Author version of paper published as:
Ahmed, W., Yusuf, R., Hasan, I., Goonetilleke, A. and Gardner, T. (2010) **Quantitative PCR assay of sewage-associated *Bacteroides* markers to assess sewage pollution in an urban lake in Dhaka, Bangladesh.** Canadian Journal of Microbiology, Vol. 56 (10), pp. 838–845.

Quantitative PCR assay of sewage-associated *Bacteroides* markers to assess sewage pollution in an urban lake in Dhaka, Bangladesh

W. Ahmed [1, 2], R. Yusuf [3], I. Hasan [4], A. Goonetilleke [2], and T. Gardner [1, 2]

[1] *Department of Environment and Resource Management, 80 Meiers Road, Indooroopilly, Brisbane 4068, Australia*

[2] *School of Urban Development, Queensland University of Technology, GPO Box 2434, Brisbane 4001, Australia*

[3] *School of Environmental Science and Management, Independent University Bangladesh, Dhaka 1212, Bangladesh*

[4] *Department of Biochemistry and Molecular Biology, Rajshahi University, Rajshahi 6205, Bangladesh*

Running title: Sewage pollution in environmental waters

Correspondence to: W. Ahmed, Department of Natural Resources, and Water, 80 Meiers Road Indooroopilly, Brisbane 4068, Australia (e-mail: shuhat@yahoo.com).

Abstract: This paper aimed to assess the magnitude of sewage pollution in an urban lake in Dhaka, Bangladesh by using Quantitative PCR (qPCR) of sewage-associated *Bacteroides* HF183 markers. PCR was also used for the quantitative detection of ruminant wastewater-associated CF128 markers along with the enumeration of traditional fecal indicator bacteria, namely, enterococci. The number of enterococci in lake water samples ranged from 1.1×10^4 to 1.9×10^5 CFU/100 ml of water. From the 20 water samples tested, 14 (70%) and 7 (35%) were PCR positive for the HF183 and CF128 markers, respectively. The numbers of the HF183 and CF128 markers in lake water samples were 3.9×10^4 to 6.3×10^7 and 9.3×10^3 to 6.3×10^5 genomic units (GU)/100 ml of water, respectively. The high numbers of enterococci and the HF183 markers indicate sewage pollution and potential health risks to those who use the lake water for non-potable purposes such as bathing and washing clothes. This is the first study that investigated the presence of microbial source tracking (MST) markers in Dhaka, Bangladesh where diarrhoeal diseases is one of the major causes of childhood mortality. The molecular assay as used in this study can provide valuable information on the extent of sewage pollution, thus facilitating the development of robust strategies to minimise potential health risks.

Keywords: sewage pollution; fecal indicator bacteria; microbial source tracking; host-specific molecular markers; public health risk

Introduction

Fecal pollution is a major concern in relation to water bodies that are used for potable and non-potable purposes due to possible exposure to a wide array of disease causing microorganisms (Hörman et al. 2004). It has been estimated that there are approximately 1.5 million deaths per year in developing countries as a result of unsafe drinking water and inadequate sanitation (Prüss-Üstun et al. 2008). A range of point and non-point sources have been identified in research literature as the potential sources of such pollution (Ahmed et al. 2005; McLellan 2004). Microbiological safety of water is generally assessed by enumerating fecal indicator bacteria such as *E. coli* and enterococci (USEPA 2000). These indicators are commonly present in the feces of warm-blooded animals including humans, and therefore do not provide specific information regarding their sources.

The source identification of fecal pollution (i.e., humans vs. animals) in environmental waters is vitally important in order to implement appropriate mitigation strategies to minimize public health impacts. Human sourced fecal pollution poses a great health risk due to exposure to a wide array of relevant pathogens and has the lowest threshold of public acceptability. However, the identification and designation of pollution sources in environmental waters is difficult due to the cosmopolitan nature (i.e., shared by a diversity of warm-blooded animals) of fecal indicator bacteria (Field and Samadpour 2007).

Over the last decade researchers have developed a range of microbial source tracking (MST) tools that can be used to distinguish human sourced fecal pollution from that of animals (Ahmed et al. 2005; Bernhard and Field 2000, Fong et al. 2005; Parveen et al. 1997; Scott et al. 2005). In recent years, PCR detection of host-specific molecular markers such as anaerobic bacterial gene markers (Bernhard and Field 2000), bacterial toxin gene markers (Khatib et al. 2002) and viral markers (Fong et al. 2005) has gained popularity due to their ability to identify the sources of fecal pollution in environmental waters**.** Quantitative PCR (qPCR) methods have also been applied to quantify such markers in environmental waters (Kildare et al. 2007; Reischer et al. 2007; Seurinck et al. 2005). Among these markers, sewage-associated *Bacteroides* are reported to be host-specific (i.e., human) and geographically stable. They have been widely used to detect the sources of sewage pollution in developed countries such as USA (Bernhard et al. 2003), France (Gourmelon et al. 2007), Portugal, UK, Ireland (Gawler et al. 2007), and Australia (Ahmed et al. 2008). Quantitative PCR methods have also been developed to quantify *Bacteroides* markers in environmental waters in Belgium (Seurinck et al. 2006), Hawaii (Betancourt and Fujioka 2006), Japan (Okabe et al. 2007), Austria (Reischer et al. 2007), and USA (Layton et al. 2006). However, to-date none of these MST methods has been applied for the identification of fecal pollution sources in developing countries, particularly in Southeast Asia such as Bangladesh where diarrhoeal diseases, namely,

cholera due to poor water quality is a serious health concern. The World Health Organization has estimated that in 2002 in Bangladesh, 68 out of 1,000 deaths were due to diarrhoeal disease namely cholera caused by *Vibrio cholerae* (Death and DALY estimates 2002). The source of the infection is typically cholera patients when their untreated fecal materials enter into the environmental waters or drinking water supplies.

The primary aim of the study discussed in the paper was to assess the magnitude of sewage pollution in an urban lake in Dhaka, Bangladesh by using qPCR of sewage-associated *Bacteroides* HF183 markers. Water samples were also tested for ruminant wastewater-specific CF128 markers using qPCR along with the enumeration of traditional fecal indicator bacteria, namely, enterococci using culture based method. The results suggest that the study lake is highly polluted with human-derived sewage and could pose human health risks.

Materials and methods

Host group sampling and DNA extraction

To determine the sensitivity and specificity of the HF183 and CF128 markers, 45 fecal samples were collected from five host groups, including humans, cattle, dogs, cats, and chickens. Human fecal samples (n= 15) were collected from healthy individuals. Cattle fecal samples (n = 12) were collected from two farms located outside the City, Dhaka. Dog (n = 5) and cat (n = 6) fecal samples were collected from different houses in the neighbourhood. Chicken fecal samples (n = 7) were collected from a local poultry market. A fresh fecal sample (approximately 0.2 - 0.5 g) was collected from each individual animal, transported on ice to the laboratory, and processed within 8 h. DNA was extracted from fresh feces (i.e., 150 – 200mg) by using QIAamp DNA stool kit (Qiagen, Valencia, CA, USA). Extracted DNA samples were resuspended in 200 µl AE buffer, and stored at -80°C.

Study Lake

The study lake is located between Baridhara Diplomatic Zone and Gulshan Residential areas of Dhaka, the capital of Bangladesh (Figure 1). The approximate coordinates are 23°48’4”N and 90°25’2”E. The lake is approximately 1,150 meters long and 100 meters wide. The local residents use lake water for bathing and washing clothes. The potential sources of the pollution include sewage flow via stormwater drains and the discharge of fecal materials from slum-like establishments located along the Gulshan part of the lake. During a sanitary survey, a small number of cattle (n =2) and stray dogs were observed which can also be potential sources of fecal pollution. After rainfall (i.e., > 100 mm), the lake overflows and submerges surrounding residential areas for a few days due to poor drainage conditions.

Water sampling

Environmental water samples (n = 20) were collected on the 8th of August, 2009 from 20 locations (sites BL1 – BL20) in the lake (Figure 1). Samples were collected from the bank of the Baridhara Diplomatic Zone due to ease of access. Grab samples were taken from 30 cm below the water surface two weeks after a rainfall event >150 mm. Sterile 500 ml plastic containers were used for sample collection, and immediately transported to the adjacent laboratory at the Independent University of Bangladesh (IUB) for microbiological analysis.

Enumeration of enterococci

The membrane filtration method was used to process the environmental samples for enterococci enumeration. Sample serial dilutions were made and filtered through 0.45 µm pore size (47 mm diameter) nitrocellulose membranes (Advantec, Tokyo, Japan), and placed on membrane-Enterococcus indoxyl-D-glucoside (mEI agar) (Difco, Detroit, MI, USA) for the isolation of enterococci. Agar plates were incubated at 41°C for 48 h. For bacterial enumeration, all the samples were tested in triplicate.

DNA extraction

Water samples were processed according to a previously published method described elsewhere (Bernhard and Field 2000). Briefly, 100 ml of water samples were filtered through 0.45 µm pore size nitrocellulose membranes (Advantec) and DNA was extracted by using DNeasy blood and tissue kit (Qiagen). Extracted DNA were resuspended in 200 µl AE buffer, and stored at -80°C. Quantitative PCR analysis was performed at University of the Sunshine Coast (USC), Australia. DNA isolated from fecal and water samples were lyophilized prior to transport to USC. At USC, the samples were resuspended in 200 µl buffer AE (Qiagen) and stored at -20°C. Prior to analysis, the concentration of DNA was determined using a spectrophotometer.

PCR assays

Quantitative PCR was used to quantify the HF183 and CF128 markers. HF183 markers were amplified by using forward primer (5´- ATC ATG AGT TCA CAT GTC CCG - 3´) (Bernhard and Field 2000) and a reverse primer (5´- TAC CCC GCC TAC TAT CTA ATG - 3´) (Seurinck et al. 2005). CF128 markers were amplified using previously published primer set (5´- CCA ACY TTC CCG WTA CTC - 3´; forward primer and 5´- CAA TCG GAG TTC TTC GTG - 3´; reverse primer) (Bernhard and Field 2000). The qPCR mixture contained 12.5 µl of Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen, Carlsbad, CA), 300 nM of each primer, and 5 µl of template DNA. The HF183 qPCR consisted of 2 min at 50°C, 10 min at 95°C followed by 45 cycles of 30 s at 95°C, 1 min at 53°C, and then extension of 1 min at 60°C. The CF128 qPCR consisted of 2 min at 50°C, 15 min at 95°C followed by 40 cycles of 30 s at 94°C, 30 s at 59°C, and then extension of 1 min at 72°C.

The HF183 and CF128 positive controls were isolated from sewage and cattle wastewater, respectively. The PCR-amplified product was purified using a QIAquick PCR purification kit (Qiagen) and cloned into a pGEM-T Easy vector system (Promega, Madison, WI), transferred into *E. coli* JM109-competent cells, and plated on LB agar plates containing ampicillin, IPTG, and X-gal as recommended by the manufacturer. Recombinant plasmids with corresponding inserts were purified using a plasmid mini kit (Qiagen). DNA sequencing was carried out at the Australian Genome Research Facility at St. Lucia, Queensland, Australia. Standards for qPCR of the HF183 and CF128 were prepared from the plasmid DNA. A tenfold serial dilution was prepared from the plasmid DNA, ranging from 10^6 to 10^0 copies per µl of DNA. The reproducibility of the qPCR assays was assessed by determining intra-assay repeatability and inter-assay variations. The Coefficient of Variation (CV) was calculated using five dilutions (10^6 to 10^2 gene copies) of the HF183 and CF128 plasmid DNA. Each dilution was quantified in triplicate.

For the qPCR standards, the concentration was plotted against the cycle number at which the fluorescence signal increased above the threshold value (C_T value). The amplification efficiency (E) was determined by running the standards, and was estimated from the slope of the standard curve by the formula $E = (10^{-1/slope}) - 1$. A reaction with 100% efficiency generates a slope of -3.32. To separate the specific product from non-specific products, including primer dimers, melting curve analysis was performed for each PCR run. During melting curve analysis, temperature was increased from 55°C to 95°C at approximately 2°C per min. Non-specific products were not observed in the melting curve analysis.

Quantitative PCR limit of detection in freshwater

To determine the qPCR limit of detection of the HF183 markers, sewage samples were collected from a Sewage treatment plant (STP) in Brisbane, Australia and suspended in freshwater collected from the Brisbane River (n = 3) at a ratio 1:1 (62.5 ml water:62.5 ml fresh sewage). Sample serial dilutions (10^{-1} to 10^{-10}) were made and DNA extraction was performed for each dilution and tested with the qPCR. The numbers of enterococci were enumerated using the membrane filtration method for each dilution according to the method described above.

Recovery efficiency

The recovery efficiency was determined by spiking autoclaved freshwater samples (n=3) with known numbers of *Bacteroides vulgatus* ATCC 8482. For recovery efficiency, *B. vulgatus* was used due to the fact that the *Bacteroides* species that carries the HF183 and CF128 markers has never been isolated. *B. vulgatus* was grown overnight according to a previously published method (Seurinck *et al.* 2005), and cell numbers were determined by using flow cytometry. Ten fold serial dilutions ($2.1 \pm 0.4 \times 10^7$, $2.1 \pm 0.4 \times 10^6$, $2.1 \pm 0.4 \times 10^5$ and 2.1 ± 0.4

× 10^4 per 100 ml freshwater) of the stationary phase of *B. vulgatus* were made, and filtered through the membranes. DNA extraction was performed according to the method described above. For the quantification of *B. Vulgates*, a bacterial qPCR assay was used. The primers and cycling condition of this assay is described elsewhere (Boon et al. 2003). A 10-fold dilution series (ranging from 10^6 to 10^2) of DNA extract from *B. vulgatus* were used as qPCR standards for the *B. vulgatus* bacterial assay. The recovery efficiency was calculated using the following equation: Recovery (%) = (No. of cells after filtration/No. of cells before filtration) x 100.

Check for PCR inhibitors in DNA isolated from environmental waters

An experiment was conducted to determine the potential presence of PCR inhibitory substances in representative DNA samples (n = 5) isolated from environmental waters collected from the study lake. DNA samples were serially diluted (i.e., 10-fold and 100-fold), and finally undiluted, 10-fold and 100-fold dilutions were spiked with 10^3 gene copies of the ruminant wastewater associated CF128 *Bacteroides* markers (Bernhard and Field 2000). CF128 markers were chosen as the prevalence of these markers assumed to be low in samples collected from the lake compared to the HF183 markers. Before spiking, all five DNA samples were tested to determine the background level of the CF128 markers. None of the samples were found to be positive for the CF128 markers. The threshold cycle (C_T) values of the spiked DNA samples were compared to those of the DNA samples from distilled water spiked with the 10^3 gene copies of the CF128 markers.

DNA sequencing

To verify the identity of the PCR product obtained from lake water samples using the HF183 and CF128 *Bacteroides* primers, the PCR-amplified sequences from the HF183 and CF128 primer sets were cloned into the pGEM®-T Easy Vector system (Promega, Madison, WI, USA). Plasmids were extracted using the QIAprep Spin® Miniprep kit (Qiagen). Bidirectional sequences were obtained using T7 and SP6 long sequencing primer targeting sites on either side of the insert. DNA sequencing was carried out and analysed using Bioware Jellyfish Software.

Results

Sensitivity and Specificity

The HF183 markers were detected in 13 human fecal samples out of 15 samples tested, and was not detected in 28 out of 30 animal fecal samples. However, fecal samples from one dog and one cat were positive. The CF128 markers were detected in 9 out of 12 cattle fecal samples and were not detected in fecal samples from humans and other animal host groups. The numbers of the HF183 and CF128 markers in positively identified samples are

given in Table 1. The overall sensitivity of the HF183 primer in detecting human-specific HF183 markers in human fecal samples was 0.87. For the CF128 markers this figure was 0.75. The overall specificity of the HF183 markers to differentiate between sewage and animal feces was 0.93. For CF128 this figure was 1.0.

Quantitative PCR standards and reproducibility

Tenfold serial dilutions of the HF183 and CF128 plasmid DNA were analysed in order to determine the reaction efficiencies. The standard curves had a linear range of quantification from 10^6 to 10^1 HF183 and CF128 markers per µl of DNA extract. The amplification efficiencies were between 95-99% for both markers. The correlation coefficient (r^2) was > 0.97. The amplification of the correct PCR products were verified by analysing the melting curves. The reproducibility of the qPCR assays was determined by assessing intra-assay and inter-assay Coefficient of Variation (CV) of the standards. These values were less than 2% and 5% respectively for both assays, indicating high reproducibility.

Limit of detection and recovery efficiency

Prior to testing of environmental samples, the limit of detection of qPCR was evaluated for the HF183 markers. The qPCR was able to detect the HF183 markers up to dilution 10^{-8} in freshwater spiked with sewage. At this dilution, no culturable enterococci were found. The estimated recovery efficiency in freshwater samples ranged from 86% to 50% with the greatest variability occurring at lower cell counts (Table 2). The mean detection efficiency was 68% ± 10%.

PCR inhibitors

For the spiked distilled water, the mean C_T value for the CF128 marker was 27.3 ± 0.5. For surface water samples, the mean C_T values for undiluted, 10-fold and 100-fold diluted DNA are shown in Table 3. One-way ANOVA was performed to determine the differences between the C_T values obtained for DNA isolated from distilled water and those obtained for surface water samples. However, significant ($P < 0.001$) differences were observed between the C_T values for spiked distilled water DNA and undiluted DNA from lake water samples, indicating that the undiluted DNA extracted from the lake contained PCR inhibitory substances. Significant ($P < 0.001$) differences were also observed between spiked distilled water DNA and 10-fold diluted DNA. However, no significant differences ($P > 0.05$) were observed between the C_T values for spiked distilled water DNA and 100-fold diluted DNA indicating that 100-fold dilution of DNA is required to remove the effects of PCR inhibitory substances.

Numbers of enterococci

The number of enterococci in surface water samples ranged from 1.1 x 10^4 to 1.9 x 10^5 CFU/100 ml of water. Site BL18 had the lowest number (1.1 x 10^4 CFU per 100 ml) and site BL12 had the highest number (1.9 x 10^5 CFU per 100 ml) (Table 4). The numbers of enterococci were also extremely high in sites BL3 and BL9. These are located near stormwater pipes draining into the lake. All water samples tested exceeded the recommended guidelines values for fresh and marine waters (USEPA 1986, WHO 2003).

Quantitative PCR results for environmental water samples

Of the 20 samples tested, 14 (70%) were positive for the HF183 markers and the number in these positively identified samples ranged from 3.9 x 10^4 to 6.3 x 10^7 genomic units (GU) per 100 ml of water (Table 4). The highest number (6.3 x 10^7 GU per 100 ml) of markers was found in water samples collected from site BL3 followed by BL6 (6.1 x 10^7) and BL9 (5.3 x 10^7). Of the 20 water samples tested, 7 (35%) were PCR positive for the ruminant wastewater-associated markers. The number of the CF128 markers in these positively identified samples ranged from 9.3 x 10^3 to 6.3 x 10^5 genomic units (GU) per 100 ml of water. In all, 4 (20%) samples were positive for both HF183 and CF128 markers. Up to 3 amplicons (i.e., amplified from environmental DNA) of each marker were sequenced and verified to ensure that they were > 97% identical to the published sequence.

Discussion

The high numbers of the HF183 markers in lake water samples indicate the presence of sewage pollution. Stormwater pipes and slum-like establishments in the vicinity may have introduced sewage pollution to the lake. Prior to water sampling, a sanitary survey was undertaken to identify possible sources of fecal pollution other than human sources. A small number of stray dogs and few cattle were observed near the lake. The HF183 markers have been isolated occasionally from the feces of companion animals such as dogs and cats (Ahmed *et al.* 2008). As such, prior to field application, the specificity of the HF183 and CF128 markers were evaluated by testing fecal samples from five host groups. It is acknowledged that the number of samples tested for specificity assay was small. Nonetheless, the HF183 marker displayed high specificity. These markers were only detected in two fecal samples from companion animals and the numbers were low compared to human feces (see Table 1). The high specificity of the HF183 markers in distinguishing human from animal feces has been reported from many locations throughout the world (Ahmed et al. 2008; Gourmelon et al. 2007; Seurinck et al. 2006). To rule out the possibility of dog-sourced fecal pollution, dog *Bacteroides* BacCan markers were used to

identify dog sourced fecal pollution in water samples (Kildare et al. 2007). Only two samples from sites BL9 and BL10 gave positive signals (data not shown). Therefore, considering the limited number of non-point sources of fecal pollution and high-specificity of the HF183 markers, the most likely source of the HF183 marker in the lake is considered to be from fecal discharges from slum-like establishments and storm water pipes.

All water samples were screened for the HF183 and CF128 markers using qPCR. A significant number of water samples (i.e., 70%) were positive for the HF183 markers. However, these markers could not be detected in sites BL17, BL18, BL19 and BL20. This could be due to the fact that these sites were located approximately 1 km from the suspected sources of sewage pollution. The CF128 markers were detected in seven samples, however, the numbers were much lower than those found in cattle feces.

The numbers of the HF183 markers in human feces and sewage (collected from influent) range from $8.4 \pm 0.1 \times 10^5$ to $7.2 \pm 1.1 \times 10^9$ per gram of wet feces and $5.9 \pm 0.7 \times 10^9$ to $3.1 \pm 0.3 \times 10^{10}$ GU per litre, respectively (Seurinck et al. 2006). Similar numbers of the HF183 have been detected in the feces collected from humans in the present study (see Table 1). In this study, the number of the HF183 markers in lake water samples ranged from 3.9×10^4 to 6.3×10^7 GU per 100 ml of water, which is one to two orders of magnitude lower than human feces or sewage. Such data is cause for significant concern and indicate the presence of sewage pollution. The number of enterococci was also high which indicated a high level of fecal pollution. The numbers of enterococci were higher near the stormwater pipes, thus confirming that these pipes are contributing to the fecal load in the lake. It has to be noted that the samples were collected after a rainfall event when the fecal indicator numbers are expected to be high in environmental waters.

Nonetheless, the high levels of fecal indicator bacteria and the HF183 *Bacteroides* markers indicate sewage pollution and health risks to those who use the lake water for non-potable uses such as bathing and washing clothes. The overflow from the lake after heavy rain could also pose serious health risks to the residents surrounding the lake due to the possibility of cross contamination with the reticulated supply water. The significant increase in diarrhoeal cases after flood or storm events are common in Bangladesh (Hashizume et al. 2008).

The qPCR assays as employed in this study could provide a scientifically reliable approach to source tracking and for assessing the magnitude of fecal pollution within a short time frame and generate valuable data which could be used to develop strategies to minimise fecal pollution or at the minimum to educate local residents regarding the health risks associated with exposure to such pollution. Local laboratories could employ such methods for rapid assessment of sewage pollution in environmental waters. In the near future we intend to

evaluate the specificity of the markers thoroughly by testing more animal species as well as testing more water samples from the study area at different time periods.

Acknowledgements

This study was funded by Queensland Department of Environment and Resource Management (DERM). This was a joint project between DERM, Queensland University of Technology (QUT) and the Independent University of Bangladesh (IUB). We thank IUB for providing laboratory space. We also thank Ms. Warda Ashraf for processing the DNA samples.

References

Ahmed, W., Neller, R., and 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., Stewart, J., Powell, D., and Gardner, T. 2008. Evaluation of *Bacteroides* markers for the detection of human faecal pollution. Lett. Appl. Microbiol. 46: 237-242.

Bernhard, A.E., and Field, K.G. 2000. A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. 66: 4571–4574.

Bernhard, A.E., Goyard, T., Simonich, M., and Field, K.G. 2003. Application of a rapid method for identifying fecal pollution sources in a multi-use estuary. Water Res. 37: 909–913.

Betancourt, W.Q., and Fujioka, R.S. 2006. *Bacteroides* spp. as reliable marker of sewage contamination in Hawaii's environmental waters using molecular techniques. Wat. Sci. Technol. 54: 101-107.

Boon, N., Top, E.M., Verstraete, W., and Siciliano, S.D. 2003. Bioaugmentation as a tool to protect the structure and function of an activated sludge microbial community against a 30-chloroaniline shock load. Appl. Environ. Microbiol. 69: 1511-1520.

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

Fong, T.T., Griffin, D.W., and 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.

Gawler, A.H., Beecher, J.E., Brandão, J., Carroll, N.M., Falcão, L., Gourmelon, M., Masterson, B., Nunes, B., Porter, J., Rincé, A., Rodrigues, R., Thorp, M., Walters, J.M., and Meijer, W.J. 2007. Validation of host-specific

Bacteriodales 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. Water Res. 41: 3780-3784.

Gourmelon, M., Caprais, M.P., Ségura, R., Le Mennec, C., Lozach, S., Piriou, J.Y., and Rincé, A. 2007. Evaluation of two library-independent microbial source tracking methods to identify sources of fecal contamination in French estuaries. Appl. Environ. Microbiol. 73: 4857-4866.

Hashizume, M., Wagatsuma, Y., Faruque, A.S.G., Sack, D.A., Hayashi, T., Hunter, P.R., and Armstrong, B. 2008. Factors determining vulnerability diarrhoea during and after severe floods in Bangladesh. J. Water. Health. 6: 323-332.

Hörman, A., Rimhanen-Finne, R., Maunula, L., von Bonsdorff, C.-H., Torvela, N., Heikinheimo, A., and Hanninen, M.-L. 2004. *Campylobacter* spp., *Giardia* spp., *Cryptosporidium* spp., Noroviruses, and indicator organisms in surface water in south-western Finland, 2000–2001. Appl. Environ. Microbiol. 70: 87–95.

Khatib, L.A., Tsai, Y.L. and Olson, B.H. 2002. A biomarker for the identification of cattle fecal pollution in water using the LTIIa toxin gene from enterotoxigenic *E. coli*. Appl. Microbiol. Biotechnol. 59: 97–104.

Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., and Wuertz, S. 2007. 16S r RNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: a Bayesian approach. Water Res. 41: 3701-3715.

Layton, A., MacKay, L., Williams, D., Garrett, V., Gentry, R. and Sayler G. 2006. Development of *Bacteroides* 16S rRNA Gene TaqMan-Based Real-Time PCR Assays for Estimation of Total, Human, and Bovine Fecal Pollution in Water. Appl. Environ. Microbiol.72: 4214 - 4224.

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

Okabe, S., Okayama, N., Savichtcheva, O., and Ito, T. 2007. Identification and quantification of host-specific *Bacteroides-Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in freshwaters. Appl. Microbiol. Biotechnol. 74: 890–901.

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

Prüss-Üstun, A., Bos, R., Gore, F., and Bartam, J. 2008. Safer Water, Better Health: Costs, Benefits and Sustainability of Interventions to Protect and Promote Health. Geneva: World Health Organization.

Reischer, G.H., Kasper, D.C., Steinborn, R., Farnleitner, A.H., and Mach, R.L. 2007. A quantitative real-time PCR assay for the highly sensitive and specific detection of human faecal influence in spring water from a large alpine catchment area. Lett. Appl. Microbiol. 44: 351-356.

Scott, T.M., Jenkins, T.M., Lukasik, J., and Rose, J.B. 2005. Potential use of a host associated molecular marker in *Enterococcus faecium* as an index of human pollution. Environ. Sci. Technol. 39: 283-287.

Seurinck, S., Verdievel, M., Verstraete, W. and Siciliano, S.D. 2006. Identification of human fecal pollution sources in a coastal area: a case study at Oostende (Belgium). J. Water Health 4: 167–175.

U.S. Environmental Protection Agency. 1986. Ambient water quality criteria for bacteria. EPA440/5-84-002. Washington, DC 20460:Office of Water Regulations and Standards Division.

U.S. Environmental Protection Agency. 2000. Improved Enumeration Methods for the Recreational Water Quality Indicators: Enterococci and Escherichia coli. EPA/821/R-97/004. Office of Science and Technology, Washington, DC, 55p.

WHO. 2003. Guidelines for safe recreational-water environments, coastal and fresh-waters. Geneva, Switzerland.

WHO 2006. Mortality country fact sheet.

http://www.who.int/whosis/mort/profiles/mort_searo_bgd_bangladesh.pdf

449 **Table 1:** Polymerase chain reaction results for the numbers of sewage associated *Bacteroides* HF183 and

450 ruminant-associated CF128 markers in host groups in Dhaka, Bangladesh

451

Host-groups	No. of samples	Sewage-associated *Bacteroides* HF183			Ruminant wastewater-associated *Bacteroides* CF128		
		Positive	Numbers per 100 mg of feces	Negative	Positive	Numbers per 100 mg of feces	Negative
Humans	15	13	1.2×10^6 - 3.9×10^8	2	0	0	15
Cattle	12	0	0	12	9	5.1×10^6 - 6.3×10^7	3
Dogs	5	1	7.8×10^4	4	0	0	5
Cats	6	1	4.6×10^3	5	0	0	6
Chickens	7	0		7	0	0	7

452

453 **Table 2:** Detection efficiency with the qPCR assay for autoclaved freshwater samples with known numbers of *B.*

454 *vulgatus* cells

455

Spiked cells per 500 ml of water	Detection efficiency ± SD (%)
$2.1 \pm 0.4 \times 10^7$	80 ± 6
$2.1 \pm 0.4 \times 10^6$	73 ± 7
$2.1 \pm 0.4 \times 10^5$	66 ± 9
$2.1 \pm 0.4 \times 10^4$	46 ± 6

456

457

Table 3: Effects of PCR inhibitors on the PCR detection of spiked cattle wastewater associated CF128 *Bacteroides* markers in DNA isolated from surface water samples as opposed to distilled water samples

Samples	Threshold cycle (C_T) value for the real-time PCR		
	Undiluted DNA	10-fold dilution	100-fold dilution
Distilled water	27.3 ± 0.5	-	-
BL4	35.6 ± 0.8	31.1 ± 0.4	27.3 ± 0.6
BL5	34.9 ± 0.7	30.0 ± 0.6	28.1 ± 0.8
BL6	36.2 ± 0.8	32.1 ± 0.6	27.9 ± 0.6
BL7	35.3 ± 0.6	31.4 ± 0.9	27.6 ± 0.7
BL8	35.1 ± 0.9	31.6 ± 0.6	27.4 ± 0.6

476 **Table 4:** Numbers of enterococci and sewage-associated *Bacteroides* HF183 markers in surface waters samples

477 collected from Baridhara Lake, Dhaka, Bangladesh

478

Sampling sites	Number (cfu) of enterococci per 100 ml	Number of human-specific *Bacteroides* HF183 marker per 100 ml	Presence/absence of cattle-specific *Bacteroides* CF128 markers per 100 ml
BL1	9.0×10^4	3.3×10^6	1.2×10^4
BL2	8.0×10^4	2.1×10^5	ND
BL3	1.8×10^5	6.3×10^7	3.1×10^4
BL4	5.2×10^4	2.1×10^7	ND
BL5	4.0×10^4	8.9×10^6	ND
BL6	5.5×10^4	6.1×10^7	ND
BL7	6.0×10^4	2.1×10^6	ND
BL8	3.5×10^4	4.8×10^5	ND
BL9	1.3×10^5	5.3×10^7	5.1×10^4
BL10	3.2×10^4	9.1×10^5	ND
BL11	2.2×10^4	ND	ND
BL12	1.9×10^5	5.1×10^7	6.3×10^5
BL13	2.2×10^4	1.1×10^5	ND
BL14	4.7×10^4	3.9×10^4	ND
BL15	1.9×10^4	ND	ND
BL16	5.6×10^4	1.0×10^5	ND
BL17	2.6×10^4	ND	3.7×10^4
BL18	1.1×10^4	ND	9.3×10^3
BL19	2.0×10^4	ND	ND
BL20	1.4×10^4	ND	7.3×10^4

479

480 ND: Not detected

481

482

483

484

485

486

487

488

489

490

491

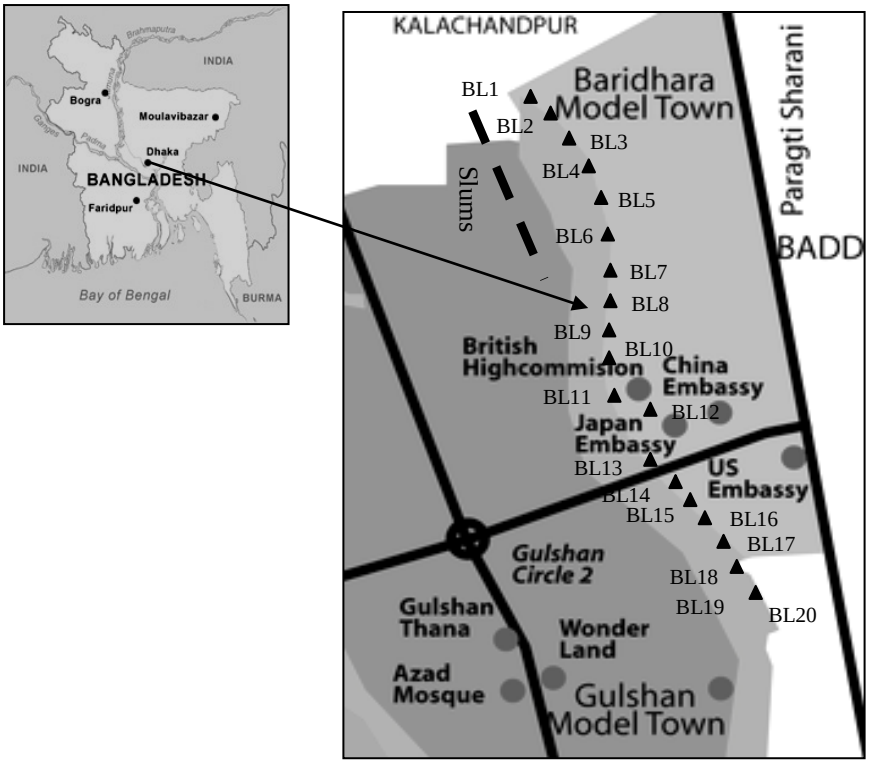

Figure 1: A map showing the sampling sites (▲) on the Baridhara Lake, Dhaka, Bangladesh