Quantitative detection of human- and canine-associated Bacteroides genetic markers from an urban coastal lagoon

Serhat A. Yasar, Toby J. T. Mills, Zehra I. Uluturk, Jason M. S. Ruszczyk, Rebecca J. LeBard and Brett A. Neilan,*

School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia
School of Environmental and Life Sciences, The University of Newcastle, Callaghan, Australia
Natural Environment Unit, Warringah Council, Sydney, Australia
*Correspondence author. E-mail: brett.neilan@newcastle.edu.au

ABSTRACT

The contamination of water catchments by nonpoint source faecal pollution is a major issue affecting the microbial quality of receiving waters and is associated with the occurrence of a range of enteric illnesses in humans. The potential sources of faecal pollution in surface waters are diverse, including urban sewage leaks, surface runoff and wildlife contamination originating from a range of hosts. The major contributing hosts require identification to allow targeted management of this public health concern. In this study, two high-performing Microbial Source Tracking (MST) assays, HF183/Bac242 and BacCan-UCDmodif, were used for their ability to detect host-specific Bacteroides 16Sr RNA markers for faecal pollution in a 12-month study on an urban coastal lagoon in Sydney, Australia. The lagoon was found to contain year-round high numbers of human and canine faecal markers, as well as faecal indicator bacteria counts, suggesting considerable human and animal faecal pollution. The high sensitivity and specificity of the HF183/Bac242 and BacCan-UCDmodif assays, together with the manageable levels of PCR inhibition and high level DNA extraction efficiency obtained from lagoon water samples make these markers candidates for inclusion in an MST “toolbox” for investigating host origins of faecal pollution in urban surface waters.

Key words: Bacteroides, faecal indicator bacteria, microbial source tracking, qPCR, urban surface water

HIGHLIGHTS

• This long-term study (twelve months with bi-monthly sampling) validated the use of HF183/BacCan242 and BacCan-UCDmodif to quantify host-associated faecal inputs in urban surface waters.
• Correlations between MST markers, traditional FIB measurements, and physiochemical parameters indicate that the latter may be used to predict the likelihood of human- and canine-associated faecal inputs at this site.

1. INTRODUCTION

The monitoring and management of recreational waters by environmental agencies is vital to the protection of users and the environment (World Health Organization 2003; NHMRC 2008). The presence of faecal contamination in recreational waters, as measured by the presence of faecal indicator bacteria (FIB), has been shown to predict the likelihood of pathogen presence and disease transmission (Cabelli et al. 1982; Wade et al. 2003). However, the detection of traditional FIBs does not allow for discrimination between host sources of the faecal pollution. Of particular importance is the need to identify the presence of faecal pollution from human sources such as sewerage, as these sources are more likely to contain human pathogens when compared with non-human sources. In addition, FIB from non-faecal sources can potentially result in false-positive results for faecal contamination (Davies et al. 1995; Desmarais et al. 2002; Yamahara et al. 2009). Identifying the potential source of the faecal pollution entering a recreational water body is important, allowing environmental managers to effectively manage the different human health risks associated with different faecal sources (Ashbolt et al. 2010).

In an attempt to address the differential risk associated with different host sources, a large number of microbial source tracking (MST) techniques have been developed over the past two decades (Harwood et al. 2005; Stoeckel & Harwood...
Although the pioneering methods were primarily library dependent (Hagedorn et al. 1999; Harwood et al. 2005), the discovery of host-specific genetic markers has led to the rise of library-independent MST methods (Bernhard & Field 2000b; Layton et al. 2006; Ahmed et al. 2016). These library-independent methods have the advantage of not requiring time-consuming and expensive reference libraries, and instead rely on a molecular assay to detect host-specific genetic markers directly from environmental samples. One of the first markers to be targeted by such an assay, and arguably the most discussed, is the 16S rRNA gene of human-associated Bacteroidales (Bernhard & Field 2000a). This gene has since been used as a target for a number of human-associated assays (Seurinck et al. 2005; Layton et al. 2006; Reischer et al. 2007).

The HF183 marker qPCR assay targeting the 16S rRNA gene of a human-associated Bacteroidales has emerged as a highly specific and sensitive assay for the detection of sewage, and remains one of the most widely tested techniques in MST field studies globally (Bernhard & Field 2000a; Seurinck et al. 2005; Gourmelon et al. 2010; Litton et al. 2010). Despite the assay’s widespread coverage in literature, relatively few case studies have been conducted using this marker in Australia (Ahmed et al. 2012). It must be noted however that there is inconsistent nomenclature in the literature regarding the HF183 marker assays, with the HF183 marker forward primer paired with numerous reverse primers and/or different real-time PCR chemistries. These differences in assay design must be considered when making comparisons. Additionally a frequently used HF183 marker reverser primer (Seurinck et al. 2005) was not named in the initial publication. As a result this primer has accumulated a variety of names for the same primer sequence, including Bac242R, HF264R, HF183R and SSHBacr (McLain et al. 2009; Stoeckel et al. 2009; Gourmelon et al. 2010; Staley et al. 2012; Kobayashi et al. 2013). This article uses Bac242R to describe the aforementioned modified HF183 marker assay reverse primer.

In addition to human-associated Bacteroides markers, a range of methods have also been described to target animal-associated Bacteroides. Bacteroides markers have been proposed for ruminants (Layton et al. 2006; Kildare et al. 2007), cattle (Reischer et al. 2007), porcine (Mieszkin et al. 2009; Devane et al. 2013) and birds (Lu et al. 2008; Ryu et al. 2012), amongst others. Kildare et al. (2007) described the first reported marker targeting the canine-associated Bacteroides 16S rRNA gene, BacCan-UCD. Canine-associated faecal pollution of beaches and lagoons can result in poor water quality due to the high levels of FIB found in dog faeces (Wright et al. 2009). MST strategies need to be capable of targeting the primary hosts most likely to contribute faecal pollution in order to facilitate accurate assessment and remediation of the human health risk associated with mixed faecal contamination of recreational waters.

In the present study we validate the use of two MST assays to quantitatively assess the extent of human- and canine-associated faecal pollution in an urbanised coastal lagoon in Sydney, Australia. A recent sanitary survey of the lagoon has found that sewerage overflows and dogs are the two main contributors of faecal pollution into the lagoon (Sydney Water 2012). The HF183/Bac242 and the BacCan-UCDmodif qPCR assays, targeting the 16S rRNA gene of human- and canine-associated Bacteroides, have been used to detect and quantify sewage and dog inputs respectively. In particular, HF183/Bac242 was found to be one of the most suitable markers for detecting sewage in waters when considering host specificity, sensitivity, and environmental marker concentrations and decay (Ahmed et al. 2010b, 2016; Dick Linda et al. 2010; Kobayashi et al. 2013; Nshimyimana et al. 2014; Riedel et al. 2014). These markers were assayed in parallel with intestinal enterococci and E. coli counts, as well as a range of physicochemical parameters to test for any correlation with traditional FIB and water quality measurements. The specificity and distribution of host-specific MST markers can vary geographically, and hence it is important to test markers against a range of local faecal sources. This study has tested specific HF183/Bac242 and BacCan-UCDmodif marker assays against faecal samples from a range of local sources to validate their suitability for the case study. The findings of this study will not only enable environmental managers to implement targeted remediation strategies to deal with specific faecal inputs, but will also validate MST as a viable tool for water quality monitoring by local councils.

2. METHODS

2.1. Study site

Manly Lagoon is a shallow coastal lagoon located at the northern end of Manly Beach in Sydney, Australia. The lagoon is approximately 2 km long (west-east) and 150 m at its widest. The surface area of the lagoon is approximately 0.1 km² and it is fed by a catchment approximately 18 km² in size, which is primarily residential, and contains several recreational reserves and parklands. Potential sources of faecal pollution include stormwater runoff and sewerage overflow from nearby residential areas, as well as surface runoff from parklands and reserves. The lagoon is considered to be one of the
most polluted water bodies in New South Wales, Australia (Integrated Catchment & Environmental Management Research Group 2004), and at the time of the study was closed for all recreational purposes due to concerns over water quality.

2.2. Sample collection
Water sampling was conducted fortnightly, excluding post heavy rain events, over a 12 month period \( (n = 378) \) between July 2010 and July 2011. Sampling was conducted during low tide and in dry weather. At each collection date, three separate water samples within a 30 m radius were collected, with the use of a motorboat, from each of the six sites \( (n = 18) \) along the lagoon (Figure 1). For sites 1–4, the top 1 m of the water column was collected using a 1 m long PVC pipe with a one-way valve (radius 14 mm, 600 ml capacity). The pipe was rinsed three times with lagoon water before each sampling. Due to shallow depths at sites 5 and 6, grab samples from approximately 30 cm below the water surface were taken. All water samples were collected in sterile 600 ml PET bottles. All sewage samples were collected from a nearby sewerage outflow. Approximately 1 L of sewage was collected every 10 minutes over a 1-hour period (total of 6 L). Samples were then pooled and 3 × 1 L composite sample were collected in sterile containers containing the dechlorinating agent sodium thiosulphate, to prevent bacteriocidal action during sample transport. Animal faecal samples were collected from nearby reserves and parklands. Dog \( (n = 8) \) and duck \( (n = 4) \) samples (various species) were collected using sterile utensils and specimen jars immediately after witnessing shedding by each animal. Human faecal samples \( (n = 18) \) were obtained from the Sydney Adventist Hospital, Wahroonga, Australia (ethics approval: UNSW Ref # AD11071). All samples were transported to the laboratory on ice and stored at 4 °C until processed. All samples were processed within 6 h of arrival at the laboratory.

2.3. Physical and chemical data
Physical and chemical data, along with weather information, was recorded at the time of each collection. Temperature, dissolved oxygen, pH, conductivity and turbidity were measured at two depths (top half, bottom half) for each site using Hydralab Quanta probe (Hydrolab Corporation, Austin, USA). Rainfall data for the previous three days prior to collection was collected from the Bureau of Meteorology website (www.bom.gov.au).

Figure 1 | Map of Manly Lagoon and surrounding catchment. Sampling sites 1–6 are represented by black squares (⬛), and north indicated by an arrow (↑). For the purpose of this study, sites 1–4 are considered downstream, and sites 5 and 6 are considered upstream. Image source: Google Maps © 2020.
2.4. Enumeration of faecal indicators

Enumeration of *E. coli* and enterococci was conducted using the membrane filtration method, as described by Australian and New Zealand Standard AS/NZS 4,276.1, 4,276.7 and 4,276.9. Briefly, water samples (*n* = 393) were filtered in duplicates through sterile 47 mm 0.45 μm membranes (Millipore, Australia) and placed onto m-FC agar (Oxoid, Australia) and m-Enterococcus agar (Becton Dickinson, Australia) for the enumeration of *Escherichia coli* and enterococci, respectively. Plates were incubated at 42 ± 0.5 °C for 24 h (*E. coli*) and 48 h (enterococci). The *E. coli* (blue) and enterococci (purple) colonies were counted immediately after incubation and converted to CFU/100 ml.

2.5. DNA extraction

DNA from water samples was extracted using the PowerWater DNA Isolation Kit (MoBio) modified with an additional lysis step. Briefly, 150–600 ml (recorded) of water was filtered through a sterile 47 mm 0.22 μM membrane (Millipore) and placed into the lysis tube containing 1 ml of the lysis buffer provided. Varying volumes of water were extracted to account for differences in turbidity and resulting filter capacity. The membrane was then incubated at 70 °C for 20 min, after which the manufacturer’s instructions were followed. All results were normalized to per 100 ml of water. DNA from faecal samples was extracted using a PowerSoil DNA Isolation Kit (MoBio), modified with an additional lysis step. Approximately 0.25 g of faecal DNA was placed into the lysis tube provided and heated to 70 °C for 20 min, after which manufacturer’s instructions were followed. DNA was eluted with 100 μl elution buffer and stored at −20 °C until processing. DNA concentration and purity was measured using a NanoDrop 2000 spectrophotometer (NanoDrop).

2.6. Construction of positive control plasmids

PCR products, amplified using the primer sets for each marker shown in Table 1, were purified using the PCR-M Clean-Up System (Viogene), according to manufacturer’s instructions, and cloned into a PCR 2.1-TOPO vector using TA cloning (Invitrogen). The plasmid was transformed into *E. coli* DH5α cells and recombinant bacteria were selected on LB agar plates containing ampicillin, IPTG (isopropyl beta-D-1-thiogalactopyranoside) and X-gal (5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside) as recommended by the manufacturer. Presumptive transformants (white colonies) were screened for the presence of the plasmid containing the control insert using colony PCR. One insert-positive clone for each marker was selected for overnight growth in liquid LB medium (containing ampicillin). Plasmid isolation was conducted using a plasmid mini kit (Qiagen) and inserts were sequenced to confirm specificity. Plasmid copy numbers were calculated by measuring the DNA concentration using a NanoDrop 2000 spectrophotometer (NanoDrop) and the molecular weight of the plasmid.

2.7. Quantification standards and limit of detection

Determination of amplification efficiencies and development of standard curves for the qPCR assays were conducted using plasmids containing the marker sequence specific for each assay. Quantification standard curves were constructed by making 10-fold dilutions (10⁻⁶–10⁻¹⁰) of each marker-containing plasmid. Limits of quantification for each assay (ALOQ) were calculated as the lowest concentration of marker-containing plasmid that was within the linear range of the standard curve during

---

Table 1 | Primers used in real-time and conventional DNA amplification assays described in this study

<table>
<thead>
<tr>
<th>Assay</th>
<th>Primer</th>
<th>Oligonucleotide Sequences (5’-3’)</th>
<th>Size of Product (bp)</th>
<th>Annealing Temp (°C)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Human – associated</td>
<td>HF183F</td>
<td>ATCATGAGTTCACATGTCCCG TACCCCGCCCTACTATAATG</td>
<td>81</td>
<td>53</td>
<td>Bernhard &amp; Field (2000b)</td>
</tr>
<tr>
<td><em>Bacteroides</em></td>
<td>Bac242R</td>
<td></td>
<td></td>
<td></td>
<td>Seurinck <em>et al.</em> (2005)</td>
</tr>
<tr>
<td>Canine – associated</td>
<td>BacCan-545F1</td>
<td>GGAGCGCAGACGGGTTTT CAATCGGAGTGCTCTGATATCTA</td>
<td>143</td>
<td>60</td>
<td>Kildare <em>et al.</em> (2007)</td>
</tr>
<tr>
<td><em>Bacteroides</em></td>
<td>BacUni-690r1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>General <em>Bacteroides</em></td>
<td>Bac32F</td>
<td>AACGCTAGCTACAGGCTT CAATCGGAGTGCTCTG</td>
<td>675</td>
<td>53</td>
<td>Bernhard &amp; Field (2000a)</td>
</tr>
<tr>
<td>Bac708R</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16S rDNA – General</td>
<td>P338F</td>
<td>ACTCTACGGGAGGCAGCAG ATTACCGGCGCTG</td>
<td>192</td>
<td>53</td>
<td>Ovreas <em>et al.</em> (1997)</td>
</tr>
<tr>
<td><em>Bacterial</em></td>
<td>P518R</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

Downloaded from http://iwaponline.com/wst/article-pdf/84/7/1732/948327/wst084071732.pdf by guest on 05 December 2021
qPCR amplification. The limit of detection (A_LOD) for the HF183/Bac242 assay was calculated by spiking sterile water with 10-fold dilutions of sewage (n = 3) to final concentrations of 10 ml sewage/100 ml water – 1 μl sewage/100 ml water. DNA extraction was conducted for each dilution using the PowerWater Isolation Kit (Mo Bio) as described previously. The lowest dilution at which HF183/Bac242 markers were detectable using the qPCR assay was calculated as the A_LOD for the HF183/Bac242 assay, expressed as μl of sewage/100 ml water. The A_LOD of the BacCan-UCDmodif assay was calculated by suspending 2 g (wet weight) of a pooled canine (n = 8) faecal sample in 100 ml sterile water. Triplet serial dilutions (n = 3) were then conducted to final concentrations of 1 g canine faeces/100 ml water – 1 ng canine faeces/100 ml water. DNA extraction was conducted for each dilution using a MoBio PowerSoil DNA Isolation Kit (MoBio) as described previously. The lowest dilution at which the BacCan-UCDmodif marker was detectable using the qPCR assay was calculated as the A_LOD for the BacCan-UCDmodif assay, expressed as μg of canine faeces/100 ml water.

2.8. Testing for PCR inhibitors

The possible presence of PCR inhibitors in DNA isolated from lagoon water samples was investigated in 3 representative samples (sites 1, 3, and 6) every second month (n = 18), adapted from (Ahmed et al. 2010a). DNA samples were diluted to 10^0, 10^-1 and 10^-2 and spiked with 10^3 copies of the HF183/Bac242 marker. The HF183/Bac242 marker was then quantified using the HF183/Bac242 assay for the spiked dilutions, un-spiked versions of the same dilutions, and sterile water spiked with 10^3 copies of the HF183/Bac242 marker. The number of HF183/Bac242 markers detected for each spiked sample was normalized against the background levels detected in the un-spiked samples, and compared with the number of markers detected in the spiked sterile water. The lowest spiked DNA dilution that did not exhibit a significant (P > 0.05) difference in copy number to the spiked sterile water was considered free of PCR inhibitors.

2.9. Determination of DNA extraction recovery efficiency

The recovery efficiency of the DNA extraction protocol would ideally be conducted by spiking sterilised lagoon water with known amounts of Bacteroides cells containing the HF183/Bac242 and BacCan-UCDmodif markers. The recovery efficiency was calculated by using a general bacterial real-time PCR assay, since a strain carrying these markers had not been isolated at the time of the study for use as a control. Bacteroides vulgatus ATCC 8482 cultures were grown overnight in Brain Heart Infusion (BHI) broth as previously described (Seurinck et al. 2003). Cells were harvested during stationary phase (17 h) and enumerated using a haemocytometer. Non-filtered, autoclaved lagoon water was spiked with B. vulgatus to a final concentration of 10^7 cells/100 ml, 10^6 cells/100 ml and 10^5 cells/100 ml. Spiking was conducted in triplicate for each dilution (n = 3). Spiked lagoon water (100 ml) was filtered through 0.22 μm membranes and DNA extraction was performed as described previously using MoBio PowerWater DNA Isolation Kit (MoBio). Bacteroides vulgatus was quantified using a previously described general bacterial real-time PCR assay (Boon et al. 2003), taking into consideration the 16S rRNA gene copy number of the strain, and primers listed in Table 1. A 10-fold dilution series of DNA was used as a standard for the general bacterial real-time assay. Recovery efficiency was calculated using the following formula, Recovery (%) = (concentration recovered)/(concentration spiked) × 100.

2.10. Calculation of host-specific marker concentrations in environmental samples

Amplification efficiencies (E) were calculated by estimating the slope of a linear standard curve constructed with at least 5 10-fold dilutions (10^-6–10^-10) of known amounts of marker-containing plasmids. The log concentration of marker copies in each reaction was plotted against the Ct value obtained, resulting in the generation of a standard curve. The amplification efficiency for each marker was estimated from the slope of the respective standard curve using the formula:

\[ E = 10^{-1/slope - 1}. \]  

(1)

A DNA reaction with a 100% (E = 1) efficiency will theoretically double the amount of amplicon each cycle, hence the initial concentration of template (C₀) can be calculated from the resulting concentration of product (Cₓ) after n cycles, using the formula:

\[ Cₓ = C₀(1 + E)^n \]  

(2)
2.11. Real time PCR

The HF183/Bac242 and BacCan-UCDmodif markers were quantified using qPCR utilizing SYBR green chemistry. All reactions were conducted in a Rotor-Gene Q real-time PCR machine (Qiagen), and analyzed using the Rotor-Gene Q series software v. 2.0.3 (Qiagen). A standard curve of marker-containing plasmid DNA (5 dilutions) was included in each run. For both assays, the 20 μl reaction consisted of 10 μl SensiFast™ SYBR No-ROX kit (Bioline, Australia), 250 nM of each primer and 2 μl of template DNA. Primers used in the assays are shown in Table 1. The BacCan-UCD assay was modified for use with the SYBR green real-time PCR chemistry, denoted BacCan-UCDmodif, specifically the TaqMan probe (BacUni-656p) and second reverse primer (BacUni690r2) were removed.

For the HF183/Bac242 assay, the cycling conditions 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 1 min at 60 °C. Cycle threshold values were evaluated at a set threshold of 0.06. Melt curve analysis was conducted and reactions which exhibited amplification and a disassociation temperature between 80 and 83 °C were considered positive for the HF183/Bac242 marker.

For the BacCan-UCDmodif assay, the cycling conditions consisted of 2 min at 50 °C, 10 min at 95 °C, and then 40 cycles of 15 s at 95 °C and 1 min at 60 °C. Melt curve analysis was conducted and reactions that exhibited amplification and a disassociation temperature between 84 °C and 86 °C were considered positive for the BacCan-UCDmodif marker.

2.12. Statistical analysis

All statistical analyses were performed using GraphPad Prism version 6.00 for Mac (GraphPad Software). One-way ANOVAs were used to determine differences between Ct values obtained in the inhibition experiments. Variability in the within-assay and between-assay values were analysed using Student’s t-tests. All regression calculations were conducted using parametric (Pearson’s) correlation analyses. All statistical analyses were calculated to within a 95% confidence interval (α = 0.05).

3. RESULTS AND DISCUSSION

3.1. Host specificity of the HF183/Bac242 and BacCan-UCDmodif assays

The host specificity of the HF183/Bac242 and BacCan-UCDmodif assays was tested against raw sewage, human, dog and duck faeces. The HF183/Bac242 marker tested positive against 25% of the dog faecal samples (n = 8) and 100% of the human faecal and sewage samples (n = 16 and n = 2, respectively). Duck faecal samples did not contain the HF183/Bac242 marker (n = 6). The BacCan-UCDmodif canine marker was detected in 100% of the dog samples, and did not cross amplify with any of the human or duck faecal samples. For the two dog faecal samples that tested positive to the HF183/Bac242 marker, the concentration was significantly lower than the average HF183/Bac242 marker concentration in human faeces (data not shown). All faecal samples tested positive for the general Bacteroides marker using conventional PCR.

3.2. Amplification efficiencies and quantitation

Standard curves were determined for the HF183/Bac242 and BacCan-UCDmodif assays. The amplification efficiencies for each assay were calculated from the slopes of the standard curves (Equation (1)), and were 97.8% for the HF183/Bac242 assay, and 96.1% for the BacCan-UCDmodif assay. A logarithmic non-linear regression curve was fitted to the data with a best fit (r²) value of 0.989 for the HF183/Bac242 marker and 0.99 for the BacCan-UCDmodif marker. The standard curves had a linear range of quantification (A_LOGQ) from 1 × 10¹⁻¹ to 10⁵ markers per 25 μl reaction for the HF183/Bac242 assay, and from 2.5 × 10⁻¹ to 2.5 × 10⁵ markers per 25 μl reaction for the BacCan-UCDmodif assay. All environmental samples that exhibited above-threshold amplification within the A_LOGQ, and exhibited the correct melt curve profile were deemed to contain the target gene. The corresponding melt curve disassociation temperatures were 82 ± 2 °C for the HF183/Bac242 marker, and 85 ± 1 °C for the BacCan-UCDmodif marker.

The sensitivity of the HF183/Bac242 and BacCan-UCDmodif qPCR assays to detect their respective markers in environmental samples was investigated by determining the amount of source material required for positive detection (A_LOD). The HF183/Bac242 and BacCan-UCDmodif qPCR assays were able to detect the presence of their respective markers in as little as 50 μl of sewage (HF183/Bac242) and 1 mg of dog faeces (BacCan-UCDmodif) added to 100 ml water and filtered onto a 0.22 μm membrane. At these source concentrations, no amplicons were observed using conventional PCR with the same primer pairs, suggesting that the qPCR assay is more sensitive at detecting the HF183/Bac242 marker than conventional PCR. Lower concentrations of source material (1–25 μl/100 ml sewage and 0.5–0.1 mg of dog faeces) did not yield detectable...
amplicons with either the qPCR assay or conventional PCR. This suggests that 50 μl sewage and 1 mg of dog faeces per 100 ml of water is the minimum amount of source material required for the described DNA extraction method to yield amplifiable DNA.

3.3. PCR inhibition and effect of matrix on DNA extraction efficiency

To test for the effect of PCR inhibitors on the quantification of the human- and canine-associated markers, an inhibition assay was conducted on representative samples. A one-way ANOVA revealed a significant degree of variation between the normalised copy numbers of spiked undiluted eDNA and the no-eDNA control ($F = 42.49, p < 0.01$). This suggests that inhibition was occurring in qPCR reactions using undiluted eDNA as a template. However, a significant variance was not observed in reactions using diluted ($10^{-1}$) eDNA ($F = 1.176, p = 0.321$). This suggests that PCR inhibition was not occurring in qPCR reactions using eDNA dilutions of 1:10 as a template. Subsequently, all environmental samples were diluted 1:10 for qPCR quantification to mitigate the effect of PCR inhibitor on marker quantitation.

The DNA extraction efficiency of the PowerWater DNA extraction kit (MoBio, USA), and the effect the water matrix from Manly Lagoon has on this efficiency was investigated. The water matrices investigated were filtered lagoon water, non-filtered lagoon water and a distilled water control. The recovery efficiencies for all matrices were cell-concentration dependent, with the lower cell concentrations yielding the highest efficiencies (Table 2). The recovery efficiencies ranged from $27.8\%$ (100 ml non-filtered water spiked with $10^7$ cells) to $90.07\%$ (100 ml filtered water spiked with $10^5$). For any one matrix, there was approximately a $\sim 40\%$ higher recovery efficiency for cell concentrations of $10^5$ cells/100 ml when compared with concentrations of $10^7$ cells/100 ml. However, the greatest variability in recovery efficiencies occurred at $10^5$/100 ml cell concentrations (95% CI of $\pm 12.0\%$–$17.2\%$). The water matrix that yielded the highest recovery efficiency for all cell concentrations was the non-filtered lagoon water. There was no significant difference in recovery efficiency between the filtered lagoon water and distilled water. Even though the lagoon water matrices were triple-sterilised (via autoclave) prior to spiking with *B. vulgatus*, DNA extractions were performed on un-spiked sterilised lagoon water to check for the presence of background DNA. No amplification was observed with universal 16S rDNA primers.

3.4. Enumeration of FIB

Box-and-whisker plots showing the $\log_{10}$ distributions of intestinal enterococci and thermotolerant *E. coli* for the six sites over 12 months is represented in Figure 2. The upper whisker of each box represents the 95th percentile cfu count for each site. All six sites tested positive for both enterococci and *E. coli* for all collection dates, with the upstream sites (sites 5 and 6) consistently yielding higher cfu counts when compared to the downstream sites (sites 1–4). The current NHMRC (2008) guidelines use 95th percentile enterococci counts, CFU per 100 ml, to assess water quality into four Microbial Assessment Categories (MAC). These MAC limits ($A \leq 40$ CFU/100 mL, $B = 41$ to 200 CFU/100 mL, $C = 201$ to 500 CFU/100 mL and $D > 501$ CFU/100 mL) are represented in Figure 2(a) as horizontal broken lines. Based on these limits, sites 2, 3 and 4 fall into a MAC of ‘B’, sites 5 and 6 fall into a MAC of ‘C’ and site 6 falls into a MAC of ‘D’.

3.5. Quantification of the HF183/Bac242 and BacCan-UCDmodif markers

The HF183/Bac242 and BacCan-UCDmodif markers were quantified using qPCR from 126 water samples collected from Manly Lagoon over a 12-month period. Frequencies of detection and marker quantification ranges are summarized for each marker in Table 3. The concentration of the HF183/Bac242 marker across all sites ranged from 3.1 to 5.77 $\log_{10}$ CN/100 ml of water, whereas the BacCan-UCDmodif marker concentrations ranged from 3.32 $\log_{10}$ to

<table>
<thead>
<tr>
<th>Spiked colonies/100 ml (n)</th>
<th>% Recovery efficiency ± 95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-filtered lagoon water</td>
</tr>
<tr>
<td>1 $\times$ 10⁵ (3)</td>
<td>90.7 ± 17.2</td>
</tr>
<tr>
<td>1 $\times$ 10⁶ (3)</td>
<td>55.2 ± 6.4</td>
</tr>
<tr>
<td>1 $\times$ 10⁷ (3)</td>
<td>41.9 ± 6.5</td>
</tr>
</tbody>
</table>
Figure 2 | Box and whisker plots showing the log_{10} distribution of intestinal enterococci (Panel a) and *E. coli* (Panel B) CFU counts for the six sites over 12 months. The lower and upper edges of the box are the 25th and 75th percentiles, respectively; the upper and lower whiskers are the 5th and 95th percentiles, respectively. The horizontal line within the box is the median. Outliers are plotted as individual markers (·). Calculations are based on 63 measurements for each site. The NHMRC (2008) recreational water 95th percentile enterococci limits for the Microbial Assessment Categories of B, C and D are represented by the horizontal lines (···), (-·-) and (-··-), respectively.
4.60 log$_{10}$ CN/100 ml. The HF183/Bac242 marker had a higher frequency of detection than the BacCan-UCDmodif marker, with 100% of samples testing positive, whereas only 68.3% of the samples tested positive for the BacCan-UCDmodif marker. Of the 40 samples that did not contain the BacCan-UCDmodif marker, 29 (72.5%) were in the downstream sites. Site 6 (upstream) was the site with the highest frequency of BacCan-UCDmodif marker detections, with 85.7% of the samples testing positive. On average, and across all sites and collection dates, the HF183/Bac242 marker concentrations were significantly higher than the BacCan-UCDmodif marker concentrations (paired $t$ test, $t = 13.84$, $p < 0.0001$).

A paired $t$ test revealed that there was no significant difference ($t = 0.345$, $p = 0.734$) in the mean HF183/Bac242 marker concentrations between the downstream sites (3.1–5.66 log$_{10}$ CN/100 ml) and the upstream sites (3.5–5.77 log$_{10}$ CN/100 ml). The vast majority of the samples contained the HF183/Bac242 marker at concentrations over 4 log$_{10}$ CN/100 ml (96.8%). Roughly half (46.8%) of the samples contained the HF183/Bac242 marker at concentrations above 5 log$_{10}$ CN/100 ml. No significant differences were observed between the downstream and upstream groups, nor between the four downstream or the two upstream sites. However, a paired $t$ test revealed that the upstream sites were found to contain higher levels of the BacCan-UCDmodif marker than the downstream sites, albeit with a weak $t$-ratio ($t = 2.64$, $p = 0.016$). The concentration range of BacCan-UCDmodif markers was between 3.32 and 4.40 log$_{10}$ CN/100 ml for the downstream sites, and 3.62 to 4.60 log$_{10}$ CN/100 ml for the upstream sites. Although a significant difference was observed in the BacCan-UCDmodif levels between the downstream and upstream sites, no such difference was observed between the four downstream sites, or between the two upstream sites.

A total of 86 (68.3%) and 27 (21.4%) water samples yielded BacCan-UCDmodif concentrations of over 3 log$_{10}$ and 4 log$_{10}$ CN/100 ml, respectively. Of the downstream samples, 65.5% yielded marker concentrations of over 3 log$_{10}$ CN/100 ml, with only 13.1% containing concentrations over 4 log$_{10}$/100 ml. The upstream sites had a larger proportion of samples that contained higher marker levels, with 38.1% of samples testing positive in the range of 4–5 log$_{10}$ CN/100 ml.

A one-way ANOVA revealed that the HF183/Bac242 marker concentrations at each site were not significantly different ($F = 0.318$, $p = 0.901$). However, the same statistical test applied to the BacCan-UCDmodif concentrations showed a significant difference between the site means ($F = 3.182$, $p = 0.011$). A Tukey’s multiple comparison test revealed this difference to only be between site 2 and site 6, with a mean difference of –0.3386 log$_{10}$/100 ml.

3.6. Correlation between FIB, markers and physical/chemical parameters

A parametric (Pearson’s) correlational analysis was performed to investigate possible covariance between the various physical, chemical, bacterial and molecular parameters measured in this study. The correlation coefficients (Pearson’s $r$) values are summarised in a significance matrix (Table 4). Of the physical/chemical parameters tested, turbidity was the only parameter exhibiting a positive correlation with all four bio-markers (FIB, HF183/Bac242 and BacCan-UCDmodif), albeit with weak coefficients ($r = 0.190–0.413$). Conductivity (salinity) and dissolved oxygen (DO) were both correlated (negatively) with the two FIB markers, but were only correlated with either one of the two molecular markers (DO with HF183/Bac242 and conductivity with BacCan-UCDmodif). Rainfall data was only correlated with HF183/Bac242 ($r = 0.407$) and *E. coli* ($r = 0.217$). Water temperature and pH did not correlate with any of the bio-markers. *E. coli* and *Enterococcus* counts were moderately correlated with each other ($r = 0.691$), and weakly with the BacCan-UCDmodif marker ($r = 0.329$ and 0.276, respectively). The HF183/Bac242 marker did not correlate with any of the other bio-markers.

### Table 3 | Detection frequencies and quantification ranges for the HF183/Bac242 and BacCan-UCDmodif markers

<table>
<thead>
<tr>
<th>Detection range</th>
<th>HF183/Bac242</th>
<th></th>
<th>BacCan-UCDmodif</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Downstream % (n)</td>
<td>Upstream % (n)</td>
<td>All sites % (n)</td>
<td>Downstream % (n)</td>
</tr>
<tr>
<td>Not detected</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>34.5 (29)</td>
</tr>
<tr>
<td>3 Log$<em>{10}$—4 Log$</em>{10}$</td>
<td>3.6 (3)</td>
<td>2.4 (1)</td>
<td>3.2 (4)</td>
<td>52.4 (44)</td>
</tr>
<tr>
<td>4 Log$<em>{10}$—5 Log$</em>{10}$</td>
<td>45.2 (38)</td>
<td>50.0 (21)</td>
<td>46.8 (59)</td>
<td>13.1 (11)</td>
</tr>
<tr>
<td>5 Log$<em>{10}$—6 Log$</em>{10}$</td>
<td>51.2 (43)</td>
<td>47.6 (20)</td>
<td>50 (63)</td>
<td>–</td>
</tr>
</tbody>
</table>

Downstream sites are sites 1–4, upstream sites are sites 5 and 6. For each assay $n = 126$. 

<table>
<thead>
<tr>
<th>Log$_{10}$ CN/100 ml</th>
<th>HF183/Bac242</th>
<th>BacCan-UCDmodif</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>4.60</td>
<td>5.77</td>
</tr>
<tr>
<td>4</td>
<td>5.00</td>
<td>5.66</td>
</tr>
<tr>
<td>5</td>
<td>5.32</td>
<td>5.57</td>
</tr>
<tr>
<td>6</td>
<td>5.66</td>
<td>5.77</td>
</tr>
</tbody>
</table>

...
3.7. Discussion

This study evaluated the use of two qPCR-based Bacteroides 16S rRNA gene markers to quantify host-associated faecal pollution in Manly Lagoon, Sydney. This small coastal estuarine lagoon is regarded as one of the most polluted recreational waters in the eastern seaboard of Australia (Integrated Catchment and Environmental Management Research Group 2004). Much of this pollution is attributed to faecal contamination from a variety of possible sources in the lagoon’s catchment, including sewerage overflows, storage drains and animals effluent mainly from dogs. The lagoon’s small size in relation to its 18 km² catchment means that the microbial quality of the lagoon is heavily impacted by rainfall events, sending microbial indicator counts to well above recreational guideline levels. It is primarily for this reason that the lagoon was closed for all recreational activities.

The assays, HF183/Bac242 (Bernhard & Field 2000b; Seurinck et al. 2005; Stoeckel et al. 2009) and BacCan-UCD (Kildare et al. 2007) assays targeting human- and canine-associated Bacteroides 16S rRNA markers, respectively, are considered two of the best-performing MST markers for inferring faecal contamination originating from their respective hosts (Boehm et al. 2013). In this work, both assays were assessed using SYBR green chemistry with the modified BacCan-UCD assay denoted BacCan-UCD modif.

The specificity and sensitivity of the assays were found to be high when tested against faecal samples from local hosts; hence, geographical variance was not found to affect assay performance in this respect. The HF183/Bac242 marker was found to be present in all sewage and human faecal sources tested, as well as in two canine faecal samples. The potential cross-reactivity of the HF183/Bac242 assay with dog faeces has been previously documented (Malla et al. 2018). The limits of detection and quantification were found to be adequate for the detection of small amounts of faecal contamination in water samples and were consistent with previously reported values.

The two MST assays were evaluated in a 12-month ‘wet-weather’ case study of Manly Lagoon with the two hosts identified as the most likely sources of pollution (Sydney Water 2012). In addition to the MST markers, the traditional microbial indicators, E. coli and intestinal enterococci, and a range of physico-chemical parameters, were also measured. The rationale behind this was twofold; firstly, to evaluate a potential correlation between all quality measures, and secondly, to provide an additional tier of confidence to the detection of faecal contamination. This kind of multi-tiered approach is considered the current ‘best-practice’ method for MST work, as it enables more robust conclusions to be drawn.

The present study found that intestinal enterococci counts for Manly Lagoon were consistently above the 2008 NHMRC guidelines for managing risks in recreational water for both primary- and secondary-contact recreational waters, particularly at the two upstream sites. This finding was consistent with previous BeachWatch data that found the lagoon to have a low microbial quality rating, particularly after rainfall.

The likelihood of faecal contamination, as suggested by the high enterococci counts, was supported by the findings of the MST study with ubiquitous detection of the HF183/Bac242 marker at all sites along the lagoon, and all sampling times.

Both the HF183/Bac242 and the BacCan-UCDmodif markers were consistently detected at all sites, indicating that faecal contamination from human and canine sources was widespread in the catchment. The copy numbers of these markers were higher for HF183/Bac242, potentially suggesting that sewage was the main contributor of faecal pollution. Consistent with
this, a recent study of three Sydney estuarine waters found sewage-associated MSTs, including the HF183 marker (Green et al. 2014), to be present in concentrations 3–5 times higher following rainfall compared with dry samples (Ahmed et al. 2020).

These results inform the potential suitability of Manly Lagoon for secondary-contact recreational activities, such as rowing and boating during dry or near-dry periods, which has not been adequately addressed previously. However, these copy number values need to be interpreted with caution, as higher copy-numbers have not yet been directly linked to higher faecal loads, pathogen titers, or public health risks.

A further level of confidence was added to the microbial data by the correlational analysis of the physico-chemical data. Turbidity measurements were positively correlated with all microbial indicators (FIB and MST), whereas conductivity and DO were both negatively correlated with FIB counts, but were only negatively correlated with either the HF183/Bac242 (DO) or BacCan-UCDmodif (conductivity) markers. The findings of the present study support the use of turbidity measurements to potentially act as a rapid and cost-effective early-warning indicator for low microbial quality, especially when combined with traditional FIBs and other physicochemical measurements, such as increased rainfall, decreased DO and decreased conductivity.

These parameters could be easily and continuously monitored via remote data loggers, the results of which can be analysed via multivariable statistical models to trigger further microbial investigation, such as MST analysis, if needed. Advances in the field of in-situ molecular detection technology, such as the Environmental Sample Processor (ESP) (Preston et al. 2011), could mean that further microbial analysis can be potentially automated in the not-so-distant future, making this an exciting time for both water quality researchers and policy makers alike.

4. CONCLUSION

1. This long-term study (twelve months with bi-monthly sampling) validated the use of HF183/Bac242 and BacCan-UCDmodif to quantify host-associated faecal inputs in urban surface waters.
2. Correlations between MST markers, traditional FIB measurements, and physio-chemical parameters indicate that the latter may be used to predict the likelihood of human- and canine-associated faecal inputs at this site.
3. High sensitivity and specificity, along with manageable PCR inhibition and high-level DNA extraction efficiency, support HF183/Bac242 and BacCan-UCDmodif use in urban surface waters.

DATA AVAILABILITY STATEMENT

All relevant data are included in the paper or its Supplementary Information.

REFERENCES

Bernhard, A. E. & Field, K. G. 2000b A PCR assay to discriminate human and ruminant feces on the basis of host differences in Bacteroides-Prevotella genes encoding 16S rRNA. Applied and Environmental Microbiology 66 (10), 4571–4574.


gull-associated conventional and quantitative PCR assays in artificial and environmental waters. Journal of Environmental Management 136, 112–120.