated at 37 °C for 24 h. After incubation, the bacterial growths for different antibiotics and concentrations on each plate were recorded using four ratings: 1 – for no growth, 2 – for filmous growth, 3 – for restricted growth and 4 – for full growth of colonies (Carroll et al., 2005). Isolates that did not exhibit any growth on the control plates was not used in the analysis. In the current study, of the 571 *E. coli* isolates tested, 380 were resistant to one or more antibiotics, and were used to develop the ARA-database. Isolates which were not resistant to any of the antibiotics tested from the sources were not included in the database as they could not be used to distinguish among the sources.

2.6. Lake sampling

Five sampling sites (i.e. T1–T5) were chosen at various points along the length of the lake (Fig. 1). Water samples ($n = 20$) were collected fortnightly on four occasions. Sample site T1 was located in the upper reaches of the lake which receives water from an upstream golf-course. Sample site T2 was located near a dog kennel. Sample sites T3 and T4 both were located near sewerage pipes. Sample site T5 was located at lake mouth which is exclusively used for swimming and other recreational activities. From each site up to 39 colonies of enterococci (where possible) and 23 colonies of *E. coli* (where possible) were biochemically fingerprinted. The biochemical fingerprints of enterococci and *E. coli* isolates from each site were compared to the corresponding BF-database. However, only *E. coli* isolates from water samples were subjected to the antibiotics panel and their ARPs were compared to the ARA-database. In addition, the ARPs and BF data of *E. coli* isolates from water samples were combined together and compared with the composite of BF-ARA-database.

2.7. Data analysis and source tracking

To identify the dominant source(s) of faecal pollution, a Sp-analysis (Kühn et al., 1991) was performed between the populations of faecal indicators from the sources and water samples based on their BPTs, ARPs, and a combination of both. In brief, phenotypic similarity between bacterial populations in two samples was calculated as Sp-coefficient and determined as $(Sx + Sy)/2$, where Sx is the similarity of population x in population y and Sy is the similarity of population y in population x. In this way, the Sp-coefficient calculates the proportion of isolates that are identical (i.e. having the same BPT/ARP) in two or more compared bacterial populations (Kühn et al., 1991). It is high (maximum 1) if two populations contain isolates with similar BPT/ARP, and is low (minimum 0) if the populations contain isolates with different BPT/ARP. Comparison of several bacterial populations yielded a matrix of Sp-coefficient which can be clustered to show the degree of similarity between populations in a two dimensional dendrogram. All data handling, including optical readings and calculation of Sp-coefficient were performed using the PhPlate software version 4001.

3. Results

3.1. Abundance of indicator bacteria in the lake

The number of enterococci in water samples ranged from 140 to 760 colony-forming units (cfu) 100 ml^{-1} with the highest number observed at site T3 (760 cfu 100 ml^{-1}) followed by 680 cfu 100 ml^{-1} at site T2. For *E. coli*, these figures ranged from 252 to 3125 cfu 100 ml^{-1} with the highest number at site T2 (3125 cfu 100 ml^{-1}) followed by 2533 cfu 100 ml^{-1} at site T3. The abundance of indicator bacteria in all water samples exceeded the Australia and New Zealand Environment and Conservation Council (ANZECC) recreational water quality guidelines for fresh and marine waters of 150 faecal coliforms 100 ml^{-1} and 35 enterococci 100 ml^{-1} for primary contact (ANZECC, 2000).

3.2. Antibiotic resistance patterns of Escherichia coli strains

All 571 *E. coli* isolates collected from sources/host groups were subjected to antibiotics panel and their ARPs were determined by numerical scoring scheme. Of these, 380 (67%) isolates were resistant to one or more antibiotics at different concentrations (Table 1). In general, isolates from all sources showed lower resistance to gentamycin (80 μg ml^{-1}), ofloxacin (5, 10, 15 and 20 μg ml^{-1}), tetracycline (80 μg ml^{-1}) and moxolactam (10, 15, and 20 μg ml^{-1}). The best discrimination between human and animals was found on the basis of amoxicillin (15 and 20 μg ml^{-1}), cephalothin (10 μg ml^{-1}), erythromycin (20 and 100 μg ml^{-1}), chlortetracycline (20 μg ml^{-1}), and moxolactam (5, 10, 15, and 20 μg ml^{-1}). Isolates from waterfowls were the least resistant to antibiotics.

3.3. Population similarity (Sp) analysis among the suspected sources

Initially, a Sp-analysis was performed by comparing bacterial populations from each source with others to identify the degree of similarities among the sources. Table 2 shows the number of enterococci and *E. coli* isolates tested from the sources for their

Table 1
Percentage of antibiotic resistance of *Escherichia coli* isolates tested from sources and water samples

Antibiotic (μg ml^{-1})	% of isolates resistant				
	STP (n = 140)	Waterfowls (n = 66)	Dogs (n = 98)	Chickens (n = 76)	Water samples (n = 224)
Amoxicillin (5-10-15-20)	94-91-78-69	36-27-26-7	68-55-44-33	62-57-46-32	41-38-35-33
Cephalothin (10-25-50-100)	87-31-19-11	41-33-21-14	45-40-32-28	43-38-38-22	34-30-24-22
Erythromycin (25-50-80-100)	89-48-22-8	27-24-27-6	62-49-44-33	59-57-32-28	51-43-21-10
Gentamycin (20-40-60-80)	13-9-2-0	17-5-0-0	17-14-11-6	22-14-5-0	12-8-7-5
Ofloxacin (5-10-15-20)	10-4-2-1	9-2-0-0	11-4-0-0	4-1-0-0	8-5-4-3
Chlortetracycline (20-40-60-80)	54-38-15-6	32-29-21-14	35-32-27-20	36-28-18-8	43-36-25-15
Tetracycline (20-40-60-80)	32-22-8-0	11-9-3-0	24-21-14-9	25-14-12-1	14-12-8-7
Moxolactam (5-10-15-20)	26-15-14-10	0-0-0-0	7-3-2-1	14-6-4-1	5-3-2-1

Table 2
A correlation coefficient matrix showing the degree of similarity among enterococci (Ent) and *Escherichia coli* populations among sources based on their biochemical fingerprints

Sources	Number of isolates		Population similarity (Sp) coefficient							
			STP		Waterfowls		Dogs		Chickens	
	Ent	*E. coli*	Ent	*E. coli*	Ent	*E. coli*	Ent	*E. coli*	Ent	*E. coli*
STP	121	171	1.00	1.00						
Waterfowls	156	130	0.12	0.17	1.00	1.00				
Dogs	116	131	0.03	0.09	0.09	0.13	1.00	1.00		
Chickens	131	139	0.11	0.07	0.27	0.31	0.16	0.21	1.00	1.00

biochemical fingerprints and subjected to Sp-analysis. For enterococci biochemical fingerprints, the highest Sp-coefficient was found between waterfowls and chickens (0.27) followed by dogs and chickens (0.16). Enterococci populations from the STP generally showed low Sp-coefficients with those of animals. Similar patterns were observed when *E. coli* populations from STP and animals were subjected to Sp-analysis except that the Sp-coefficients among animals were higher than those of enterococci.

Of the 571 *E. coli* isolates tested, 380 were resistant to one or more antibiotics at different concentrations. The highest Sp-coefficient was found between dog and chicken (0.17). In contrast, *E. coli* populations from the STP showed low Sp-coefficients with those of animals (Table 3). The BF (11 data-points) and ARA (32 data-points) data were combined together for each isolate and a composite database (i.e. BF-ARA-database) comprising of 380 isolates was developed. When this composite database was subjected to Sp-analysis, very low Sp-coefficients were found among the sources as opposed to those values obtained by either BF or ARA-database alone (Table 4).

3.4. Source tracking

Table 5 shows the number of enterococci and *E. coli* isolates tested from each sampling site. To predict the source(s) of faecal pollution, enterococci, and *E. coli* populations from each site were compared with the populations from the sources. When using BF-database, enterococci populations from waterfowls showed the highest Sp-coefficient (mean ± SD; 0.46 ± 0.09) with water samples followed by STP (0.31 ± 0.06) (Table 5). Similar patterns were also observed when *E. coli* BF-database was used (0.32 ± 0.03; waterfowl, and 0.27 ± 0.09; STP). Both bacterial populations from all sampling sites showed higher similarities with waterfowls. In contrast, bacterial populations from dogs and chickens generally showed low similarities with water samples,

Table 3
A correlation coefficient matrix showing the degree of similarity among *Escherichia coli* populations among sources based on their antibiotic resistance patterns

Sources	Number of isolates	Population similarity (Sp) coefficient			
		STP	Waterfowls	Dogs	Chickens
STP	141	1.00			
Waterfowls	66	0.01	1.00		
Dogs	98	0.03	0.08	1.00	
Chickens	75	0.02	0.05	0.17	1.00

Table 4
A correlation coefficient matrix showing the degree of similarity among *Escherichia coli* populations among sources based on the composite database of biochemical fingerprints and antibiotic resistance patterns

Sources	Number of isolates	Population similarity (Sp) coefficient			
		STP	Waterfowls	Dogs	Chickens
STP	141	1.00			
Waterfowls	66	0.01	1.00		
Dogs	98	0.01	0.03	1.00	
Chickens	75	0.02	0.02	0.08	1.00

Table 5
Comparison of population similarity (Sp) coefficient based on biochemical fingerprinting of enterococci (Ent) and *Escherichia coli* isolates from sources and water samples collected from sites T1 to T5

Sources	Population similarity (Sp) coefficient to water samples										Mean population similarity (Sp) value	
	T1		T2		T3		T4		T5			
	Ent (n = 116)	*E. coli* (n = 85)	Ent (n = 97)	*E. coli* (n = 92)	Ent (n = 97)	*E. coli* (n = 87)	Ent (n = 98)	*E. coli* (n = 83)	Ent (n = 100)	*E. coli* (n = 88)	Ent	E. coli
STP	**0.32**	0.22	**0.27**	**0.38**	**0.35**	**0.29**	**0.40**	**0.31**	0.24	0.14	0.31 ± 0.06	0.27 ± 0.09
Waterfowl	**0.26**	**0.31**	**0.48**	**0.37**	**0.46**	**0.31**	**0.47**	**0.27**	**0.46**	**0.34**	0.46 ± 0.09	0.32 ± 0.03
Dog	0.09	0.03	0.15	**0.29**	0.11	0.17	0.13	0.07	0.10	0.13	0.11 ± 0.02	0.13 ± 0.10
Chicken	0.07	0.11	0.13	0.09	0.16	0.03	0.06	0.04	0.04	0.02	0.09 ± 0.05	0.05 ± 0.03

Populations showing similarity above 0.25 (maximum 1) have been highlighted.

Table 6
Comparison of population similarity (Sp) coefficient based on antibiotic resistance of *Escherichia coli* isolated from sources and water samples collected from sites T1 to T5

Sources	Population similarity (Sp) coefficient to water samples					Mean population similarity (Sp) value
	T1 (n = 31)	T2 (n = 69)	T3 (n = 46)	T4 (n = 52)	T5 (n = 26)	
STP	0.19	0.35	0.29	0.34	0.20	0.27 ± 0.07
Waterfowl	0.11	0.14	0.27	0.09	0.19	0.16 ± 0.07
Dog	0.02	0.11	0.07	0.04	0.11	0.07 ± 0.04
Chicken	0.06	0.05	0.06	0.01	0.03	0.04 ± 0.02

except a higher value (0.29 for *E. coli*) was found between dogs and water samples for site T2. High similarity values were also observed for both bacterial populations from STP and water samples with higher values found in sites T2 and T3 located below the submerged sewerage pipes collecting domestic wastewater.

Of the 435 *E. coli* isolates tested from the lake, 224 (51%) were resistant to one or more antibiotics. Table 6 shows the number of antibiotic resistant *E. coli* isolates at each sampling site. When *E. coli* populations from each site were compared to those of the ARA-database, the highest similarity (mean Sp-coefficient ± SD: 0.27 ± 0.07) was found between STP and water samples followed by waterfowls (0.16 ± 0.07). The Sp-coefficients between STP and water samples was higher than animals with the highest found in sampling sites T2 and T4 (Table 6). *E. coli* populations from dogs and chickens generally showed low similarities with water samples. The mean Sp-coefficients between STP and water samples was significantly higher than dogs and chickens (P = 0.01).

The BF and ARA data were combined for each environmental isolate and compared to the composite BF-ARA-database. Comparison of *E. coli* populations with this database also showed higher Sp-coefficients for STP (0.14 ± 0.09) and waterfowls (0.09 ± 0.02) with water samples than dogs (0.04 ± 0.03) and chickens (0.01 ± 0.01) (P = 0.001).

4. Discussion

Database-dependent methods have been extensively used to identify the sources of faecal pollution in various aquatic environments (Ahmed et al., 2005; Carroll et al., 2005; Hagedorn et al., 2003; Harwood et al., 2000; Scott et al., 2004). However, the application of these methods could be limited by factors such as the lack of stability of bacterial characters, statistical analysis and accuracy (Gordon, 2001; Harwood et al., 2003; Ritter et al., 2003; Stoeckel et al., 2004). Furthermore, the lack of host-specificity and cosmopolitan nature of the indicator bacteria are major drawbacks for DD methods (Gordon, 2001), although it has been reported that certain enterococci and *E. coli* could be host-specific (Ahmed et al., 2005; Clermont et al., 2008). In view of these observations, it has been suggested that DD methods could probably be suitable for a small catchment with limited sources of faecal pollution rather than a large mixed landuse catchment which generally receives pollutants from diverse sources.

In this study, waterfowls, chickens and dogs were identified as potential sources of faecal pollution. Additionally, a STP servicing the study area was also identified as a potential source as the pipes in the reticulation network was quite old. It is acknowledged that the databases used in this study comprised of relatively smaller number of isolates when compared to the large databases reported in the literature (Ahmed et al., 2005; Hagedorn et al., 2003; Harwood et al., 2000; Wiggins et al., 2003). However, our study area had only a small population of dogs (⩽20) and chickens (⩽25). In addition, the number of waterfowls observed during sampling was also low (⩽30). Therefore, developing a database containing thousands of isolates from such small number of sources was not practical. Given the small size of the catchment, and limited number of polluting sources (i.e. dogs, chickens and waterfowls), and the fact that two faecal indicators, and two source tracking methods were applied, it was hypothesised that these databases would be able to identify the dominant sources of faecal pollution.

Before evaluating the predictive capability of each database, the discriminatory ability of each method was assessed individually to ensure that each database was capable of distinguishing between sources especially STP from animals. The Sp-coefficients among the sources determined by enterococci BF-database was generally low (except between waterfowl and chicken; Sp-coefficient of 0.27). A similar pattern was also observed among the sources when *E. coli* BF-database was used except that in this case the Sp-coefficient between waterfowls and dogs was high (0.31). Nonetheless, low Sp-coefficients between bacterial populations from STP and animals indicated that both the databases could be used to distinguish between them. The Sp-coefficients among the sources determined by *E. coli* ARA-database also showed very low similarities among the sources, especially between STP and animals (<0.04) compared to BF-databases. This is because the ARA-database was comprised of relatively smaller number of isolates (i.e. 380) than the BF-database (i.e. 571) resulting in lower similarities (i.e. higher discrimination) among the sources. Furthermore, the ARA-database was based on 32 data-points for each isolate compared to the BF-database which comprised of 11 data-points for each isolate resulting in higher discrimination among the sources. The BF and ARA data was also combined for each isolate and the Sp-analysis was performed to evaluate the discriminatory ability of the composite database. This approach yielded a relatively low Sp-coefficient (<0.03) between STP and animals indicating that the

combination of two datasets is far more discriminatory than either method alone. However, it has to be noted that such low similarities as reported in this study depends on the number of sources and the number of isolates tested from each source. A database comprising of a large number of isolates from multiple sources will commonly yield higher similarities among sources due to the cosmopolitan nature of faecal indicator bacteria (Ahmed et al., 2006). Therefore, successful application of this approach could be limited to a small catchment with limited sources of faecal pollution and will require in-depth knowledge of the study area.

Enterococci and *E. coli* populations from surface waters showed higher similarities with waterfowl and STP as opposed to dog and chicken. Bacterial populations from sites T2, T3, and T4 had higher similarities to STP, and interestingly sites T2 and T3 are located near the submerged pipes. In contrast, bacterial populations from dogs and chickens showed low similarities with water samples except site T2 which showed a high similarity (0.29) between dogs and water samples. This was not surprising as this site was located near a dog kennel, and therefore, could be more exposed to faecal pollution from dogs. The ARA-database of *E. coli* also showed higher similarity between STP and water samples than animals with the highest found in site T2 (0.35). The fact that both indicator bacteria and as well as all databases showed higher similarities between the bacterial populations from STP and water samples from sites T2 and T3 (where the submerged collection pipes are located), indicated possible of wastewater leakage into the lake. We also verified these findings by using a composite BF-ARA-database which resulted in similar findings although, the overall similarities between the sources and water samples were lower than the values obtained by using BF or ARA-database alone because of the increased discriminatory power of the composite datasets. Hence, in order to obtain a clear indication of the dominant sources of pollution, a large number of isolates from the sources and water samples could be tested when a composite approach is used. However, it has to be noted that the combination approach may not be suitable because of associated cost for typing a large number of isolates from both the sources and water samples.

In conclusion, the study has demonstrated the application BF and ARA along with Sp-analysis of faecal indicator bacteria to identify the dominant sources of faecal pollution in a small recreational coastal lake. Each database individually identified the major sources of faecal pollution with a good degree of discrimination. When the datasets from these databases were combined, a higher discrimination was found. However, it still pointed the same sources of pollution. The Sp-analysis, as used in this is a simple, rapid and reliable approach and could be used for comparing bacterial populations from sources with water samples to predict the sources of faecal pollution. However, this approach should be limited to small catchments with limited sources of faecal pollution.

Acknowledgements

We thank Mr. Jack Tucker, Mr. Daniel Owen, and Mr. Peter MacDougall for their cooperation in collecting samples. This study was funded by Caloundra City Council, Queensland, Australia.

References

Ahmed, W., Neller, R., Katouli, M., 2005. 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, 4461–4468.

Ahmed, W., Neller, R., Katouli, M., 2006. Population similarity of enterococci and *Escherichia coli* in surface waters: a predictive tool to trace the sources of faecal contamination. J. Water Health 4, 347–356.

Ahmed, W., Stewart, J., Gardner, T., Powell, D., Brooks, P., Sullivan, D., Tindale, N., 2007. Sourcing faecal pollution: a combination of library-dependent and library-independent methods to identify human fecal pollution in non-sewered catchments. Water Res. 41, 3771–3779.

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

Baker, K.H., Herson, D.S., 1999. Detection and occurrence of indicator organisms and pathogens. Water Environ. Res. 37, 909–913.

Bernhard, A.E., Field, K.G., 2003. Application of a rapid method for identifying faecal pollution sources in a multi-use estuary. Water Res. 37, 909–913.

Blanch, A.R., Belanche-Muñoz, L., Edbon, J., Gantzer, C., Lucena, F., Ottoson, J., Kourtis, C., Iversen, A., Kühn, I., Mocé, L., Muniesa, M., Schwartzbrod, J., Skraber, S., Papageorgiou, G.T., Taylor, H., Wallis, J., Jofre, J., 2006. Integrated analysis of established and novel microbial and chemical methods for microbial source tracking. Appl. Environ. Microbiol. 72, 5915–5926.

Carroll, S., Hargreaves, M., Goonetilleke, A., 2005. Sourcing faecal pollution from on-site wastewater treatment systems in surface waters using antibiotic resistance analysis. J. Appl. Microbiol. 99, 471–482.

Clermont, O., Lescat, M., O'Brien, C.L., Gordon, D.M., Tenaillon, O., Denamur, E., 2008. Evidence for a human-specific *Escherichia coli*. Environ. Microbiol. 10, 1000–1006.

Edge, T.A., Hill, S., 2007. Multiple lines of evidence to identify the sources of fecal pollution at a freshwater beach in Hamilton Harbour, Lake Ontario. Water Res. 41, 3585–3594.

Field, K.G., Samadpour, M., 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Res. 41, 3517–3538.

Fong, T.T., Griffin, D.W., Lipp, E.K., 2005. Molecular assays for targeting human and bovine enteric viruses in coastal waters and their application for library-independent source tracking. Appl. Environ. Microbiol. 71, 2070–2078.

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

Hagedorn, C., Crozier, J.B., Mente, K.A., Booth, A.M., Graves, A.K., Nelson, N.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, 792–799.

Harwood, V.J., Whitlock, J., Withington, V., 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of faecal contamination in tropical waters. Appl. Environ. Microbiol. 66, 3698–3704.

Harwood, V.J., Wiggins, B., Hagedorn, C., Ellender, R.D., Gooch, J., Kern, J., Samadpour, M., Chapman, A.C.H., Robinson, B.J., Thomson, B.C., 2003. Phenotypic library-based microbial source tracking methods: efficacy in the California collaborative study. J. Water Health 1, 153–166.

Iversen, A., Kühn, I., Franklin, A., Möllby, R., 2002. High prevalence of vancomycine resistant enterococci in Swedish sewage. Appl. Environ. Microbiol. 68, 2838–2842.

Johnson, L.K., Brown, M.B., Carruthers, E.A., Ferguson, J.A., Dombek, P.E., Sadowsky, M.J., 2004. Sample size, library composition, and genotypic diversity among natural populations of *Escherichia coli* from different animals influence accuracy of determining sources of fecal pollution. Appl. Environ. Microbiol. 70, 4478–4485.

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, 3171–3177.

Love, D.C., Sobsey, M.D., 2007. Simple and rapid F+ coliphage culture, latex agglutination, and typing assay to detect and source track fecal contamination. Appl. Environ. Microbiol. 73, 4110–4118.

Manero, A., Blanch, A.R., 1999. Identification of enterococcus spp. with a biochemical key. Appl. Environ. Microbiol. 65, 4424–4430.

McLellan, S.L., 2004. Genetic diversity of *Escherichia coli* isolated from urban rivers and beach water. Appl. Environ. Microbiol. 70, 4658–4665.

Moore, D.F., Harwood, V.J., Ferguson, D.M., Lukasik, J., Hannah, P., Getrich, M., Brownell, M., 2005. Evaluation of antibiotic resistance analysis and ribotyping for identification of faecal pollution sources in an urban watershed. J. Appl. Microbiol. 99, 618–628.

Parveen, S., Murphree, R.L., Edmiston, L., Kaspar, C.W., Portier, K.M., Tamplin, M.L., 1997. Association of multiple-antibiotic resistance profiles with point and nonpoint sources of *Escherichia coli* in Apalchicola Bay. Appl. Environ. Microbiol. 63, 2607–2612.

Ritter, K.J., Carruthers, E., Carson, C.A., Ellender, R.D., Harwood, V.J., Kingsley, K., Nakatsu, C., Sadowsky, M., Shear, B., West, B., Whitlock, J.E., Wiggins, B.A., Wilbur, J.D., 2003. Assessment of statistical methods in library-based approaches in microbial source tracking. J. Water Health 1, 209–223.

Scott, T.M., Caren, J., Nelson, G.R., Jenkins, T.M., Lukasik, J., 2004. Tracking sources of faecal pollution in a south California watershed by ribotyping *Escherichia coli*: a case study. Environ Forensics 5, 15–19.

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 faecal pollution. Environ. Sci. Technol. 39, 283–287.

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

Stoeckel, D.M., Mathes, M.V., Hyer, K.E., Hagedorn, C., Kator, H., Lukasik, J., O'Brien, T.L., Fenger, T.W., Samadpour, M., Strickler, K.M., Wiggins, B.A., 2004. Comparison of seven protocols to identify faecal contamination sources using *Escherichia coli*. Environ. Sci. Technol. 38, 6107–6117.

Stoeckel, D.M., Harwood, V.J., 2007. Performance, design, and analysis in microbial source tracking studies. Appl. Environ. Microbiol. 73, 2405–2415.
Wiggins, B.A., Cash, P.W., Creamer, W.S., Dart, S.E., Garcia, P.P., Gerecke, T.M., Han, J., Henry, B.L., Hoover, K.B., Johnson, E.L., Jones, K.C., McCarthy, J.G., McDonough, J.A., Mercer, S.A., Noto, M.L., Park, H., Phillips, M.S., Purner, S.M., Smith, B.M., Stevens, E.N., Varner, A.K., 2003. Use of antibiotic resistance analysis for representativeness testing of multi-watershed libraries. Appl. Environ. Microbiol. 9, 3399–3405.