

1 Author version of paper published as:

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

5 Copyright 2010 NRC Research Press, National Research Council of Canada

6
7
8 **Quantitative PCR assay of sewage-associated *Bacteroides* markers to**
9 **assess sewage pollution in an urban lake in Dhaka, Bangladesh**

10
11
12
13
14 **W. Ahmed^{1,2}, R. Yusuf³, I. Hasan⁴, A. Goonetilleke², and T. Gardner^{1,2}**

15
16
17
18 ¹ *Department of Environment and Resource Management, 80 Meiers Road, Indooroopilly, Brisbane*
19 *4068, Australia*

20 ² *School of Urban Development, Queensland University of Technology, GPO Box 2434, Brisbane*
21 *4001, Australia*

22 ³ *School of Environmental Science and Management, Independent University Bangladesh, Dhaka*
23 *1212, Bangladesh*

24 ⁴ *Department of Biochemistry and Molecular Biology, Rajshahi University, Rajshahi 6205,*
25 *Bangladesh*

26
27
28 Running title: Sewage pollution in environmental waters

29
30
31
32 *Correspondence to: W. Ahmed, Department of Natural Resources, and Water, 80 Meiers Road*
33 *Indooroopilly, Brisbane 4068, Australia (e-mail: shuhat@yahoo.com).*

34 **Abstract:** This paper aimed to assess the magnitude of sewage pollution in an urban lake in Dhaka, Bangladesh
35 by using Quantitative PCR (qPCR) of sewage-associated *Bacteroides* HF183 markers. PCR was also used for the
36 quantitative detection of ruminant wastewater-associated CF128 markers along with the enumeration of
37 traditional fecal indicator bacteria, namely, enterococci. The number of enterococci in lake water samples
38 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
39 (35%) were PCR positive for the HF183 and CF128 markers, respectively. The numbers of the HF183 and
40 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
41 (GU)/100 ml of water, respectively. The high numbers of enterococci and the HF183 markers indicate sewage
42 pollution and potential health risks to those who use the lake water for non-potable purposes such as bathing and
43 washing clothes. This is the first study that investigated the presence of microbial source tracking (MST)
44 markers in Dhaka, Bangladesh where diarrhoeal diseases is one of the major causes of childhood mortality. The
45 molecular assay as used in this study can provide valuable information on the extent of sewage pollution, thus
46 facilitating the development of robust strategies to minimise potential health risks.

47
48
49
50
51
52
53
54
55
56
57
58
59
60

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

63 **Introduction**

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

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

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

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

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

102

103 **Materials and methods**

104 **Host group sampling and DNA extraction**

105 To determine the sensitivity and specificity of the HF183 and CF128 markers, 45 fecal samples were collected
106 from five host groups, including humans, cattle, dogs, cats, and chickens. Human fecal samples ($n = 15$) were
107 collected from healthy individuals. Cattle fecal samples ($n = 12$) were collected from two farms located outside
108 the City, Dhaka. Dog ($n = 5$) and cat ($n = 6$) fecal samples were collected from different houses in the
109 neighbourhood. Chicken fecal samples ($n = 7$) were collected from a local poultry market. A fresh fecal sample
110 (approximately 0.2 - 0.5 g) was collected from each individual animal, transported on ice to the laboratory, and
111 processed within 8 h. DNA was extracted from fresh feces (i.e., 150 – 200mg) by using QIAamp DNA stool kit
112 (Qiagen, Valencia, CA, USA). Extracted DNA samples were resuspended in 200 μ l AE buffer, and stored at
113 -80°C.
114

115 **Study Lake**

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

124 **Water sampling**

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

130 **Enumeration of enterococci**

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

136 **DNA extraction**

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

144 **PCR assays**

145 Quantitative PCR was used to quantify the HF183 and CF128 markers. HF183 markers were amplified by using
146 forward primer (5'- ATC ATG AGT TCA CAT GTC CCG - 3') (Bernhard and Field 2000) and a reverse primer
147 (5'- TAC CCC GCC TAC TAT CTA ATG - 3') (Seurinck et al. 2005). CF128 markers were amplified using
148 previously published primer set (5'- CCA ACY TTC CCG WTA CTC - 3'; forward primer and 5'- CAA TCG
149 GAG TTC TTC GTG - 3'; reverse primer) (Bernhard and Field 2000). The qPCR mixture contained 12.5 μl of
150 Platinum SYBR Green qPCR SuperMix-UDG (Invitrogen, Carlsbad, CA), 300 nM of each primer, and 5 μl of
151 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
152 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
153 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.

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

165 For the qPCR standards, the concentration was plotted against the cycle number at which the
166 fluorescence signal increased above the threshold value (C_T value). The amplification efficiency (E) was
167 determined by running the standards, and was estimated from the slope of the standard curve by the formula $E =$
168 $(10^{-1/\text{slope}}) - 1$. A reaction with 100% efficiency generates a slope of -3.32. To separate the specific product from
169 non-specific products, including primer dimers, melting curve analysis was performed for each PCR run. During
170 melting curve analysis, temperature was increased from 55°C to 95°C at approximately 2°C per min. Non-
171 specific products were not observed in the melting curve analysis.

172 **Quantitative PCR limit of detection in freshwater**

173 To determine the qPCR limit of detection of the HF183 markers, sewage samples were collected from a Sewage
174 treatment plant (STP) in Brisbane, Australia and suspended in freshwater collected from the Brisbane River ($n =$
175 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
176 extraction was performed for each dilution and tested with the qPCR. The numbers of enterococci were
177 enumerated using the membrane filtration method for each dilution according to the method described above.

178 **Recovery efficiency**

179 The recovery efficiency was determined by spiking autoclaved freshwater samples ($n=3$) with known numbers of
180 *Bacteroides vulgatus* ATCC 8482. For recovery efficiency, *B. vulgatus* was used due to the fact that the
181 *Bacteroides* species that carries the HF183 and CF128 markers has never been isolated. *B. vulgatus* was grown
182 overnight according to a previously published method (Seurinck *et al.* 2005), and cell numbers were determined
183 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

184 $\times 10^4$ per 100 ml freshwater) of the stationary phase of *B. vulgatus* were made, and filtered through the
185 membranes. DNA extraction was performed according to the method described above. For the quantification of
186 *B. Vulgates*, a bacterial qPCR assay was used. The primers and cycling condition of this assay is described
187 elsewhere (Boon et al. 2003). A 10-fold dilution series (ranging from 10^6 to 10^2) of DNA extract from *B.*
188 *vulgatus* were used as qPCR standards for the *B. vulgatus* bacterial assay. The recovery efficiency was calculated
189 using the following equation: Recovery (%) = (No. of cells after filtration/No. of cells before filtration) x 100.

190 **Check for PCR inhibitors in DNA isolated from environmental waters**

191 An experiment was conducted to determine the potential presence of PCR inhibitory substances in representative
192 DNA samples (n = 5) isolated from environmental waters collected from the study lake. DNA samples were
193 serially diluted (i.e., 10-fold and 100-fold), and finally undiluted, 10-fold and 100-fold dilutions were spiked
194 with 10^3 gene copies of the ruminant wastewater associated CF128 *Bacteroides* markers (Bernhard and Field
195 2000). CF128 markers were chosen as the prevalence of these markers assumed to be low in samples collected
196 from the lake compared to the HF183 markers. Before spiking, all five DNA samples were tested to determine
197 the background level of the CF128 markers. None of the samples were found to be positive for the CF128
198 markers. The threshold cycle (C_T) values of the spiked DNA samples were compared to those of the DNA
199 samples from distilled water spiked with the 10^3 gene copies of the CF128 markers.

200 **DNA sequencing**

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

207

208 **Results**

209 **Sensitivity and Specificity**

210 The HF183 markers were detected in 13 human fecal samples out of 15 samples tested, and was not detected in
211 28 out of 30 animal fecal samples. However, fecal samples from one dog and one cat were positive. The CF128
212 markers were detected in 9 out of 12 cattle fecal samples and were not detected in fecal samples from humans
213 and other animal host groups. The numbers of the HF183 and CF128 markers in positively identified samples are

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

217 **Quantitative PCR standards and reproducibility**

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

225 **Limit of detection and recovery efficiency**

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

231 **PCR inhibitors**

232 For the spiked distilled water, the mean C_T value for the CF128 marker was 27.3 ± 0.5 . For surface water
233 samples, the mean C_T values for undiluted, 10-fold and 100-fold diluted DNA are shown in Table 3. One-way
234 ANOVA was performed to determine the differences between the C_T values obtained for DNA isolated from
235 distilled water and those obtained for surface water samples. However, significant ($P < 0.001$) differences were
236 observed between the C_T values for spiked distilled water DNA and undiluted DNA from lake water samples,
237 indicating that the undiluted DNA extracted from the lake contained PCR inhibitory substances. Significant ($P <$
238 0.001) differences were also observed between spiked distilled water DNA and 10-fold diluted DNA. However,
239 no significant differences ($P > 0.05$) were observed between the C_T values for spiked distilled water DNA and
240 100-fold diluted DNA indicating that 100-fold dilution of DNA is required to remove the effects of PCR
241 inhibitory substances.

242

243

244

245 **Numbers of enterococci**

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

251 **Quantitative PCR results for environmental water samples**

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

260

261 **Discussion**

262 The high numbers of the HF183 markers in lake water samples indicate the presence of sewage
263 pollution. Stormwater pipes and slum-like establishments in the vicinity may have introduced sewage pollution
264 to the lake. Prior to water sampling, a sanitary survey was undertaken to identify possible sources of fecal
265 pollution other than human sources. A small number of stray dogs and few cattle were observed near the lake.
266 The HF183 markers have been isolated occasionally from the feces of companion animals such as dogs and cats
267 (Ahmed *et al.* 2008). As such, prior to field application, the specificity of the HF183 and CF128 markers were
268 evaluated by testing fecal samples from five host groups. It is acknowledged that the number of samples tested
269 for specificity assay was small. Nonetheless, the HF183 marker displayed high specificity. These markers were
270 only detected in two fecal samples from companion animals and the numbers were low compared to human feces
271 (see Table 1). The high specificity of the HF183 markers in distinguishing human from animal feces has been
272 reported from many locations throughout the world (Ahmed *et al.* 2008; Gourmelon *et al.* 2007; Seurinck *et al.*
273 2006). To rule out the possibility of dog-sourced fecal pollution, dog *Bacteroides* BacCan markers were used to

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

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

283 The numbers of the HF183 markers in human feces and sewage (collected from influent) range from
284 $8.4 \pm 0.1 \times 10^5$ 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,
285 respectively (Seurinck et al. 2006). Similar numbers of the HF183 have been detected in the feces collected from
286 humans in the present study (see Table 1). In this study, the number of the HF183 markers in lake water samples
287 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
288 human feces or sewage. Such data is cause for significant concern and indicate the presence of sewage pollution.
289 The number of enterococci was also high which indicated a high level of fecal pollution. The numbers of
290 enterococci were higher near the stormwater pipes, thus confirming that these pipes are contributing to the fecal
291 load in the lake. It has to be noted that the samples were collected after a rainfall event when the fecal indicator
292 numbers are expected to be high in environmental waters.

293 Nonetheless, the high levels of fecal indicator bacteria and the HF183 *Bacteroides* markers indicate
294 sewage pollution and health risks to those who use the lake water for non-potable uses such as bathing and
295 washing clothes. The overflow from the lake after heavy rain could also pose serious health risks to the residents
296 surrounding the lake due to the possibility of cross contamination with the reticulated supply water. The
297 significant increase in diarrhoeal cases after flood or storm events are common in Bangladesh (Hashizume et al.
298 2008).

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

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

306

307 **Acknowledgements**

308 This study was funded by Queensland Department of Environment and Resource Management (DERM).

309 This was a joint project between DERM, Queensland University of Technology (QUT) and the Independent

310 University of Bangladesh (IUB). We thank IUB for providing laboratory space. We also thank Ms. Warda

311 Ashraf for processing the DNA samples.

312

313

314

315

316

317

318

319

320

321

322

323

324

325

326

327

328

329

330 **References**

- 331 Ahmed, W., Neller, R., and Katouli, M. 2005. Host species-specific metabolic fingerprint database for
332 enterococci and *Escherichia coli* and its application to identify sources of fecal contamination in surface waters.
333 Appl. Environ. Microbiol. 71: 4461–4468.
334
- 335 Ahmed, W., Stewart, J., Powell, D., and Gardner, T. 2008. Evaluation of *Bacteroides* markers for the detection
336 of human faecal pollution. Lett. Appl. Microbiol. 46: 237-242.
337
- 338 Bernhard, A.E., and Field, K.G. 2000. A PCR assay to discriminate human and ruminant feces on the basis of
339 host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. Appl. Environ. Microbiol. 66: 4571–4574.
340
- 341 Bernhard, A.E., Goyard, T., Simonich, M., and Field, K.G. 2003. Application of a rapid method for identifying
342 fecal pollution sources in a multi-use estuary. Water Res. 37: 909–913.
343
- 344 Betancourt, W.Q., and Fujioka, R.S. 2006. *Bacteroides* spp. as reliable marker of sewage contamination in
345 Hawaii's environmental waters using molecular techniques. Wat. Sci. Technol. 54: 101-107.
346
- 347 Boon, N., Top, E.M., Verstraete, W., and Siciliano, S.D. 2003. Bioaugmentation as a tool to protect the structure
348 and function of an activated sludge microbial community against a 30-chloroaniline shock load. Appl. Environ.
349 Microbiol. 69: 1511-1520.
350
- 351 Field, K.G., and Samadpour, M. 2007. Fecal source tracking, the indicator paradigm, and managing water
352 quality. Water Res. 41: 3517-3538.
353
- 354 Fong, T.T., Griffin, D.W., and Lipp, E.K. 2005. Molecular assays for targeting human and bovine enteric viruses
355 in coastal waters and their application for library-independent source tracking. Appl. Environ. Microbiol. 71:
356 2070–2078.
357
- 358 Gawler, A.H., Beecher, J.E., Brandão, J., Carroll, N.M., Falcão, L., Gourmelon, M., Masterson, B., Nunes, B.,
359 Porter, J., Rincé, A., Rodrigues, R., Thorp, M., Walters, J.M., and Meijer, W.J. 2007. Validation of host-specific

360 *Bacteroidales* 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries
361 of the European Union. *Water Res.* 41: 3780-3784.
362
363 Gourmelon, M., Caprais, M.P., Ségura, R., Le Mennec, C., Lozach, S., Piriou, J.Y., and Rincé, A. 2007.
364 Evaluation of two library-independent microbial source tracking methods to identify sources of fecal
365 contamination in French estuaries. *Appl. Environ. Microbiol.* 73: 4857-4866.
366
367 Hashizume, M., Wagatsuma, Y., Faruque, A.S.G., Sack, D.A., Hayashi, T., Hunter, P.R., and Armstrong, B.
368 2008. Factors determining vulnerability diarrhoea during and after severe floods in Bangladesh. *J. Water. Health.*
369 6: 323-332.
370
371 Hörman, A., Rimhanen-Finne, R., Maunula, L., von Bonsdorff, C.-H., Torvela, N., Heikinheimo, A., and
372 Hanninen, M.-L. 2004. *Campylobacter* spp., *Giardia* spp., *Cryptosporidium* spp., Noroviruses, and indicator
373 organisms in surface water in south-western Finland, 2000–2001. *Appl. Environ. Microbiol.* 70: 87–95.
374
375 Khatib, L.A., Tsai, Y.L. and Olson, B.H. 2002. A biomarker for the identification of cattle fecal pollution in
376 water using the LTIIa toxin gene from enterotoxigenic *E. coli*. *Appl. Microbiol. Biotechnol.* 59: 97–104.
377
378 Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., and Wuertz, S. 2007. 16S r RNA-
379 based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: a
380 Bayesian approach. *Water Res.* 41: 3701-3715.
381
382 Layton, A., MacKay, L., Williams, D., Garrett, V., Gentry, R. and Sayler G. 2006. Development of *Bacteroides*
383 16S rRNA Gene TaqMan-Based Real-Time PCR Assays for Estimation of Total, Human, and Bovine Fecal
384 Pollution in Water. *Appl. Environ. Microbiol.* 72: 4214 - 4224.
385
386 McLellan, S.L. 2004. Genetic diversity of *Escherichia coli* isolated from urban rivers and beach water. *Appl.*
387 *Environ. Microbiol.* 70: 4658–4665.
388

389 Okabe, S., Okayama, N., Savichtcheva, O., and Ito, T. 2007. Identification and quantification of host-specific
390 *Bacteroides-Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in freshwaters. Appl.
391 Microbiol. Biotechnol. 74: 890–901.
392
393 Parveen, S., Murphree, R.L., Edminston, L., Kaspar, C.W., Portier, K.M., and Tamplin, M.L. 1997. Association
394 of multiple antibiotic resistance profiles with point and non-point sources of *Escherichia coli* in Apalachicola
395 Bay. Appl Environ Microbiol 63: 2607–2612.
396
397 Prüss-Üstun, A., Bos, R., Gore, F., and Bartam, J. 2008. Safer Water, Better Health: Costs, Benefits and
398 Sustainability of Interventions to Protect and Promote Health. Geneva: World Health Organization.
399
400 Reischer, G.H., Kasper, D.C., Steinborn, R., Farnleitner, A.H., and Mach, R.L. 2007. A quantitative real-time
401 PCR assay for the highly sensitive and specific detection of human faecal influence in spring water from a large
402 alpine catchment area. Lett. Appl. Microbiol. 44: 351-356.
403
404 Scott, T.M., Jenkins, T.M., Lukasik, J., and Rose, J.B. 2005. Potential use of a host associated molecular marker
405 in *Enterococcus faecium* as an index of human pollution. Environ. Sci. Technol. 39: 283-287.
406
407 Seurinck, S., Verdievel, M., Verstraete, W. and Siciliano, S.D. 2006. Identification of human fecal pollution
408 sources in a coastal area: a case study at Oostende (Belgium). J. Water Health 4: 167–175.
409
410 U.S. Environmental Protection Agency. 1986. Ambient water quality criteria for bacteria. EPA440/5-84-002.
411 Washington, DC 20460:Office of Water Regulations and Standards Division.
412
413 U.S. Environmental Protection Agency. 2000. Improved Enumeration Methods for the Recreational Water
414 Quality Indicators: Enterococci and *Escherichia coli*. EPA/821/R-97/004. Office of Science and Technology,
415 Washington, DC, 55p.
416
417 WHO. 2003. Guidelines for safe recreational-water environments, coastal and fresh-waters. Geneva, Switzerland.
418

419 WHO 2006. Mortality country fact sheet.
420 http://www.who.int/whosis/mort/profiles/mort_searo_bgd_bangladesh.pdf
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448

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 <i>Bacteroides</i> HF183			Ruminant wastewater-associated <i>Bacteroides</i> CF128		
		Positive	Numbers per 100 mg of feces	Negative	Positive	Numbers per 100 mg of feces	Negative
Humans	15	13	$1.2 \times 10^6 - 3.9 \times 10^8$	2	0	0	15
Cattle	12	0	0	12	9	$5.1 \times 10^6 - 6.3 \times 10^7$	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 \pm 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

458 **Table 3:** Effects of PCR inhibitors on the PCR detection of spiked cattle wastewater associated CF128
 459 *Bacteroides* markers in DNA isolated from surface water samples as opposed to distilled water
 460 samples
 461

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

462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475

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 <i>Bacteroides</i> HF183 marker per 100 ml	Presence/absence of cattle-specific <i>Bacteroides</i> 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

492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524

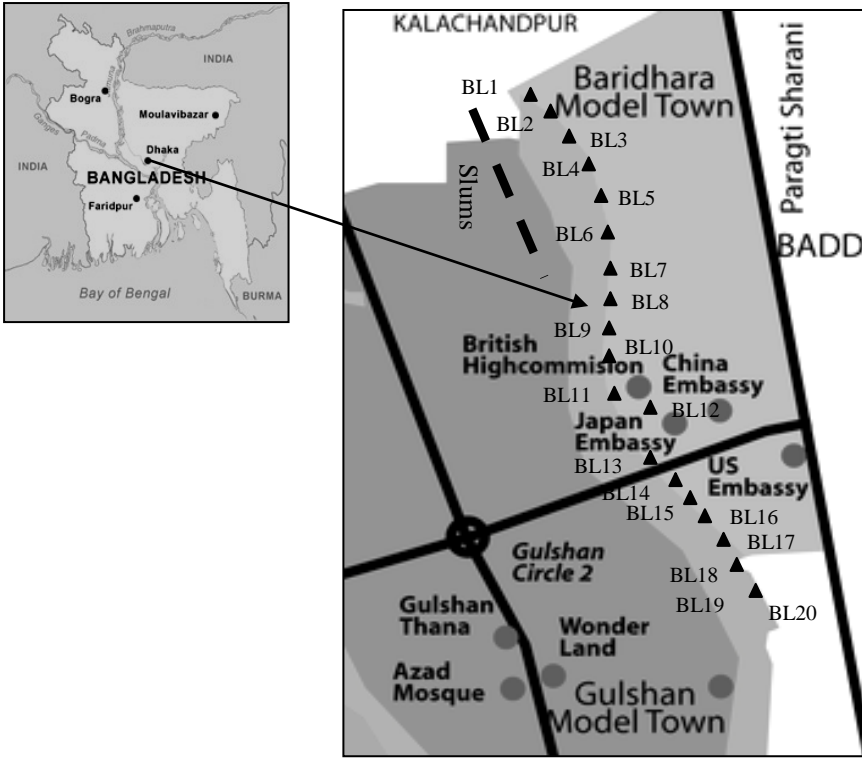


Figure 1: A map showing the sampling sites (▲) on the Baridhara Lake, Dhaka, Bangladesh