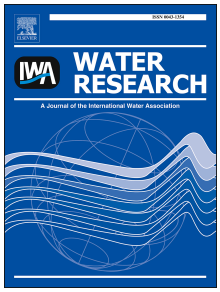
Accepted Manuscript



Bacteroidales markers for microbial source tracking in Southeast Asia

Jean Pierre Nshimyimana, Mercedes C. Cruz, R. Janelle Thompson, Stefan Wuertz

PII: S0043-1354(17)30287-7

DOI: 10.1016/j.watres.2017.04.027

Reference: WR 12824

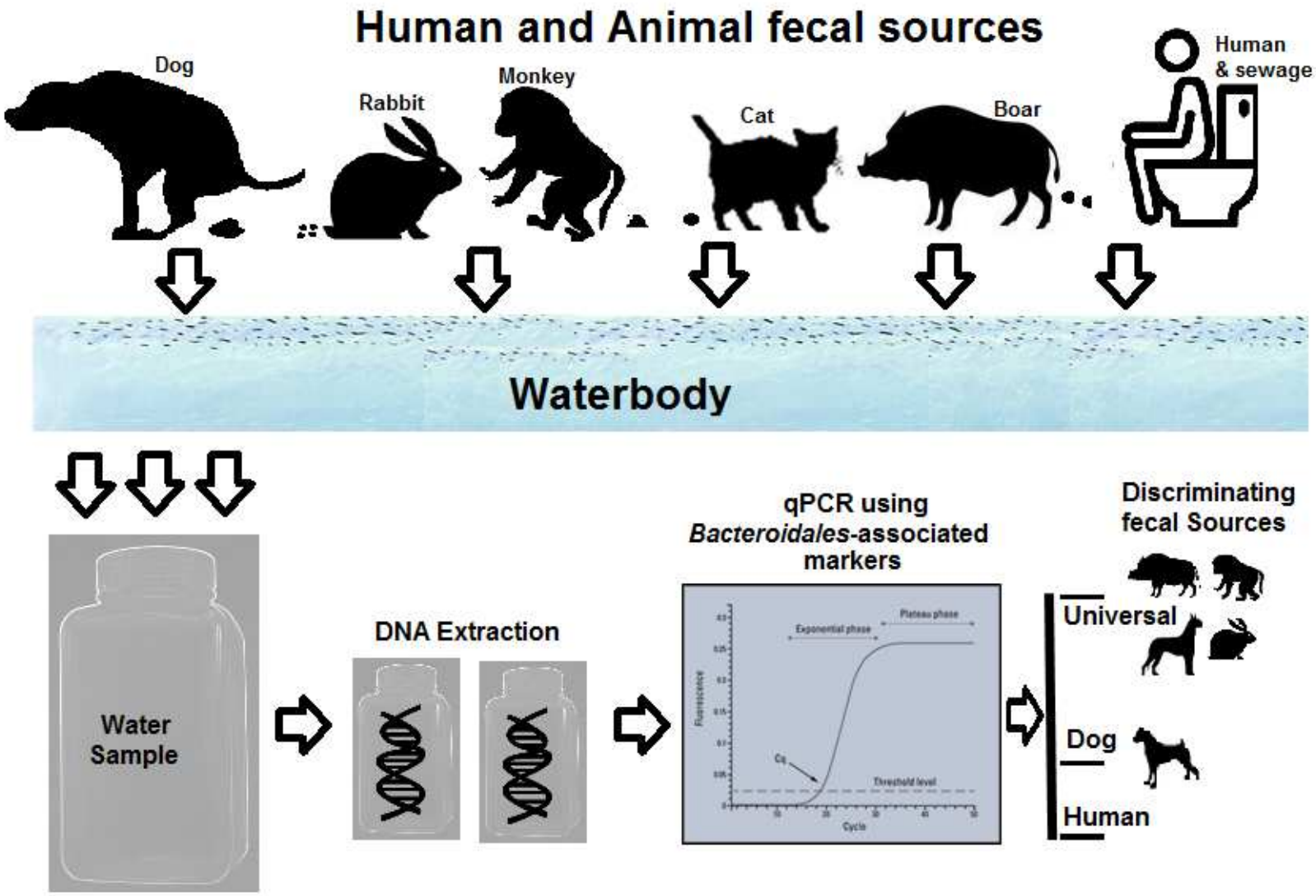
To appear in: *Water Research*

Received Date: 6 October 2016

Revised Date: 5 April 2017

Accepted Date: 8 April 2017

Please cite this article as: Nshimyimana, J.P., Cruz, M.C., Thompson, R.J., Wuertz, S., *Bacteroidales* markers for microbial source tracking in Southeast Asia, *Water Research* (2017), doi: 10.1016/j.watres.2017.04.027.


Human and Animal fecal sources
Dog
Rabbit
Monkey
Cat
Boar
Human
& sewage
Waterbody
Water
Sample
DNA Extraction
qPCR using
Bacteroidales-associated
markers
Exponential phase
Plateau phase
Cq
Threshold level
Cycle
Fluorescence
Discriminating
fecal Sources
Universal
Dog
Human

Bacteroidales Markers for Microbial Source Tracking in Southeast Asia

Nshimyimana, Jean Pierre[1, 2, 3, 4], *Cruz, Mercedes C.*[2], *Thompson, R. Janelle*[3, 4], *and Wuertz, Stefan*[1, 2, *]

[1] School of Civil and Environmental Engineering, Nanyang Technological University (NTU), 50 Nanyang Avenue, Singapore 639798, Singapore

[2] Singapore Center for Environmental Life Sciences Engineering, NTU, 60 Nanyang Dr., Singapore 637551, Singapore

[3] Department of Civil and Environmental Engineering, Massachusetts Institute of Technology (MIT), 77 Massachusetts Avenue, Cambridge, MA 02139, USA

[4] Centre for Environmental Sensing and Modeling, Singapore-MIT Alliance for Research and Technology, 1 Create Way, Singapore 138602, Singapore

*Correspondence to swuertz@ntu.edu.sg

Abstract

The island city country of Singapore served as a model to validate the use of host-associated *Bacteroidales* 16S rRNA gene marker assays for identifying sources of fecal pollution in the urban tropical environment of Southeast Asia. A total of 295 samples were collected from sewage, humans, domesticated animals (cats, dogs, rabbits and chicken), and wild animals (birds, monkeys and wild boars). Samples were analyzed by real time PCR using five human-associated assays (HF183-SYBR Green, HF183, BacHum, BacH and *B. thetaiotaomicron* α-1-6, mannanase (B. theta), one canine-associated assay (BacCan), and a total *Bacteroidales* assay (BacUni). The best performing human-associated assay was B. theta with a diagnostic sensitivity of 69% and 100% in human stool and sewage, respectively, and a specificity of 98%. BacHum achieved the second highest sensitivity and specificity for human stool at 65% and 91%, respectively. The canine-associated *Bacteroidales* assay (BacCan) had a sensitivity and specificity above 80% and was validated for tracking fecal pollution from dogs. BacUni demonstrated a sensitivity and specificity of 100% for mammals, thus BacUni was confirmed for total *Bacteroidales* detection in the region. We showed for the first time that rabbit fecal samples cross-react with human-associated assays (HF183-SYBR Green, HF183, BacHum and BacH) and with BacCan. Our findings regarding the best performing human-associated assays differ from those reported in Bangladesh and India, which are geographically close to Southeast Asia, and where HF183 and BacHum were the preferred assays, respectively.

Keywords: Fecal pollution; microbial source tracking; *Bacteroidales*; Quantitative PCR; Singapore; Southeast Asia

1. Introduction

Methods based on specific genetic targets, such as host-associated *Bacteroidales* 16S rRNA gene markers, have been developed to discriminate human fecal from other fecal animal sources (Bernhard and Field 2000a, b, Santo Domingo et al. 2007, Feachem 1975). Such markers have successfully differentiated human fecal pollution sources from various animal sources including ruminants, dogs and birds (e.g. geese) (Bernhard and Field 2000b, Kildare et al. 2007, Green et al. 2012) in a wide range of geographic regions that include Canada (Wilkes et al. 2013), the US Gulf of Mexico (Harwood et al. 2009), Japan (Okabe et al. 2007), Kenya (Jenkins et al. 2009), Tanzania (Pickering et al. 2011, Pickering et al. 2012, Mattioli et al. 2013), and New Zealand (Green, 2012).

The most widely used human-associated *Bacteroidales* marker, HF183, was specific for human stool and sewage detection in some locations but not others, indicating variability of human-associated *Bacteroidales* markers across different regions. HF183 was identified as a highly performing marker in California, USA (Layton et al. 2013), four Atlantic Rim Countries (France, Ireland, Portugal and the United Kingdom) (Gawler et al. 2007), Bangladesh (Ahmed et al. 2010) and Southeast Queensland in Australia (Ahmed et al. 2008a, Ahmed et al. 2008b). BacHum was validated for human fecal source tracking in California (Kildare et al. 2007), Kenya (Jenkins et al. 2009) and Odisha (India) (Odagiri et al. 2015), while a separate study identified *Bacteroides thetaiotaomicron-specific α-1-6* mannanase (B. theta) (Yampara-Iquise et al. 2008) as the most efficient genetic marker for human stool and sewage source tracking in Darwin (northern Australia) (Neave et al. 2014). Additional regional differences in efficacy of host-associated *Bacteroidales* for human microbial source tracking were reported for the markers

BacH and BacHum, after testing them against 280 human and animal fecal samples from sixteen countries across six continents (Reischer et al. 2013).

Urban waterways in Southeast Asia are subject to fecal pollution from a variety of sources. Singapore, an urban island in Southeast Asia, is a regional exception where advanced rainwater collection and sewage management infrastructures have limited fecal pollution of waterways and water bodies. Although several studies have noted high concentrations of FIB associated with septic systems or sewage, high environmental baseline levels of FIB are common under tropical conditions (Nshimyimana et al. 2014, Ekklesia et al. 2015a, Ekklesia et al. 2015b). Southeast Asian cities in Vietnam, Indonesia, Cambodia and Thailand listed high *E. coli* concentrations in the dry season, averaging 4.3 log cfu/100 ml, with a demonstrated need to identify sources of fecal contamination (Widmer et al. 2013). However, discrimination and tracking of human fecal non-point sources of pollution in Southeast Asia has seen limited application and potential *Bacteroidales* assays have not yet been validated.

Seven host-associated *Bacteroidales* assays (one universal, five human, and one canine assay) were selected for this study to validate performance in the tropical urban environment. The five human-associated *Bacteroidales* assays have been previously used and validated for temperate and tropical environments. Six of the tested assays used a dual-labeled hydrolysis probe (Taqman) and one, HF183-SYBR Green, was based on the DNA binding dye SYBR Green. HF183-SYBR Green (Ahmed et al. 2010) and BacHum (Odagiri et al. 2015) are the only markers that have been validated under tropical conditions (Bangladesh and Odisha, India, respectively). *Bacteroides thetaiotaomicron*-specific *α-1-6,* mannanase (B. theta) was suggested as an alternative human fecal pathogen indicator for Singapore (Liang et al. 2015) and reported

92 to have limited cross-reaction with animal feces compared to other human-associated
93 *Bacteroidales* markers such as HF183 (Yampara-Iquise et al. 2008, Aslan and Rose 2013).

94 The two human markers HF183 TaqMan® and BacH and the canine marker (BacCan)
95 were selected based on a multi-laboratory (27 laboratories) round-robin comparison that
96 evaluated the performance of 41 MST assays using 64 blind fecal samples in California, USA
97 (Boehm et al. 2013). BacCan was one of two recommended canine-associated assays in that
98 study (Schriewer et al. 2013). The *Bacteroidales* universal assay (BacUni) was selected because
99 it has been used in tropical environments of Kenya and India, and was identified as abundant in
100 human sources and many types of animal fecal sources in the USA (Kildare et al. 2007, Jenkins
101 et al. 2009, Odagiri et al. 2015, Silkie and Nelson 2009).

102 The objectives of this study were (i) to find the best assay(s) for detecting and monitoring
103 human and sewage fecal pollution and (ii) to test the efficacy of a high performing canine-
104 associated *Bacteroidales* assay (BacCan) in Singapore. We evaluated the performance of seven
105 host-associated *Bacteroidales* marker assays against 295 fecal samples from humans, sewage,
106 and eight relevant domesticated and wild animals. Both sensitivity and specify were compared
107 to select the best marker for microbial source tracking in the tropical urban environment of
108 Singapore and Southeast Asia, a region hosting one of the fastest growing human populations in
109 the world.

110 **2. Materials and Methods**

111 **2.1 Study area**

112 The study was carried out in the tropical urban environment of Singapore, an island
113 located in Southeast Asia on the tip of the Malaysian peninsula. This island city country is

known for its urbanized environment dominated by high density residential housing, with over 300 parks and four natural reserves that enhance the Singapore greenery (NParks 2016). The Natural reserves and parks are integrated with water bodies where few of them are used for secondary recreational activities (e.g., Jurong Lake, Kallang River, MacRitchie and Marina Bay Reservoirs). The animals commonly found in these natural reserves include wild boars (*Sus scrofa*), monkeys (*Macaca fascicularis*) and other mammals, birds and reptiles. Moreover, Singapore residents also keep domesticated animals as pets or livestock including dogs, cats, rabbits, and chickens. Despite the abundance of wild and domesticated animal species, the impact from human fecal pollution is of highest concern and thus water quality monitoring programs need to be able to detect when human fecal pollution enters waterways or water bodies.

2.2 Fecal sample collection

A total of 295 fresh fecal samples were collected from domesticated animals (i.e., dogs (n=50), cats (n=50), rabbits (n=40) and chickens (n=16)), wild animals (i.e., boars (n=14)), monkeys (n=10) and birds (n=50)), human volunteers (n=35), and raw sewage from an urban tropical wastewater treatment plant (n=30). Ten individual fecal samples for each type of animal and, where applicable, an additional one to four pooled samples of ten samples each were processed for further steps of the study (Supplemental Table 1). Pooling samples was done by mixing 2 mg of each animal fecal sample or 2 ml from each sewage sample in a 50-ml Falcon® tube (Corning Inc., Tewksbury, MA USA) and diluting the fecal samples with 15 ml of PBS before removing 5 ml of each pooled sample for DNA extraction. Animals were selected in consultation with the Singapore National Parks Board (NParks) by the following criteria: i) animals that live and walk around water bodies and could potentially contribute to non-point sources of fecal microbial contamination, ii) population size on the island, iii) frequency and

137 quantity of droppings, and iv) movement patterns such as grazing in groups of more than three
138 animals. To ensure that all animal fecal samples collected were fresh or had recently been
139 dropped, dry droppings or those found after a rainfall event were not collected. In addition,
140 sightings of animals in the vicinity of the sampling location and fresh animal footprints were
141 other indicators used to identify fresh droppings during wild animal tracking. NParks staff
142 provided animal behavior training in order to ensure that animals were not stressed by our
143 sampling events.

144 Domesticated animals were sampled approximately two hours after every morning
145 feeding at the rescue house of the Society for the Prevention of Cruelty to Animals in Singapore.
146 The rescue house was selected because it provided an opportunity to sample multiple types and
147 individuals of healthy domesticated animals.

148 Each fecal sample was collected using a 15-ml Falcon® tube (Corning Inc., Tewksbury,
149 MA USA) and a stool nucleic acid collection and preservation bottle (Norgen Biotek Corp.,
150 Thorold, Ontario, Canada) for humans, monkeys and wild boars. Collected fecal samples were
151 frozen on site, stored and carried to the laboratory using a high capacity liquid nitrogen freezer
152 (Taylor-Wharton Cryogenics LLC, Minnetonka, MN USA). Sewage samples were collected by
153 the grab method and one liter was filtered onto Sterivex 0.22-μm pore size filters (Millipore,
154 Billerica, MA, USA). Filter pieces were removed from the cartridge and aseptically sliced into
155 eight to ten fragments using a flame-sterilized scalpel prior to analysis. Once at the lab, fecal and
156 sewage samples were stored at -80^{0}C until processing.

157 Human stool samples were collected using an anonymous sampling strategy carefully
158 designed to protect the identity of the human volunteers. Every individual volunteer was given a
159 human stool sampling kit customized for this study. The kit included a Commode Specimen

Collection System (STL Medical Supply, Inc., Saint Louis, Missouri, USA), a pair of gloves, a stool nucleic acid collection and preservation bottle (Norgen Biotek Corp., Thorold, Ontario, Canada) and a biological waste bag used for disposing all human stool sampling related wastes. Human stool samples were then labeled with a random number and submitted to the Singapore Center for Environmental Life Science Engineering (SCELSE) collection point from where they were immediately transported and stored in a -80 degree freezer until sample processing.

The Nanyang Technological University (NTU) Institutional Review Board (NTU-IRB-2015-03-008) approved sample collection and other proceedings of this study. Human volunteers provided the stool samples willingly and without monetary compensation. All volunteers signed a consent form giving permission to use their stool samples in this study. Animal fecal samples were collected after receiving animal behavior training and a permit from NParks (NP/RP14-131 of 9 January 2015). All field sampling of animal feces was supervised by an NParks field technician assigned to the area where animals were being tracked.

2.3 Genomic DNA extraction and Inhibition removal

The PowerFecal® DNA isolation kit (Mo Bio, Carlsbad, CA USA) was used to extract the genomic DNA of all collected fecal and sewage samples following the manufacturer recommendations. Briefly, 0.25 grams of each fecal sample or the entire fragmented Sterivex® filter were added to bead tubes with lysis buffer. Samples were then vigorously homogenized using a vortex for cell lysis. Removal of inhibitors was performed using Mo Bio Inhibitor Removal Technology® (IRT) (Mo Bio, Carlsbad, CA USA). Genomic DNA was eluted in 100 µl of the elution buffer provided and samples were aliquoted and stored at -20^{0}C degrees until further analysis. DNA concentration and quality were measured and assessed using Qubit™ 2.0 Fuorometer (Invitrogen™, Waltham, MA, USA) and NanoDrop® ND-1000 (Waltham, MA,

183 USA), respectively. Three procedural controls of water in bottles were filtered using Sterivex
184 (0.22 μm pore size) and DNA extracted using the Power DNA extraction kit tested negative by
185 qPCR for the human markers used in this study.

186 Inhibitor removal technology of the PowerFecal® DNA isolation kit was complemented
187 by the OneStep™ PCR inhibitor removal kit (Zymo Research Corps, Irvine, CA, USA). Briefly,
188 Zymo-spin™ IV-HRC columns were prepared by snapping off the base and removing the cap,
189 then spun in the centrifuge. Extracted DNA was loaded onto Zymo columns, centrifuged, and
190 after inhibition removal, genomic DNA concentration was measured using NanoDrop® ND-1000
191 (Waltham, MA, USA).

192 **2.4 qPCR assays**

193 All probe-based *Bacteroidales*-associated 16S rRNA gene marker assay mixes included
194 20 μL of qPCR mixture containing 10 μL of TaqMan® Fast Advanced Master Mix 2.0
195 (ThermoFisher Scientific, Waltham, MA USA), primers and probes of which added quantities
196 were determined based on their final concentration in the mix (Table 1), 2 μL of genomic DNA,
197 and molecular quality water to reach the final volume. The 20-μL reaction mixture for the SYBR
198 Green-based human-associated HF183 assay (HF183-SYBR Green) contained 10 μl of KAPA
199 SYBR® FAST 2X Master Mix (KAPABIOSYSTEMS, Woburn, MA, USA), 10 μM of each
200 primer and 2 μl of DNA template (Seurinck et al. 2005). The total DNA mass in each sample
201 was 1 ng total DNA/reaction as concentrations were normalized to the total DNA mass per
202 reaction (Supplemental Information). All quantification reactions were performed using the
203 LightCycler® 480 Real-Time PCR system and software v. 1.5.0 (Roche Applied Sciences,
204 Indianapolis, IN, USA) for calculating crossing point (Cp) values. Melting temperature (Tm)
205 analysis for the HF183-SYBR Green assay was performed as previously described

206 (Nshimyimana et al. 2014, Seurinck et al. 2005). The qPCR protocol for probe-based assays
207 followed TaqMan® Fast Advanced Master manufacturer's instructions. Briefly, reactions were
208 subjected to a pre-incubation step at 50°C for 2 min and 95°C for 20 s, followed by 40 cycles of
209 95°C for 3 s, 60°C for 30 s, and 40°C for 30 s. The qPCR assay for HF183-SYBR Green
210 followed KAPA SYBR® FAST 2X Master manufacturer's instructions. Briefly, reactions were
211 subjected to a pre-incubation step of 95°C for 3 min, followed by 45 cycles of 95°C for 10 s,
212 53°C for 20 s and 72°C for 1 s. Each sample, whether using probe-based or SYBR Green
213 chemistry, was analyzed in duplicate and Cp values were examined after amplification to verify
214 consistency (i.e. the Cp difference between duplicates was ≤ 1 Cp). The overall Cp difference
215 between duplicates for samples amplified in this study was averaged at 0.40 ± 0.06.

216 To ensure that the qPCR reagents were not exposed to aerosols or other laboratory
217 contamination we carried out all reagent preparation in a bench top hood fitted with a UV-light
218 for sterilization before and after use (Airclean® Systems, Creedmoor, NC, USA). Each 96-well
219 plate contained two wells without DNA template as negative controls. All negative controls were
220 identified as NTD (no-target detected) after qPCR runs, confirming that qPCR plates and
221 reagents were not contaminated during preparation (Supplemental Table 4).

222 Reproducibility on each plate run during this study were verified by two standards from a
223 serial dilution of the positive controls (plasmid DNA) of each marker (10^3 and 10^2
224 copies/reaction) of the seven analyzed in this study. Overall, the average coefficient of variability
225 (%CV) was 1.35 ± 0.72% for the 10^3 copies/reaction standard and 1.58 ± 0.54% for the 10^2
226 copies/reaction standard (Supplemental Table 3). Minimum and uniform pipetting errors across
227 this study were achieved by the use of electronic pipettes (Gilson Inc., Middleton, WI, USA)
228 calibrated once.

To ensure that inhibition removal using kit-based technologies was effective, we used a standard spiking approach by adding the target marker at 10^3 copies/reaction when amplifying a specific marker in template DNA from samples, to verify that the impact of inhibitors had been sufficiently reduced. If the qPCR with both DNA template and target marker spike did not yield the predicted concentration of target amplicon by qPCR (i.e. ≥65% target concentration to allow tolerance for intra-run variation) the sample DNA was reanalyzed after a 1:10 dilution. Fifty-eight qPCRs across the seven *Bacteroidales* assays indicated inhibition and were diluted tenfold and reanalyzed (Nshimyimana et al. 2014).

3.3 Adjustment of qPCR conditions and verification of instrument compatibility

Bacteroidales markers used in this study were previously designed, optimized, and applied in other geographic regions (Table 1). Annealing temperature and reagent concentrations (primers and probes) were according to original assays. The rest of thermocycling conditions were as recommended by master mix manufacturers (Applied Biosystems). Fast Advanced Master Mix (Applied Biosystems, Foster, CA, USA) was manufactured by the same manufacturer as Environmental Master Mix (AB 2011); it was selected for this study because it was the only TaqMan master mix optimized and compatible with both Roche (LightCycler 480) and Applied Biosystems (StepOnePlus) amplification systems. This compatibility minimized variation that might have been associated with the use of different instruments and master mixes. The HF183 marker was tested using three different assays (HF183-SYBR Green (Seurinck et al. 2005), HF183-Haugland (Haugland et al. 2010) and HF183-Green (Bernhard and Field 2000a, Green et al. 2014) and two different master mixes (SYBR Green and TaqMan) using the LightCycler 480. The comparison of measured and expected concentrations showed a strong linear relationship, confirming that the use of LightCycler 480 and different assays did not

introduce any variability in quantification (slope of 1 and $R^2 > 0.98$ (Supplemental Figure 1)). Apart from the HF183 tests, all analyzed plates were reproducible with a low coefficient of variability (Supplemental Tables 3 and 4). It should be noted that the HF183-Green procedure was modified from the recommended protocol (Green et al. 2014) to work with the Light Cycler 480 instrument which required use of a different master mix, thermal cycling parameters, and cycle threshold definitions. It is possible that these modifications from the originally reported protocol could have changed method sensitivity, specificity, or accuracy. As a result, study findings reported here represent performance of modified procedures only.

2.5 Target limit of detection

Standard curves for all seven assays tested (Table 1) were established using ten-fold serial dilutions (10^1 to 10^8 gene copies per reaction/μl) of each assay's DNA plasmid standard containing the target sequence. Standard curves of Cp values versus target DNA concentrations for each assay were generated and fit by a least-squares regression model. Confidence intervals of predicted target concentrations on measured Cp values were calculated based on propagation of error in the standard curve (Harris 1995). The limits of detection (LODs) for individual assays were determined at 99% confidential intervals as previously described (Nshimyimana et al. 2014) (Table 2). For HF183-SYBR Green, amplification of the correct target was confirmed based on the criteria that sample melting temperatures (Tm) fall within two standard deviations of the mean Tm associated with qPCR standards at concentrations of 10^1-10^6 copies per qPCR (78.9°C ± SD 0.15).

2. 6 Evaluation of assay sensitivity, specificity and accuracy

Assays for human-associated and dog-associated *Bacteroidales* genetic markers were evaluated for their sensitivity and specificity, and the best markers suitable for tracking sources

of fecal pollution in urban tropical environments were selected. Sensitivity and specificity were calculated as described in Kildare and colleagues and Odagiri and colleagues (Kildare et al. 2007, Odagiri et al. 2015). Briefly, assay sensitivity was determined based on the number of target host samples testing positive for the assayed marker (true positives; TP), divided by the total number of host samples tested (TP + false negatives; FN) (Equation 1). Assay specificity was calculated as the number of non-target host samples testing negative for the assayed marker (true negatives; TN), divided by the total number of non-target samples tested (TN + false positives; FP) (Equation 2). For this study the target sources considered were humans or dogs. We also computed the accuracy of the assays as the ratio of target and non-target host samples identified correctly by total number of samples tested (Equation 3). Pooled samples were not included for sensitivity and specificity calculations. However, their results were used to confirm similar performance in additional samples.

$$\text{Sensitivity} = \frac{\text{TP (True Positive)}}{(\text{TP+FN(False Negative)})} \quad \text{Eq. 1}$$

$$\text{Specificity} = \frac{\text{TN(True Negative)}}{(\text{Tn+FP(False Positive)})} \quad \text{Eq. 2}$$

$$\text{Accuracy} = \frac{(\text{TP+TN})}{(\text{TP+FP+TN+FN})} \quad \text{Eq. 3}$$

TP, TN, FP and FN are as explained in Equations 2 and 3.

Samples were first normalized to copies/ng DNA since sewage could not be expressed in copies/g feces. This normalization method (Layton et al. 2013) was used to assess sensitivity and specificity based on the limit of detection and allowed to take into consideration upper and lower concentrations of markers in all tested samples. *Bacteroidales* assays with high sensitivity, specificity and accuracy were selected as the best performers.

2.7 Statistical analysis

To select the most suitable human-associated *Bacteroidales* 16S rRNA genetic markers for MST in the urban tropical environment of Singapore, sensitivity and specificity were computed and compared across assays. Marker concentrations in host target samples were compared using the t-test or one-way ANOVA followed by Tukey's honest significant difference (HSD) post-hoc multiple-comparisons test using JMP Pro v. 12 (SAS Institute Inc., Cary, NC USA).

3. Results

3.1 Prevalence of *Bacteroidales* markers in human and animal samples

All five human-associated *Bacteroidales* markers were detected in all raw sewage samples (n=20) (i.e., BacHum, BacH, B. theta, HF183, and HF183-SYBR Green) and in 50-70% of human stool samples from volunteers in Singapore (n=35) (Table 3). They were also found in samples from domesticated animals (rabbits: 50% for HF183-SYBR Green, HF183 and BacHum, and 40% for BacH; cats: 10% B. theta; and chickens: 30% HF183-SYBR Green, 20% HF183, 10% BacHum and 30% BacH) (Table 3).

Abundance of markers in fecal samples was expressed per gram wet feces as discussed in Layton et al. (2013) to assess sensitivity and specificity based on the limit of detection while taking into consideration upper and lower concentrations of markers in all tested samples. Concentrations of human-associated markers in rabbit samples ranged from 7.1 to 7.7 $\log_{10}$ gene copies/g wet feces, while the range in cross-reacting chicken samples was from 3.8 to 5.2 $\log_{10}$ gene copies/g wet feces for 10% to 30% (1 to 3/10) cross reactivity. A single cat sample cross-reacted with the B. theta assay (7.1 $\log_{10}$ gene copies/g wet feces) and the canine assay detected two cat samples, recording 6.2 and 6.5 $\log_{10}$ gene copies/g wet feces, respectively (Figure 2). Overall the concentrations of markers in cross-reacting samples were lower than in target

samples (sewage and human stool) with the exception of rabbits and a single chicken sample, where the BacH concentration was similar to concentrations observed in target samples (i.e. 7.7 $\log_{10}$ gene copies/g wet feces).

Human-associated (BacHum and BacH) and canine-associated (BacCan) *Bacteroidales* assays cross-reacted with wild animals (monkeys and boars). However, the concentrations of the markers in the fecal samples from wild animals were very low (< LOD) (Figure 3).

The universal *Bacteroidales* marker (BacUni) was detected in all samples analyzed with the exception of droppings of the mynah birds and 30% of chicken fecal samples. The canine marker (BacCan) was detected in 80% of dog samples and in 2.8% and 20% of the human stool and cat fecal samples, respectively.

3.1.1 Distribution of marker concentrations in human stool and raw sewage samples

In general, sewage samples had a similar range of concentrations for four of five tested human-associated *Bacteroidales* markers (from 4.2 ± 0.5 SD to 3.5 ± 0.4 SD $\log_{10}$ gene copies/ml; Supplemental Table 2), while levels of BacH were the lowest (5.7 ± 1.5 $\log_{10}$ gene copies/g wet feces) among five tested human-associated markers. Concentrations of B. theta, BacHum and HF183 had similar concentrations in human stool samples (namely, 6.0 ± 1.0, 6.0 ± 1.6, and 6.0 ± 1.8). The marker concentrations detected in human stool samples had a comparable variance than sewage samples (4.0 ± 0.6 SD to 7.7 ± 0.8 SD $\log_{10}$ gene copies/g wet feces), likely because overall numbers of target *Bacteroidales* in human stool among individual volunteers were within a similar range as those in the sewage that are homogenized by sewage mixing in the drains during transportation to the wastewater plan (Supplemental Table 2). The canine marker was found at low concentrations in human stool and sewage compared to human markers (<3.9 ± 0.6 $\log_{10}$ gene copies/g wet feces) with the exception of two samples (> 4.0 $\log_{10}$

342 gene copies/g wet feces). The universal marker, BacUni, was detected at the highest
343 concentration in all sewage (5.2±0.3 SD $\log_{10}$ gene copies/ml), human (7.5 ±0.8 SD $\log_{10}$
344 copies/g of wet feces) and animal samples excluding birds (6.7 ± 1.6 SD $\log_{10}$ gene copies/g of
345 wet feces) (Figures 1, 2, 3, and Supplemental Table 2).

346 **3.2 Performance of host-associated *Bacteroidales* assays**

347 **3.2.1 Human associated assays**

348 Generally, all human-associated *Bacteroidales* markers were detected in sewage (100%)
349 resulting in a sensitivity of 1.0, while their sensitivity in human stool ranged from 0.51 to 0.68
350 (Table 4). B. theta and BacHum had the highest sensitivity in human stool (0.68 and 0.65,
351 respectively), and BacH and HF183 had the lowest (0.51 and 0.60, respectively). B. theta and
352 HF183-SYBR Green were the human markers with the highest (0.98) and the lowest (0.88)
353 specificity, respectively, in human stool and sewage (Table 4).

354 The accuracy of human-associated *Bacteroidales* markers was calculated by including
355 true positive human and sewage samples and true negative animal samples. B. theta had the
356 highest accuracy for detecting human stool and sewage samples (0.90, Table 4), while HF183-
357 SYBR Green and HF183 had the lowest accuracy among the five human assays tested (0.83,
358 Table 4).

359 **3.2.2 Universal assay**

360 One candidate universal assay, BacUni, was tested and amplified in raw sewage, human
361 stool and all animal samples except bird droppings (Figures 1, 2, and 3). The sensitivity and
362 specificity of BacUni were 1.0 and 1.0, respectively, in quantifying *Bacteroidales* in sewage,
363 human stool and other animal (mammal) fecal samples (Table 4). The universal assay had an
364 accuracy of 0.97 (Table 4).

365 **3.2.3 Canine-associated assay**

366 The performance of the canine assay, BacCan, was evaluated with 50 fecal samples
367 collected from domesticated dogs in Singapore. Overall the assay amplified 8 out of 10 dog fecal
368 samples and 4 of 4 pooled fecal samples of 10 individuals each (Table 3). The sensitivity and
369 specificity of BacCan were 0.80 and 0.97, respectively (Table 4). BacCan accuracy was 0.96 in
370 identifying domesticated dog fecal samples (Table 4). Sensitivity, specificity and accuracy
371 results confirmed the suitability of BacCan for tracking dog fecal sources in order to enhance
372 urban watershed management.

373

374 **4. Discussion**

375 **4.1 Performance of Canine-associated and Universal *Bacteroidales* assays**

376 BacCan had a sensitivity and specificity above 80% for target detection in dog fecal
377 samples. This high sensitivity of BacCan was previously observed in comparable studies in the
378 USA (Boehm et al. 2013, Schriewer et al. 2013), Canada (Silkie and Nelson 2009, Tambalo et al.
379 2012) and Odisha, India (Odagiri et al. 2015). However, BacCan demonstrated cross-reactivity
380 with septage, goose and cow feces in California, USA (Boehm et al. 2013) and with human stool,
381 sewage, cats, rabbits and chicken fecal samples in our study area, although the number of gene
382 copies found in non-target samples was low. The cross-reactivity of BacCan with sewage, cat
383 feces and human stool was also demonstrated during the development of the assay in the United
384 States (Kildare et al. 2007). Despite some non-target fecal samples cross-reacting with BacCan
385 in different studies, this assay has demonstrated a nearly constant sensitivity and specificity
386 (>80%) across different geographies and longitudes (this study) (Odagiri et al. 2015, Boehm et

al. 2013). Hence, we recommend BacCan as an assay for dog fecal pollution tracking in the urban tropical environment of Singapore and Southeast Asia.

The universal *Bacteroidales* assay (BacUni) showed a high sensitivity and specificity (close to 100%) to fecal samples of mammals tested in Singapore, comparable to the 100% sensitivity or specificity reported in other MST studies in the USA (Kildare et al. 2007), Kenya (Jenkins et al. 2009), and Odisha, India (Odagiri et al. 2015). BacUni has been previously detected in bird samples such as chicken, ducks and geese in USA, Canada and New Zealand (Green et al. 2011). However, it was not detected in any of the 50 Mynah birds or Javan myna (*Eudynamys scolopacea*) fecal samples used in this study. Overall the assay is highly useful for quantifying total *Bacteroidales* in mammals in Singapore and Southeast Asia (sensitivity and specificity close to 100%, Table 4).

4.2 Performance of Human-associated *Bacteroidales* Assays

The need to validate human-associated *Bacteroidales* assays across different geographical regions has become ever more apparent since recent findings of the human gut microbiome studies demonstrated variability in human gut microbial communities across different countries (Yatsunenko et al. 2012) and latitudes (Suzuki and Worobey 2014). Similarly, when two human-associated assays (BacH and BacHum) were tested on 61 human stool samples from 16 different countries across the globe the target sequences had unexpectedly low prevalence (Reischer et al. 2013), underscoring the need to thoroughly test assays at different geographic locations. The human-associated marker HF183 is a case in point; this assay was recommended for human stool and sewage source tracking in California, USA, based on a multi-laboratory validation study (Layton et al. 2013). However, a different validation study in Odisha, India, placed HF183 among poorly performing markers (Odagiri et al. 2015).

Overall concentrations of B. theta in human stool were similar to those of the other four human-associated markers tested and also comparable to ranges reported in the literature (Yampara-Iquise et al. 2008). In contrast, concentrations of B. theta in sewage were low compared to other human-associated *Bacteroidales* markers tested in this study (Supplemental Table 2). Similarly, B. theta and other designed markers targeting functional genes (e.g., gyrB, HumM2 and Mnif) exhibited lower concentrations than other human-associated *Bacteroidales* markers targeting the 16S rRNA gene in an intra-laboratory comparison study in California (Layton et al. 2013). High concentrations of B. theta in environmental and sewage samples were reported in the Great Lakes region in the USA compared to this study (Aslan and Rose 2013, Verhougstraete et al. 2015), confirming that concentrations of a target sequence in a sample (sewage, human stool or environmental) are related to assay design and type, and origin of sample.

B. theta showed the highest sensitivity and specificity (98%) to human stool (sensitivity: 69%) and sewage (sensitivity: 100%); therefore, it was selected as the best marker for Singapore and Southeast Asia (Table 4). Highest specificity of B. theta was previously reported in studies in the Great Lakes region (97%) (Aslan and Rose 2013), California (96%) (Layton et al. 2013) and Missouri (100%) (Yampara-Iquise et al. 2008) in the USA. B. theta only cross-reacted with one cat sample (1/10) among all animals (n = 295) tested in this study (Figure 2), while a validation study in the Great Lake region in the USA revealed B. theta cross reactivity for cats (1/6), gulls (3/25) and swine (3/35) (Aslan and Rose 2013). Furthermore, another recent study in the tropical environment of northern Australia demonstrated that B. theta cross-reacted with frog excrements (Neave et al. 2014). The first study in Missouri, USA, did not report any cross-reactivity for B. theta (Yampara-Iquise et al. 2008). Taken together, these findings establish

433 high-level specificity of B. theta across four distinct regions worldwide including the Great
434 Lakes region, Missouri and California in the USA, north Australia and Southeast Asia relative to
435 other human-associated *Bacteroidales* markers previously used in these regions such as HF183
436 (Layton et al. 2013, Neave et al. 2014, Yampara-Iquise et al. 2008, Aslan and Rose 2013).

437 Cross-reactivity of human-associated *Bacteroidales* markers with non-target samples
438 remains a concern because it can limit their use in the field and complicate the interpretation of
439 monitoring results. Indeed, human-associated *Bacteroidales* markers (except B. theta) tested in
440 this study (HF183-SYBR Green, HF183, BacHum and BacH) were less specific to human stool
441 and sewage as they cross-reacted with domesticated animals (rabbits and chicken, Figure 2).
442 These host-associated *Bacteroidales* assays have been previously identified as cross-reactive
443 with animal feces, often with low reported concentrations (Haugland et al. 2010). BacHum,
444 which was previously identified to cross-react with cow, dog and horse fecal samples (Kildare et
445 al. 2007), did not amplify with the dog samples in this study (Figure 2) although cross-reactivity
446 was observed at high concentrations in rabbits (> 7.0 $\log_{10}$ gene copies/g wet feces).

447 Rabbits are admired pets due in part to their size and the dense urban living conditions in
448 Singapore. According to the House Rabbit Society Singapore over 1,000 rabbits are rescued
449 every year by the Society for the Prevention of Cruelty of Animals and an undisclosed number
450 are euthanized due to lack of enough shelter facilities (http://www.hrss.net/aar/aar_care.html).
451 The number of rabbits is estimated to be similar to that of cats. Fecal waste from these
452 domesticated animals is often discharged via toilets thus mixing with sewage. Most potential
453 sewage exfiltration and transportation to open channel drains happens at the level of housing
454 developments (Ekklesia et al. 2015b). These findings suggest that rabbit feces is an important

455 potential contamination source as leaks from sewage pipes may happen before dilution and
456 homogenization of sewage.

457 Validated human-associated *Bacteroidales* markers in this study (Singapore) and other
458 tropical locations (India and Bangladesh) suggest a relationship with geography and latitudes.
459 MST validation studies carried out in Odisha, India and Bangladesh selected BacHum and
460 HF183 as the best human-associated *Bacteroidales* markers, respectively (Ahmed et al. 2010,
461 Odagiri et al. 2015). HF183 was also recently recommended for microbial source tracking in
462 California, USA (Layton et al. 2013); however, it was not the best performing assay in our
463 Southeast Asia study. BacHum, the best performing assay for human fecal pollution detection in
464 Odisha, India, was the second best performer in Singapore (Table 4).

465 This study validated the application of host-*Bacteroidales* markers for microbial source
466 tracking in Southeast Asia. A recent study established decay models for B. theta in artificial fresh
467 and seawater as a function of sunlight intensity and salinity, with one $\log_{10}$ reduction times (T_{90})
468 from 3.1 to 53 h (Liang et al. 2017). The lower concentration of B. theta in sewage in Singapore
469 suggests that the marker, if detected in environmental samples, should be viewed as an indicator
470 of recent contamination and could be complemented with another validated marker of greater
471 abundance in sewage, such as BacHum. Although the use of molecular based methodology may
472 require a high investment for regional developing countries, the outcome of applying these
473 markers could have a high return associated with sewage leak mitigation and control of fecal
474 non-point sources of pollution in the urban environment.

475 **5. Conclusions**

476 We tested five human-associated, one canine-associated and one universal *Bacteroidales* assay
477 against 295 fecal samples to identify host-associated *Bacteroidales* assays for source tracking in
478 Southeast Asia.

479 • The human-associated *Bacteroidales* assays B. theta and BacHum, the canine-
480 associated assay BacCan, and the universal *Bacteroidales* assay BacUni are validated
481 assays for use in the tropical urban environment of Southeast Asia.

482 • Rabbit fecal samples cross-reacted with all human-associated assays except B. theta.
483 This calls for caution when considering MST studies in the region and confirms the
484 usefulness of the highest specificity of B. theta to discriminate human fecal
485 contamination from other sources.

486 • Findings of this study should assist emerging densely populated cities in the Southeast
487 Asia region with limited sewage management infrastructures to efficiently identify and
488 track potential sources of human fecal pollution for mitigation of sewer leaks and
489 protection of recreational waters.

490 • Fate and transport studies of these validated MST markers under tropical conditions in
491 water and sediments are needed to correctly identify pollution events in waterways and
492 water bodies on a temporal and spatial scale.

493 • Overall, we demonstrated the need for similar validation studies before applying any
494 existing human-associated *Bacteroidales* markers in different geographical locations.
495 Regional application of the findings of this study would benefit from optimization of a
496 regional protocol under urban tropical environmental conditions.

497

498 **Acknowledgements**

499 This research was supported by the National Research Foundation Singapore and Ministry of
500 Education under the research center of excellence program and through the Singapore MIT
501 Alliance for Research and Technology's Center for Environmental Sensing and Modeling
502 interdisciplinary research program. The authors gratefully acknowledge the generous assistance
503 of the Singapore National Park Board, the Singapore Land Authority, the Society for the
504 Prevention of Cruelty to Animals Singapore, and the many volunteers who provided stool
505 samples. Jean Pierre Nshimyimana thanks Minji Kim and Odagiri Mitsunori for providing
506 positive control strains, Anisa Cokro and Anika Cokro for their help in fecal sample collection,
507 and Alex Schriewer for helpful discussions on qPCR strategies.

References

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 56(10), 838-845.

Ahmed, W., Powell, D., Goonetilleke, A. and Gardner, T. (2008a) Detection and source identification of faecal pollution in non-sewered catchment by means of host-specific molecular markers. Water Science and Technology 58(3), 579-586.

Ahmed, W., Stewart, J., Powell, D. and Gardner, T. (2008b) Evaluation of *Bacteroides* markers for the detection of human faecal pollution. Letters in Applied Microbiology 46(2), 237-242.

Anonymous .(2011) TaqMan Fast Advanced Master Mix on the Roche LightCycler 480 Real-Time PCR System, Applied Biosystems by Life Technologies.

Aslan, A. and Rose, J.B. (2013) Evaluation of the host specificity of *Bacteroides thetaiotaomicron alpha-1-6, mannanase* gene as a sewage marker. Letters in Applied Microbiology 56(1), 51-56.

Bernhard, A.E. and Field, K.G. (2000a) Identification of Nonpoint Sources of Fecal Pollution in Coastal Waters by Using Host-Specific 16S Ribosomal DNA Genetic Markers from Fecal Anaerobes. Applied and Environmental Microbiology 66(4), 1587-1594.

Bernhard, A.E. and 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.

Boehm, A., Van De Werfhorst, L.C., Griffith, J., Holden, P., Jay, J., Shanks, O., Wang, D. and Weisberg, S. (2013) Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. Water Research 47(18), 6812-6828.

Ekklesia, E., Shanahan, P., Chua, L.H.C. and Eikaas, H.S. (2015a) Temporal variation of faecal indicator bacteria in tropical urban storm drains. Water Research 68(1), 10.

Ekklesia, E., Shanahan, P., Chua, L.H. and Eikaas, H.S. (2015b) Associations of chemical tracers and faecal indicator bacteria in a tropical urban catchment. Water Research 75, 270-281.

Feachem, R. (1975) Note: An improved role for faecal coliform to faecal *streptococci* ratios in the differentiation between human and non-human pollution sources. Water Research 9(7), 689.

Gawler, A.H., Beecher, J.E., Brandão, J., Carrollf, N.M., Falcão, L., Gourmelon, M., Masterson, B., Nunes, B., Porter, J. and Rincé, A. (2007) Validation of host-specific *Bacteriodales* 16S

539 rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of
540 the European Union. Water Research 41(16), 3780.

541 Green, H., Shanks, O., Sivaganesan, M., Haugland, R. and Field, K. (2011) Differential decay of
542 human faecal Bacteroides in marine and freshwater. Environmental Microbiology 13(12), 3235-
543 3249.

544 Green, H.C., Dick, L.K., Gilpin, B., Samadpour, M. and Field, K.G. (2012) Genetic markers for
545 rapid PCR-based identification of gull, Canada Goose, duck, and chicken fecal contamination in
546 water. Applied and Environmental Microbiology 78(2), 503-510.

547 Green, H.C., Haugland, R.A., Varma, M., Millen, H.T., Borchardt, M.A., Field, K.G., Walters,
548 W.A., Knight, R., Sivaganesan, M., Kelty, C.A. and Shanks, O.C. (2014) Improved HF183
549 Quantitative Real-Time PCR Assay for Characterization of Human Fecal Pollution in Ambient
550 Surface Water Samples. Applied and Environmental Microbiology 80(10), 9.

551 Harris, D.C. (1995) 4th Edition: Quantitative Chemical Analysis. W. H. Freeman and Company,
552 New York.

553 Harwood, V.J., Brownell, M., Wang, S., Lepo, J., Ellender, R.D., Ajidahun, A., Hellein, K.N.,
554 Kennedy, E., Ye, X. and Flood, C. (2009) Validation and field testing of library-independent
555 microbial source tracking methods in the Gulf of Mexico. Water Research 43(19), 4812-4819.

556 Haugland, R.A., Varma, M., Sivaganesan, M., Kelty, C., Peed, L. and Shanks, O.C. (2010)
557 Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative
558 detection of selected *Bacteroidales* species and human fecal waste by qPCR. Systematic and
559 Applied Microbiology 33(6), 348-357.

560 Jenkins, M.W., Tiwari, S., Lorente, M., Gichaba, C.M. and Wuertz, S. (2009) Identifying human
561 and livestock sources of fecal contamination in Kenya with host-specific *Bacteroidales* assays.
562 Water Research 43(19), 4956-4966.

563 Kildare, B., Leutenegger, C., McSwain, B., Bambic, D., Rajal, V. and Wuertz, S. (2007) 16S
564 rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal
565 *Bacteroidales*: A Bayesian approach. Water Research 41(16), 15.

566 Layton, B., Cao, Y., Ebentier, D., Hanley, K., Ballesté, E., Brandão, J., Byappanahalli, M.,
567 Converse, R., Farnleitner, A., Gentry Shields, J., Gidley, M., Gourmelon, M., Lee, C., Lee, J.,
568 Lozach, S., Madi, T., Meijer, W., Noble, R., Peed, L., Reischer, G., Rodrigues, R., Rose, J.,
569 Schriewer, A., Sinigalliano, C., Srinivasan, S., Stewart, J., Van De Werfhorst, L.C., Wang, D.,
570 Whitman, R., Wuertz, S., Jay, J., Holden, P., Boehm, A., Shanks, O. and Griffith, J. (2013)
571 Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method
572 evaluation study. Water Research 47(18), 6897-6908.

573 Liang, L., Goh, S.G., Vergara, G.G.R.V., Fang, H.M., Rezaeinejad, S., Chang, S.Y., Bayen, S.,
574 Lee, W.A., Sobsey, M.D., Rose, J.B. and Gin, K.Y.H. (2015) Alternative fecal indicators and
575 their empirical relationships with enteric viruses, *Salmonella* enterica, and *Pseudomonas*
576 *aeruginosa* in surface waters of a tropical urban catchment. Applied and Environmental
577 Microbiology 81(3), 850-860.

578 Liang, L., Goh, S.G. and Gin, K.Y.H. (2017) Decay kinetics of microbial source tracking (MST)
579 markers and human adenovirus under the effects of sunlight and salinity. Science of the Total
580 Environment 574, 165-175.

581 Mattioli, M.C., Pickering, A.J., Gilsdorf, R.J., Davis, J. and Boehm, A.B. (2013) Hands and
582 water as vectors of diarrheal pathogens in Bagamoyo, Tanzania. Environmental Science and
583 Technology 47(1), 355-363.

584 Neave, M., Luter, H., Padovan, A., Townsend, S., Schobben, X. and Gibb, K. (2014) Multiple
585 approaches to microbial source tracking in tropical northern Australia. Microbiology Open 3(6),
586 860-874.

587 NParks (2016) Parks and nature reserves. Agency, S.N.P. (ed), National Parks, Singapore
588 Government.

589 Nshimyimana, J.P., Ekklesia, E., Shanahan, P., Chua, L.H.C. and Thompson, J.R. (2014)
590 Distribution and abundance of human-specific *Bacteroides* and relation to traditional indicators
591 in an urban tropical catchment. Journal of Applied Microbiology 116(5), 1369-1383.

592 Odagiri, M., Schriewer, A., Hanley, K., Wuertz, S., Misra, P.R., Panigrahi, P. and Jenkins, M.W.
593 (2015) Validation of *Bacteroidales* quantitative PCR assays targeting human and animal fecal
594 contamination in the public and domestic domains in India. Science of the Total Environment
595 502, 462-470.

596 Okabe, S., Okayama, N., Savichtcheva, O. and Ito, T. (2007) Quantification of host-specific
597 *Bacteroides-Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in
598 freshwater. Applied microbiology and Biotechnology 74(4), 890-901.

599 Pickering, A.J., Julian, T.R., Mamuya, S., Boehm, A.B. and Davis, J. (2011) Bacterial hand
600 contamination among Tanzanian mothers varies temporally and following household activities.
601 Tropical Medecine and International Health 16(2), 233-239.

602 Pickering, A.J., Julian, T.R., Marks, S.J., Mattioli, M.C., Boehm, A.B., Schwab, K.J. and Davis,
603 J. (2012) Fecal contamination and diarrheal pathogens on surfaces and in soils among Tanzanian
604 households with and without improved sanitation. Environmental Science and Technology
605 46(11), 5736-5743.

606 Reischer, G.H., Ebdon, J.E., Bauer, J.M., Schuster, N., Ahmed, W., Åström, J., Blanch, A.R.,
607 Blöschl, G., Byamukama, D., Coakley, T., Ferguson, C., Goshu, G., Ko, G., de Roda Husman,
608 A.M., Mushi, D., Poma, R., Pradhan, B., Rajal, V., Schade, M.A., Sommer, R., Taylor, H., Toth,
609 E.M., Vrajmasu, V., Wuertz, S., Mach, R.L. and Farnleitner, A.H. (2013) Performance
610 Characteristics of qPCR Assays Targeting Human- and Ruminant-Associated *Bacteroidetes* for
611 Microbial Source Tracking across Sixteen Countries on Six Continents. Environmental Science
612 and Technology 47(15), 8548-8556.

613 Santo Domingo, J.W., Bambic, D.G., Edge, T.A. and Wuertz, S. (2007) Quo vadis source
614 tracking? Towards a strategic framework for environmental monitoring of fecal pollution. Water
615 Research 41(16), 3539-3552.

616 Schriewer, A., Goodwin, K.D., Sinigalliano, C.D., Cox, A.M., Wanless, D., Bartkowiak, J.,
617 Ebentier, D.L., Hanley, K.T., Ervin, J., Deering, L.A., Shanks, O.C., Peed, L.A., Meijer, W.G.,
618 Griffith, J.F., SantoDomingo, J., Jay, J.A., Holden, P.A. and Wuertz, S. (2013) Performance
619 evaluation of canine-associated *Bacteroidales* assays in a multi-laboratory comparison study.
620 Water Research 47(18), 6909-6920.

621 Seurinck, S., Defoirdt, T., Verstraete, W. and Siciliano, S. (2005) Detection and quantification of
622 the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for
623 assessment of human faecal pollution in freshwater. Environmental Microbiology 7(2), 249-259.

624 Silkie, S.S. and Nelson, K.L. (2009) Concentrations of host-specific and generic fecal markers
625 measured by quantitative PCR in raw sewage and fresh animal feces. Water Research 43(19),
626 4860-4871.

627 Suzuki, T.A. and Worobey, M. (2014) Geographical variation of human gut microbial
628 composition. Biology Letters 10(2).

629 Tambalo, D.D., Boa, T., Liljebjelke, K. and Yost, C.K. (2012) Evaluation of two quantitative
630 PCR assays using Bacteroidales and mitochondrial DNA markers for tracking dog fecal
631 contamination in waterbodies. Microbiological Methods 91(3), 459-467.

632 Verhougstraete, M.P., Martin, S.L., Kendall, A.D., Hyndman, D.W. and Rose, J.B. (2015)
633 Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at
634 the basin scale. Proceedings of the National Academy of Sciences of the United States of
635 America 112(33), 10419-10424.

636 Widmer, K., Van Ha, N.T., Vinitnantharat, S., Sthiannopkao, S., Wangsaatmaja, S., Prasetiati,
637 M.A., Thanh, N.C., Thepnoo, K., Sutadian, A.D., Thao, H.T., Fapyane, D., San, V., Vital, P. and
638 Hur, H.G. (2013) Prevalence of *Escherichia coli* in surface waters of Southeast Asian cities.
639 World Journal of Microbiology and Biotechnology 29(11), 2115-2124.

640 Wilkes, G., Brassard, J., Edge, T.A., Gannon, V., Jokinen, C.C., Jones, T.H., Marti, R.,
641 Neumann, N.F., Ruecker, N.J., Sunohara, M., Topp, E. and Lapen, D.R. (2013) Coherence
642 among Different Microbial Source Tracking Markers in a Small Agricultural Stream with or
643 without Livestock Exclusion Practices. Applied and Environmental Microbiology 79(20), 6207-
644 6219.

645 Yampara-Iquise, H., Zheng, G., Jones, J.E. and Carson, C.A. (2008) Use of a *Bacteroides*
646 *thetaiotaomicron-specific alpha-1-6, mannanase* quantitative PCR to detect human faecal
647 pollution in water. Applied Microbiology 105(5), 1686-1693.

648 Yatsunenko, T., Rey, F.E., Manary, M.J., Trehan, I., Dominguez-Bello, M.G., Contreras, M.,
649 Magris, M., Hidalgo, G., Baldassano, R.N., Anokhin, A.P., Heath, A.C., Warner, B., Reeder, J.,
650 Kuczynski, J., Caporaso, J.G., Lozupone, C.A., Lauber, C., Clemente, J.C., Knights, D., Knight,
651 R. and Gordon, J.I. (2012) Human gut microbiome viewed across age and geography. Nature
652 486(7402), 222-227.

653

654

655 **Table 1. Oligonucleotide primers and probes used for qPCR measurements of host-associated *Bacteroidales* gene markers**

Assay	Primer or probe[1]	Concentration in 20-μl reaction mix	Oligonucleotide (5′ to 3′)	Amplicon size	Reference
HF183 (TaqMan)	HF183_F	1000 nM	ATCATGAGTTCACATGTCCG	120 bp	Bernhard and Field 2000a & Green et al. 2014
	BacR287_R	1000 nM	CTTCCTCTCAGAACCCCTATCC		
	BacP234_Probe	80 nM	FAM-CTAATGGAACGCATCCC-MGB		
HF183 SYBR Green	HF183_F	400 nM	ATCATGAGTTCACATGTCCG	83 bp	Seurinck et al. 2005
	Bac242_R	400 nM	TACCCCGCCTACTATCTAATG		
BacHum (TaqMan)	BacHum160_F	400 nM	TGAGTTCACATGTCCGCATGA	81 bp	Kildare et al. 2007
	BacHum241_R	400 nM	CGTTACCCCGCCTACTATCTAATG		
	BacHum193_P	80 nM	FAM-TCCGGTAGACGATGGGGATGCGTT-TAMRA		
BacH (TaqMan)	BacH_F	200 nM	CTTGGCCAGCCTTCTGAAAG	51 bp	Ahmed et al. 2009
	BacH_R	200 nM	CCCCATCGTCTACCGAAAATAC		
	BaCH_P1	100 nM	FAM-TCATGATCCCATCCTG-MGB		
	BaCH_P2	100 nM	FAM-TCATGATGCCATCTTG-MGB		
BacCan (TaqMan)	BacCan545_F	400 nM	GGAGCGCAGACGGGTTTT	145 bp	Kildare et al. 2007
	BacCan690_R1	400 nM	CAATCGGAGTTCTTCGTGATATCTA		
	BacCan690_R2	400 nM	AATCGGAGTTCCTCGTGATATCTA		
	BacCan656_P	80 nM	FAM-TGGTGTAGCGGTGAAA-MGB		
BacUni (TaqMan)	BacUni520_F	400 nM	CGTTATCCGGATTTATTGGGTTTA	145 bp	Kildare et al. 2007
	BacUni690_R1	400 nM	CAATCGGAGTTCTTCGTGATATCTA		
	BacUni-690_R2	400 nM	AATCGGAGTTCCTCGTGATATCTA		
	BacUni656_P	80 nM	FAM-TGGTGTAGCGGTGAAA-MGB		
B. theta (TaqMan)	BtH4515901_F	900 nM	CATCGTTCGTCAGCAGTAACA	63 bp	Yampara-Iquise et al. 2008
	BtH4515963_R	900 nM	CCAAGAAAAAGGGACAGTGG		
	BtH_P	250 nM	Roche Universal library probe #062		

656 [1] F: Forward; R: Reverse; P: Probe. FAM, 6-carboxyfluorescein; TAMRA, 6-carboxytetramethylrhodamine MGB Minor groove binder

657 **Table 2. Performance characteristics of tested *Bacteroidales* assays**

Assay	Slope	Y-intercept	R^2	Efficiency	LOD[1] (copies/μl)
HF183-SYBR Green	-3.4	36.8	0.98	93.4	21.0
HF183	-3.3	37.8	0.99	98.0	28.4
BacH	-3.3	37.4	0.97	101.0	51.5
BacHum	-3.3	38.8	0.99	99.2	43.4
B. theta	-3.3	41.0	0.99	97.0	88.0
BacCan	-3.4	39.5	0.99	95.7	86.8
BacUni	-3.3	37.6	0.99	100.5	13.3

658
659 [1] Limit of detection

660 **Table 3. Performance of human- and canine-associated and universal *Bacteroidales* assays as revealed by qPCR quantification**
661 **of fecal and raw sewage samples**

Sample	Percentage or number of positive samples							n
	Human Assay					Canine Assay	Universal Assay	
	HF183-SYBR Green	HF183	BacHum	BacH	B. theta	BacCan	BacUni	
Human								
Sewage	100	100	100	100	100	0	100	20
Human Stool	62.8	60	65.7	51.4	68.5	2.8	100	35
Wild animals								
Monkey	0	0	0	0	0	0	100	10
Wild Boars	0	0	0	0	0	0	100	10
Birds (Javan myna (Acridotheres javanicus)	0	0	0	0	0	0	0	10
Domesticated animals								
Chickens	30	20	10	30	0	0	70	10
Rabbits	50	50	50	40	0	0	100	10
Dogs	0	0	0	0	0	80	100	10
Cats	0	0	0	0	10	20	100	10
Pooled sewage and animal samples								
Sewage (n=10)	1/1	1/1	1/1	1/1	1/1	0/1	1/1	1
Rabbits (n=30)	2/3	2/3	2/3	2/3	0/3	0/3	3/3	3
Cats (n=40)	0/4	0/4	0/4	0/4	1/4	3/4	4/4	4
Dogs (n=40)	1/4	0/4	2/4	0/4	0/4	4/4	4/4	4
Birds (n=40)	0/4	0/4	0/4	0/4	0/4	0/4	0/4	4

662

Table 4. Sensitivity, specificity, and accuracy of universal, canine-associated and human-associated *Bacteroidales* assays in Singapore

Assay	Target	Sensitivity		Specificity	Accuracy
		Raw Sewage	Human Stool		
Human					
HF183-SYBR Green	Human	1	0.63	0.88	0.83
HF183		1	0.60	0.90	0.83
BacHum		1	0.65	0.91	0.85
BacH		1	0.51	0.90	0.81
B. theta		1	0.69	0.98	0.90
Universal					
BacUni	Mammal	1	1	1	1
Canine					
BacCan	Dog	-	0.80	0.97	0.96

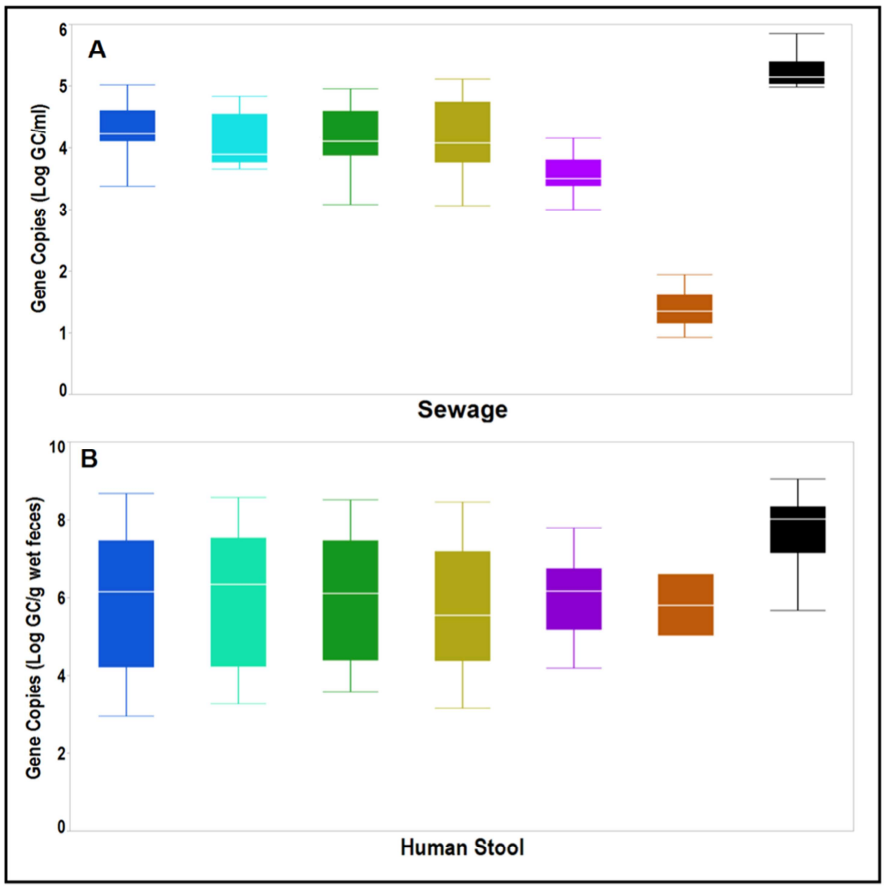


Figure 1: Concentrations of *Bacteroidales*-associated genetic markers in (A) sewage and (B) human stool samples. Each boxplot shows median, upper and lower quartiles spanning maximum and minimum observations. Colors indicate different *Bacteroidales*-associated assays tested where ■ HF183-SYBR Green, ■ HF183, ■ BacHum, ■ BacH, ■ B. theta, ■ BacCan and ■ BacUni. GC = gene copies.

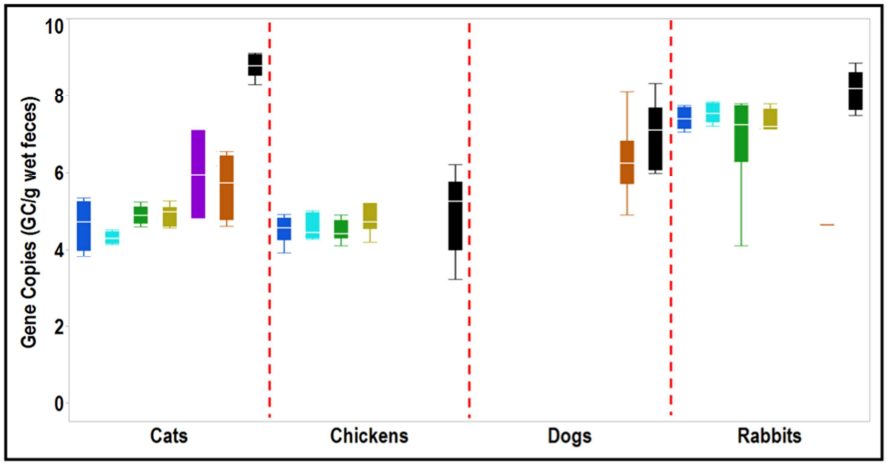


Figure 2. Concentrations of *Bacteroidales*-associated gene markers in domesticated animal fecal samples. Colors indicate different *Bacteroidales*-associated assays tested where ■ HF183-SYBR Green, ■ HF183, ■ BacHum, ■ BacH, ■ B. theta, ■ BacCan and ■ BacUni. All plotted concentrations of *Bacteroidales* markers were above the limit of detection (LOD). The red discontinuous lines separate types of fecal samples tested. GC = gene copies.

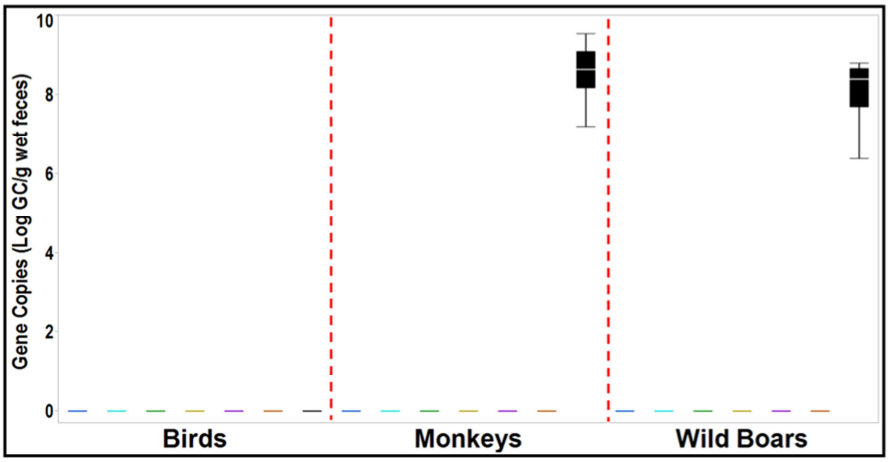


Figure 3. Concentrations of *Bacteroidales*-associated 16S rRNA gene markers in wild animal fecal samples. Colors indicate different Bacteroidales-associated assays tested where ■ HF183-SYBR Green, ■ HF183, ■ BacHum, ■ BacH, ■ B. theta, ■ BacCan and ■ BacUni. All plotted concentrations were above the limit of detection (LOD). The red discontinuous lines separate types of fecal samples tested. GC = gene copies.

- Validated *Bacteroidales*-associated assays applicable in Singapore and Southeast Asia
- B. theta validated as the most specific of five human assays tested
- Wild monkeys and boars did not cross-react with human-assays
- Rabbit cross-reacted with all human assays except B. theta