Microbial source tracking and dynamics of bacterial communities and pathogens in the urban tropical environment

Nshimiyimana, Jean Pierre

2017

Nshimiyimana, J. P. (2017). Microbial source tracking and dynamics of bacterial communities and pathogens in the urban tropical environment. Doctoral thesis, Nanyang Technological University, Singapore.

<http://hdl.handle.net/10356/72263>

<https://doi.org/10.32657/10356/72263>

NANYANG
TECHNOLOGICAL
UNIVERSITY

**Microbial Source Tracking and Dynamics of Bacterial Communities
and Pathogens in the Urban Tropical Environment**

Jean Pierre Nshimyimana

School of Civil and Environmental Engineering

2017

**Microbial Source Tracking and Dynamics of Bacterial Communities
and Pathogens in the Urban Tropical Environment**

Jean Pierre Nshimiyimana

School of Civil and Environmental Engineering

A thesis submitted to the Nanyang Technological University in fulfilment of the
requirement for the degree of Doctor of Philosophy

2017

Acknowledgements

I would like to extend my gratitude to my supervisors Professor Stefan Wuertz, Singapore Center for Environmental Life Science Engineering (SCELSE), Nanyang Technological University (NTU) and Prof. Janelle Thompson, Civil and Environmental Engineering, Massachusetts Institute of Technology (MIT) for their continuous encouragement, guidance and support in everyday research work. You enabled me to develop an in-depth understanding of my research topic.

I am also grateful to members of the Thompson Lab at MIT: Dr Kyle Pete, Dr Kevin Penn, Dr Adam Freedman, and Dr Samodha C. Fernando; members of the Center for Environmental Sensing and Modeling (CENSEM): Dr Veronique Blanc, Dr Eric Adams, Dr Peter Shanahan, Dr Benjamin Stephen Grandey, Dr Masaaki Kitajima, Dr Martin Tay, Yu Fen and Nasren Begam; members of the Gin Laboratory at National University of Singapore (NUS): Professor Karina Gin Yew-Hoong, Dr Te Shu Harn; Dr Goh Shin Giek; Dr Loh Lay Leng, Dr Charmaine Ng, and Nur Hazimah Binte Mohamed; members of the Wuertz laboratory at the University of California, Davis: Dr Minji Kim, Dr Odagiri Mitsunori, and Dr Alex Schriewer; and members of Singapore Center for Environmental Life Science Engineering (SCELSE): Ms Anisa Cokro, Ms Anika Cokro, Dr Cecilia Cruz, Dr Carla Lutz Maree, Mr Ezequiel Santillan, Dr Florentin Constancias, Mr Kumaravel Kanda Swamy, and Associate Professor Diane McDougald for their continuous guidance, encouragement, and support during my doctoral studies.

To all my colleagues in the PhD program at NTU and SMA Doctoral students under Singapore-MIT Alliance for Research and Technology (SMART) program, you have been supportive and I am grateful to have had you on this continuous PhD adventure. I am thankful to former MIT students Cameron Dixon, Kevin Foley, Erika Granger, Kathleen Kerigan, Adriana Mendez-Sagel, Shammi Quddus, Amruta Sudhalkar, and Jessica Yeager, NTU students Li-Jun Lee, Dr Eveline Ekklesia, Yu-Ling Por, and Syed Alwi Bin Sheikh Bin Hussien Alkaff of NTU for their assistance with fieldwork. I am grateful for the generous assistance of the Singapore National Park Board, Singapore Land Authority, Society for the Prevention of Cruelty to Animals Singapore and volunteers who provided stool samples.

I thank the National Research Foundation Singapore under its Campus for Research Excellence and Technological Enterprise programme under The Center for Environmental Sensing and Modeling, an interdisciplinary research group of the Singapore MIT Alliance for Research and Technology, and Ministry of Education under the research center of excellence program, Singapore Center for Environmental Life Sciences and Engineering. Your resources made my academic life and research possible in Singapore.

I owe my deepest gratitude to my wife Georgia M. Arnold for her love, patience, and continuous support. I am deeply thankful for my parents, Marie Josée Nyirakabanza and Emmanuel Mudahemuka, my brother, Mr Nsanzimana, and sisters: Mrs Akimana, Ms Dukuzimana and Mrs Mukamana, and to the family of Mr. Bill Wyman and Mrs Rosalie Smith Wyman. I am heartily thankful for your encouragement, inspiration, and moral support.

Summary

The occurrence of waterborne diseases among recreational water users in developed countries and high annual deaths associated with fecal pollution of water sources in developing countries underscore the need for comprehensive microbial water quality assessment to protect public health. The goals of my research were (1) to examine general bacterial community composition and specific sequences characteristic of sewage associated taxa (SAT) and pathogen-like sequences (PLS) as a function of water quality, (2) to validate host-associated *Bacteroidales* genetic markers for microbial sources tracking in Singapore, (3) to compare the performance of digital droplet PCR (ddPCR) and quantitative (q)PCR, and (4) to evaluate sediment-associated decay rates of the best-performing human-associated markers *B. thetaiotaomicron* α -1-6, mannanase (*B. theta*) and BacHum and those of fecal indicator bacteria (FIB) in the presence and absence of grazing protozoa.

To achieve these objectives, the V3-V4 regions of the 16S rRNA gene were sequenced in 36 samples from a Singapore water catchment using next generation sequencing (MiSeq 250/250). *Escherichia coli* was a better predictor of PLS and SAT than the human-associated genetic marker HF183-SYBR or land use.

The validation of host-associated *Bacteroidales* was achieved using 295 fecal samples from animals and seven *Bacteroidales* genetic markers, including five human-associated assays (HF183-SYBR, HF183, BacHum, BacH and *B. theta*, one canine-associated (BacCan), and a total *Bacteroidales* assay (BacUni). HF183-SYBR was a poor performer compared to other markers. Two human- (*B. theta* and BacHum) and

one canine-associated *Bacteroidales* marker (BacCan), and one universal *Bacteroidales* marker (BacUni) were validated for microbial source tracking (MST) in Singapore and Southeast Asia.

The comparison of qPCR and digital droplet PCR (ddPCR) was done with equal amounts (2 ng/μL) of DNA extracted from 180 fecal samples quantified using human-associated *Bacteroidales* markers (B. theta and BacHum). DdPCR had a high accuracy and reproducibility in detecting lower concentrations (<100 copies/reaction), while qPCR was more suited to the detection of higher concentrations (10^5 to 10^7 copies/reaction). BacHum specificity was improved when using ddPCR, while the sensitivity remained unchanged for both markers (BacHum and B. theta). B. theta concentrations detected by qPCR were 3.3 times higher than levels detected by ddPCR. BacHum concentrations measured by qPCR were 12 times higher than those of ddPCR.

Digital droplet PCR (ddPCR) was used to measure concentrations of *Bacteroidales* viable cells (PMA-treated) and DNA (non-PMA treated) to assess their decay rates in sediments under dark and anoxic conditions at a constant temperature of 30°C in microcosms experiencing grazing and reduced grazing by protozoa. Concentrations of cultivable *E. coli* and *Enterococcus* cells were also measured using Colilert and Enterolert tests, and their decay rate constants were compared to those of *Bacteroidales* genetic markers (BacHum and B. theta). *Bacteroidales* viable cells and DNA had similar ranges of 2-log removal (4 to 9 days), while *E. coli* and *Enterococcus* took 14 and 3 days, respectively. Protozoa did not have any influence on the decay of viable cells and DNA of *Bacteroidales*, and FIB (*E. coli* and *Enterococcus*).

Findings of my research provided insights on the use of advanced molecular tools for microbial water quality assessment in urban tropical environments of Singapore and Southeast Asia. Researchers with interest in microbial water quality assessment in urban tropical environments will use insights and future research direction to continue to improve microbial water quality in tropical urban environments.

List of Publications

2014

1. **Nshimiyimana, J. P.**, Ekklesia, E., Shanahan, P., Chua, L. C. H., and Thompson, J. R. (2014). Distribution and Abundance of Human Specific *Bacteroides* and Relation to Traditional Indicators in an Urban Tropical Catchment. *Journal of Applied Microbiology*, 116 (5), pp. 1369 – 13683. Link (doi):<http://onlinelibrary.wiley.com/doi/10.1111/jam.12455/abstract>

2017

2. **Nshimiyimana, J. P.**, Cruz M.C., Thompson R.J., Wuertz S. (2017). *Bacteroidales* markers for microbial source tracking in Southeast Asia. *Journal of Water Research*, 118, pp. 239-248
Link (doi): <https://doi.org/10.1016/j.watres.2017.04.027>
3. **Nshimiyimana, J. P.**, Freedman, A.J.E., Shanahan, P., Chua, L. C. H., and Thompson, R. J. (2017). Variation of bacterial communities with water quality in an urban tropical catchment. *Journal of Environmental Science & Technology. ASP* (Published date on the website: April 17, 2017)
Link (doi): <http://pubs.acs.org/doi/abs/10.1021/acs.est.6b04737>

Table of Contents

Acknowledgements	iii
Summary	v
List of Publications.....	viii
List of Table	xv
List of Figures	xvii
Chapter 1: Introduction	1
1.1 Background.....	1
1.2 Focus of the Dissertation	5
Reference: Chapter 1	9
Chapter 2. Literature Review	17
2.1 Microbial water pollution	17
2.2 Microbial water pollution in urban tropical waterways.....	17
2.2.1 Urban tropical waterways	17
2.2.2 Microbial water pollutions in urban waterways.....	19
2.2.3 Microbial water pollution in Southeast Asia.....	20
2.2.4 Microbial water pollution in Singapore	22
2.3 Human health risks related to pathogens harbored by microbial pollution ..	24
2.4 Molecular methods applied to detect, quantify and determine sources of microbial pollution.....	29
2.5 Survival and decay of source identifiers: Can we discriminate current from aged fecal pollutions?.....	33

2.6 New sequencing-based tool to study pathogens harbored by microbial pollution in urban areas	37
References: Chapter 2	39
Chapter 3. Variation of Bacterial Communities with Water Quality in an Urban Tropical Catchment	52
Abstract	52
3.1 Introduction	53
3.2. Methodology	55
3.2.1 Study design and sample collection	55
3.2.2 Illumina sequencing	56
3.2.3 Identification of pathogen-like sequences (PLS) and sewage-associated taxa (SAT)	58
3.2.3 Data Analysis	59
3.3. Results	61
3.3.1 Bacterial Community Composition (BCC) in an Urban Reservoir and Catchment	61
3.3.2 Relationship between Human <i>Bacteroides</i> Marker (HF183) and Proportion of <i>Bacteroides</i> and <i>Bacteroides dorei</i> Sequences.	64
3.3.3 Distribution and Variation of Sewage-Associated Taxa (SAT).....	66
3.3.4 Distribution and variation of Pathogen-like Sequences	69
3.3.5 Discussion	76
3.3.6 Conclusion	79

References: Chapter 3.....	80
Chapter 4. <i>Bacteroidales</i> Markers for Microbial Source Tracking in Southeast Asia	87
Abstract.....	87
4.1 Introduction	88
4.2 Materials and Methods	91
4.2.1 Study area.....	91
4.2.2 Fecal sample collection.....	91
4.2.3 Genomic DNA extraction and Inhibition removal.....	94
4.2.4 qPCR assays.....	95
4.2.4 Adjustment of qPCR conditions and verification of instrument compatibility	97
4.2. 6 Target limit of detection.....	99
4.2. 7 Evaluation of assay sensitivity and specificity	99
4.2.8 Statistical analysis	100
4. 3. Results	103
4.3.1 Prevalence of <i>Bacteroidales</i> markers in human and animal samples	103
4.3.1.1 Comparison of marker concentration in human stool and raw sewage samples.....	104
4.3.2 Performance of host-associated <i>Bacteroidales</i> assays	105
4.4. Discussion.....	109
4.4.1 Performance of Canine-associated and Universal <i>Bacteroidales</i> assays ..	109

4.4.2 Performance of Human-associated <i>Bacteroidales</i> Assays	110
4.5. Conclusions	114
References: Chapter 4.....	115
Chapter 5. Evaluation of digital droplet PCR (ddPCR) for improved Microbial Source Tracking	120
Abstract.....	120
5.1 Introduction	121
5.2. Methods	124
5.2.1. Sample selection and preparation	124
5.2.2 Design of B. thetaiotaomicron-specific alpha-1-6, mannanase (B. theta) plasmid DNA	126
5.2.3 qPCR and ddPCR assays	127
5.2. 4 ddPCR and qPCR comparison criteria.....	128
5.2.5 Statistical analysis	130
5.3 Results	134
5.3.1 Target standards	134
5.3.2 Human and animal fecal samples: Sensitivity, specificity and accuracy..	135
5.3.3 Composite and environmental samples.....	137
5.4 Discussion.....	144
5.4.1 Quantification of Standards by ddPCR and qPCR.....	144
5.4.2 Use of ddPCR for microbial sources tracking	145
5.5 Conclusions	148

References	150
Chapter 6. Decay of Human-associated <i>Bacteroidales</i> Viable Cells and DNA, and Fecal Indicator Bacteria in Tropical Freshwater Sediments.	157
Abstract.....	157
6.1 Introduction	158
6.2 Methodology.....	160
6.2.1 Experimental design.....	160
6.2.2 Enumeration of Fecal indicator bacteria	164
6.2.3 PMA treatment.....	164
6.2.4 DNA Extraction	165
6.2.5 Droplet digital PCR (ddPCR) assays	166
6.2.6 Microscopic enumeration of protozoa	167
6.2.7 Decay rate calculations and statistical analysis.....	168
6.3 Results	169
6.3.1 Physicochemical parameters	169
6.3.2 Human-associated <i>Bacteroidales</i> markers (B. theta and BacHum) demonstrated a faster decay rate in sediments	170
6.3.3 Decay profiles of fecal indicator bacteria (FIB) in sediments	175
6.3.4 The effect of protozoa on the decay of FIB and human-associated <i>Bacteroidales</i> markers	178
6.4 Discussion.....	182

6.4.1 Survival and persistence of human-associated <i>Bacteroidales</i> markers in sediments.....	182
6.4.2 Survival and persistence of fecal indicator bacteria in sediments	185
6.4.3 Decay models.....	187
6.5. Conclusions	188
References: Chapter 6.....	190
Chapter 7.	197
Summary and Future Research Considerations.....	197
7.1 Implications and Future research considerations.....	198
7.1.1 Microbial community Dynamics and water quality.....	198
7.1.2 <i>Bacteroidales</i> markers for microbial source tracking in Singapore and Southeast Asia.....	201
7.1. 3 Real-time PCR (qPCR) and Digital droplet PCR (ddPCR): Which tool is better for microbial source tracking	204
71.4 Survival and persistence of human-associated <i>Bacteroidales</i> cells and DNA in sediments	206
7.1.5 Outlook	210
References: Chapter 7	212
Appendix: Chapter 3	216
Appendix: Chapter 4	225
Appendix: Chapter 5	229

List of Table

Chapter 2

Table 1: Most common waterborne diseases.....	27

Chapter 3

Table 1. List of primers used for qPCR and library preparation for Illumina sequencing	57

Table 2. Distribution of bacterial OTUs and concentrations of water quality indicators in sampled land uses.....	63

Table 3. Contribution of <i>E. coli</i> or HF183 to variance in bacterial community structure, SAT; PLSg, and PLSs	67

Chapter 4

Table 1. Oligonucleotide primers and probes used for qPCR measurements of host-associated <i>Bacteroidales</i> gene markers.....	98

Table 2. Performance characteristics of tested <i>Bacteroidales</i> assays.....	101

Table 3. Performance of human- and canine-associated and universal <i>Bacteroidales</i> assays as revealed by qPCR quantification of fecal and raw sewage samples	102

Table 4. Sensitivity, specificity, and accuracy of universal, canine-associated and human-associated <i>Bacteroidales</i> assays in Singapore	105

Chapter 5

Table 1. Oligonucleotide primers and probes used for qPCR measurements of host-associated <i>Bacteroidales</i> gene markers	125

Table 2 . Composite samples scored positive against expected outcome determined based on previous qPCR of mixed samples.	129

Table 3. Comparison of qPCR and ddPCR limit of detections.	132

Table 4. Performance of human-associated <i>Bacteroidales</i> assays as revealed by qPCR and ddPCR quantification of human fecal samples and domesticated and wild animal fecal samples	136

Table 5. Human assays performance metrics calculated using BacHum and B.theta qPCR and ddPCR concentrations.....	142

Table 6. Ratios of qPCR to ddPCR predicted concentrations for B. theta and BacHum markers	143

Chapter 6

Table 1. Conditions and criteria for the design of sediments microcosms under anoxic conditions	163

Table 2. Chemical parameters measured in the microcosm using the miniDO ₂ T logger (Precision Measurement Engineering Inc., California, USA).....	169

Table 3. Kinetic parameters of fecal indicator bacteria, intact cells (PMA-Treated) and total DNA of human-associated <i>Bacteroidales</i> genetic markers (BacHum and B. theta) in sediment microcosms	172

Table 4. ANOVA comparison for 2-log removal (T ₉₉) in water column and sediments, and treatment and < 5µm control microcosms	175

Table 5. Protozoa Counts in Treatment and Three Controls (Log Protozoa counts/100ml).....	178

List of Figures

Chapter 2

Figure 1. Timeline (2000 – 2017) for the major milestone in microbial source tracking using <i>Bacteroidales</i> -associated markers.....	30

Chapter 3

Figure 1. Highly represented phyla across samples collected from the reservoir and catchment.	65
Figure 2. Multivariate analysis of bacterial community composition based on Principal Coordinate Analysis (PCO) of Bray-Curtis resemblance between samples..	68
Figure 3. Hierarchical clustering of the distribution of SAT (genera)..	72
Figure 4. Hierarchical clustering of the distribution of eighteen genera for which OTU sequences match human pathogens at the species level (16S rRNA identity $\geq 99\%$).	73
Figure 5. Draftsman plot of quantities considered in this study.....	74
Figure 6. The distribution of sequences from (A) sewage-associated taxa (SAT) and (B, C) pathogen-like sequences identified to genus and species level (PLSg and PLSs, respectively) as a percentage of the total bacterial community composition for different land-use categories sampled in this study..	75

Chapter 4

Figure 1. Concentrations of <i>Bacteroidales</i> -associated genetic markers in human stool and sewage samples.	107
Figure 2. Concentrations of <i>Bacteroidales</i> -associated 16S rRNA gene markers in domesticated animal fecal samples.	108
Figure 3. Concentrations of <i>Bacteroidales</i> -associated 16S rRNA gene markers in wild animal fecal samples.	108

Chapter 5

Figure 1. Concentrations of human- <i>Bacteroidales</i> markers (B. theta and BacHum) in human stool, sewage, and animal fecal samples as quantified using qPCR and Digital droplet PCR (ddPCR).....	133
Figure 2. Measured and expected concentrations of positive controls of human- <i>Bacteroidales</i> markers (B. theta and BacHum) quantified using qPCR and ddPCR, and computed using plasmid DNA of each marker.	138
Figure 3. Measured and expected concentrations of positive controls of human- <i>Bacteroidales</i> markers (B. theta and BacHum) quantified using ddPCR, and computed using plasmid DNA of each marker at concentrations below 100 copies per reaction..	139
Figure 4. Comparison of concentrations of human- <i>Bacteroidales</i> markers (BacHum and B. theta) in composite (3A and 3B) and environmental samples (3C and 3D). 140	

Chapter 6

Figure 1. Intact cells (PMA-ddPCR) concentrations over the period of 20 days as quantified using human markers B. theta (A and C) and BacHum (B and D).....	173
Figure 2: Total DNA concentrations over the period (intact cells and dead cells) of 20 days as quantified using human markers B. theta (A and C) and BacHum (B and D)..	174
Figure 3: Fecal indicator bacteria culture concentrations in treatment and control microcosms over time..	176
Figure 4: First order exponential decay ($N=N_0 \cdot e^{-kt}$) profiles for fecal indicator bacteria in the water column and sediment layers of tropical freshwater sediments microcosms under dark and anoxic conditions.....	177

Figure 5: First order exponential decay ($N=N_0 \cdot e^{-kt}$) profiles for PMA-treated and no-PMA treated concentrations of *B. theta* in water column and sediments layers of tropical freshwater microcosms under dark and anoxic conditions. 180

Figure 6: First order exponential decay ($N=N_0 \cdot e^{-kt}$) profiles for PMA-treated and no-PMA treated concentrations of *BacHum* in water column and sediments layers of tropical freshwater microcosms under dark and anoxic conditions. 181

List of Equations

Chapter 4

Eq. 1	100
Eq. 2	100
Eq. 3	100

Chapter 1: Introduction

1.1 Background

Freshwater quality management in urban environments is challenging as it relates to the need for health risk prevention and management of waterbodies used for primary or secondary recreational activities. These waterbodies are frequently located at the downstream portion of the urban watershed where the residuals of anthropogenic activities are conveyed through waterways carrying loads of microorganisms from soil/agricultural activities, leaking sewage (Sercu et al., 2009, Sauer et al., 2011), stormwater (Parker et al., 2010), runoff from concrete-lined drainages and paved urban surfaces (Sauer et al., 2011), and wildlife. Sewage can harbor pathogens that pose a threat to human health (Lin and Tong, 2011, Cai et al., 2014, Lin and Tong, 2013). Previous studies have also linked gastroenteritis illnesses to sewage-impaired beaches in the USA (Boehm et al., 2003, Cheung et al., 1990, Given et al., 2006). These studies relied on the correlation of fecal indicator bacteria (FIB) (*E. coli*, enterococci and total coliforms) used as proxies for sewage and pathogen contamination (USEPA, 1986, USEPA, 2002) with occurrence gastroenteritis illnesses among recreational water users during the period of study.

Fecal indicator bacteria (FIB) have been used extensively for water quality monitoring; however, FIB have been associated with performance limitations. They include: (1) short survival rate in waterbodies (McFeters et al., 1974); (2) identification of non-fecal strains (false positives) (Scott et al., 2002, Simpson et al., 2002); (3) not useful in disinfected water (Chern et al., 2014) ; (4) inability to discriminate human from animal

fecal sources (Field et al., 2003, Byappanahalli et al., 2012); (5) some cases of absence of correlation to the presence of pathogens (Horman et al., 2004, Stewart et al., 2008); and (6) growth after discharge in the water column (Desmarais et al., 2002, Solo-Gabriele et al., 2000). The latter reduced the FIB specificity in differentiating recent from aged human fecal contaminations (Field and Samadpour, 2007). Various research works on sub-tropical and tropical environments confirmed FIB limitations by demonstrating a wide environmental distribution of FIB in soils, sediments, beach sand, aquatic and terrestrial vegetation, and natural rivers, streams, and creeks (Hazen, 1988, Rivera, 1988, Hardina, 1991, Fujioka, 1999, Desmarais et al., 2002, Solo-Gabriele et al., 2000, Yamahara et al., 2007, Byappanahalli et al., 2012).

These FIB limitations led to the development of different alternative indicators associated with technological advancement in molecular biology. This progress was characterized by a shift from culture-based to DNA-based (non-culture-based) detection methods (Scott et al., 2002). Alternative indicators were introduced and included host-associated *Bacteroidales* (Bernhard and Field, 2000a, Field et al., 2003, Bernhard and Field, 2000b, Shanks et al., 2009b, Shanks et al., 2010, Kildare et al., 2007), F-specific RNA coliphages (Luther and Fujioka, 2004), and host-associated viruses (Gregory et al., 2006).

Molecular-based methods have addressed key FIB limitations such as source identification, discerning animal sources from human sources or other wild or domesticated animal sources, and differentiation of aged from recent contaminations (Bernhard and Field, 2000a, Bernhard and Field, 2000b, Shanks et al., 2009b, Shanks et al., 2010, Kildare et al., 2007, Bae and Wuertz, 2012, Bae and Wuertz, 2009a, Bae

and Wuertz, 2009b). However, there are challenges associated with the application of newly developed molecular-based methods. These limitations include (1) inability of host-associated *Bacteroidales* to achieve specificity and sensitivity of 100% in detecting the target without the use of a statistical conversion model (Harwood et al., 2014, Wang et al., 2010); (2) quantification with variability in data analysis such as determination of the limit of detection or limit of quantification (Stewart et al., 2013); (3) long-time persistence of DNA-signal compared to *Bacteroidales* cells or FIB in natural waters (Dick et al., 2010, Bae and Wuertz, 2012); (4) absence of geographic universality of host-associated *Bacteroidales* due to dietary differences and other reasons (Ley et al., 2008, Stewart et al., 2013); and (5) poor correlation with previous FIB levels and health risks (e.g., recreational water-related illnesses) (Santo Domingo et al., 2007).

Limitations of the newly developed molecular-based methods are being addressed and improved. The approach to ensuring efficient performance of proposed molecular-based tools include (1) validation of host-associated *Bacteroidales* before applying them at new regional levels (Gawler et al., 2007, Balleste and Blanch, 2010, Jenkins et al., 2009, Tambalo et al., 2012, Reischer et al., 2013, Ahmed et al., 2009, Odagiri et al., 2015); (2) resolution and standardization of method-based differences through cross-laboratory studies at regional levels (Boehm et al., 2013, Layton et al., 2013, Stewart et al., 2013); (3) use of more than one host-associated *Bacteroidales* during source tracking (Stewart et al., 2013); (4) use of bacterial community analysis to increase precision in source identification and pathogen detection (McLellan et al., 2010, Lin and Tong, 2011, Lin and Tong, 2013, Ibekwe et al., 2013, Cao et al., 2013, Cai et al.,

2014); and use of statistical analysis approaches to generate comparable data at regional scale (Wang et al., 2010).

The majority of identified limitations associated with FIB and molecular-based methods have been addressed in temperate climates where they have been applied for over a decade. However, tropical environments are also experiencing these challenges with most countries still using only FIB due to economic and practical reasons. Recently, there have been studies that applied molecular-based techniques in tropical countries such as Bangladesh (Ahmed et al., 2009, Knappett et al., 2011), India (Odagiri et al., 2015) , Tanzania (Pickering et al., 2011, Pickering et al., 2012, Mattioli et al., 2013), and Kenya (Jenkins et al., 2009). However, only two locations (Bangladesh and India) validated the chosen host-associated *Bacteroidales* methods before application and resulting in the selection of two different human-associated *Bacteroidales* markers in India (BacHum) (Odagiri et al., 2015) and Bangladesh (HF183) (Ahmed et al., 2009). Singapore is among the tropical countries using FIB (enterococci) and molecular-based methods for freshwater water quality monitoring (NTU, 2008, Chua et al., 2010); (Aw et al., 2009, Aw and Gin, 2010, Nshimiyimana et al., 2014, Rezaeinejad et al., 2014). However, a validation study of host-associated *Bacteroidales* methods has not yet been performed. Neither have factors affecting persistence and decay of host-associated *Bacteroidales* in sediments been thoroughly examined. Such gaps in addressing limitations related to host-associated *Bacteroidales* markers may reduce assay performance in new environments. The lack of a validation study in the tropical urban environment of Singapore and Southeast Asia has slowed the development of mitigation strategies by watershed managers, including the control of non-point fecal sources pollution.

Technical limitations are not the only challenges affecting the application of molecular-based tools for water quality monitoring and microbial source tracking in urban tropical environments. Lack of sufficient economic resources to establish state-of-the-art laboratories and training of relevant personnel are equally important. Currently, worldwide, only the USA has legally approved DNA-based enterococci (Enterococcus) techniques for water quality monitoring (USEPA, 2013). Other countries are still reviewing methods and seeking capital investments to prepare the use of these advanced techniques. Next generation sequencing, another costly but useful tool, can greatly improve bacterial community analysis and generate large data sets to better understand how microorganisms relate to water quality parameters. However, many urban tropical areas most affected by fecal pollution are associated with the poorer income levels of society. Apart from the cost, revision of standardized methods that will quantitatively link FIB, host-associated *Bacteroidales*, and bacterial community analysis are needed for future applications in source tracking.

1.2 Focus of the Dissertation

Detection and identification of non-point sources of fecal pollution in urban tropical environments is a challenge. Research progress in molecular microbiology led to the development of new molecular-based tools which improved cell detection and application of genetics in the context of fecal microbial pollution detection in natural environment (Scott et al., 2002, Harwood et al., 2005, McQuaig et al., 2006, Bernhard and Field, 2000a, Bernhard and Field, 2000b, Field et al., 2003, Layton et al., 2013, USEPA, 2012a, USEPA, 2012b, USEPA, 2013, Shanks et al., 2008, Shanks et al., 2009a, Shanks et al., 2010, Haugland et al., 2010, Green et al., 2011, Stewart et al.,

2013, Schriewer et al., 2013, Molina et al., 2014, Kildare et al., 2007, Odagiri et al., 2015).

Molecular-based tools were applied to address three research questions aiming at the improvement of water quality monitoring, fecal source identification and source identifier tracking in tropical urban environment.

(1) I applied next generation sequencing (NGS) to address the question whether bacterial community composition (BCC) varied as a function of sewage-associate taxa (SAT), pathogen-like bacteria (PLS) and water quality (Fecal Indicator Bacteria) in an urban tropical catchment.

(2) Molecular-based quantitative tools (quantitative PCR and digital droplet PCR) were used to address the question whether host-associated *Bacteroidales* 16S rRNA genetic markers could be used to discriminate human from animal fecal sources in the urban tropical environment.

(3) Molecular-based quantitative tools and laboratory-scale anoxic sediment microcosms were used to address the question whether protozoa had an effect on the decay of human-associated *Bacteroidales* markers (B. theta and BacHum) and whether these markers were appropriate for tracing sewage contamination in tropical freshwater. In addition, we evaluated whether sediments are a refuge/reservoir and potential source of *Bacteroidales* and fecal indicator bacteria to the overlying water.

The main goal of this research was to improve microbial source detection and tracking, and evaluate the relationship of fecal indicators and microbial community dynamics in tropical urban environment.

The main goal was achieved through three objectives addressing each of the research questions:

- 1 To examine the variation of bacterial community composition and subgroups (sewage-associated taxa (SAT) and pathogen-like sequences (PLS) as a function of water quality through next-generation 16S rRNA sequence analysis (V3-V4 regions) in a tropical urban environment in Singapore
- 2 To validate host-associated *Bacteroidales* genetic markers for microbial source tracking in Singapore and use of digital droplet PCR to improve depth of quantification of human-associated *Bacteroidales* markers
- 3 To evaluate sediment-associated decay rates of two human-associated microbial source tracking markers for *Bacteroidales*, B. theta and BacHum, and of the fecal indicator bacteria (FIB) *E. coli* and *Enterococcus* as well as study the role of protozoa in decay

Objectives were separated into three research phases where each topic was studied and reported in one or two of the four chapters (Chapter 3 to Chapter 6) of this dissertation.

Phase 1: Variation of microbial community composition (BCC) as a function of water quality

The first research phase examined the relations of BCC and subgroups (SAT and PLS) as a function of water quality (FIB (*E. coli*) and alternative indicator HF183/Bac287). Findings of this study are reported in Chapter 3.

Phase 2: Validation of host-*Bacteroidales* genetic markers for microbial source tracking

The second research phase focused on the combination of the field and laboratory work, and the use of molecular-quantification tools (qPCR and ddPCR) to (1) validate host-associated *Bacteroidales* genetic markers for microbial sources tracking in Singapore and Southeast Asia and (2) improve microbial source tracking and quantification depth. Findings of these experiments are reported in Chapter 4 and Chapter 5 of this dissertation.

Phase 3: Survival and decay of human-associated *Bacteroidales* genetic markers in sediments

The third research phase involved the use of microcosms and molecular-quantification tools (ddPCR) to assess sediment-associated decay rates of human-associated microbial source tracking markers for *Bacteroidales* (*B. theta* and *BacHum*) and of the fecal indicator bacteria (FIB) *E. coli* and *Enterococcus*, and to examine the role of protozoa in decay. Findings of this research work are reported in Chapter 6 of this dissertation.

Reference: Chapter 1

- AHMED, W., GOONETILLEKE, A., POWELL, D. & GARDNER, T. 2009. Evaluation of multiple sewage-associated *Bacteroides* PCR markers for sewage pollution tracking. *Water Research*, 43, 4872-7.
- AW, T. G. & GIN, K. Y. H. 2010. Environmental surveillance and molecular characterization of human enteric viruses in tropical urban wastewaters. *Applied and Environmental Microbiology*, 109, 716-730.
- AW, T. G., GIN, K. Y. H., EAN-OON, L. L., CHEN, E. X. & WOO, C. H. 2009. Prevalence and Genotypes of Human Noroviruses in Tropical Urban Surface Waters and Clinical Samples in Singapore. *Applied and Environmental Microbiology*, 75, 4984-4992.
- BAE, S. & WUERTZ, S. 2009a. Discrimination of viable and dead fecal *Bacteroidales* bacteria by quantitative PCR with propidium monoazide. *Applied and Environmental Microbiology*, 75, 2940-4.
- BAE, S. & WUERTZ, S. 2009b. Rapid decay of host-specific fecal *Bacteroidales* cells in seawater as measured by quantitative PCR with propidium monoazide. *Water Research*, 43, 4850-9.
- BAE, S. & WUERTZ, S. 2012. Survival of host-associated *Bacteroidales* cells and their relationship with *Enterococcus spp.*, *Campylobacter jejuni*, *Salmonella enterica* serovar *Typhimurium*, and adenovirus in freshwater microcosms as measured by propidium monoazide-quantitative PCR. *Applied and Environmental Microbiology*, 78, 922-32.
- BALLESTE, E. & BLANCH, A. R. 2010. Persistence of *Bacteroides* Species Populations in a River as Measured by Molecular and Culture Techniques. *Applied and Environmental Microbiology*, 76, 7608-7616.
- BERNHARD, A. E. & 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, 1587-1594.
- 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, 4571-4574.
- BOEHM, A., VAN DE WERFHORST, L. C., GRIFFITH, J., HOLDEN, P., JAY, J., SHANKS, O., WANG, D. & WEISBERG, S. 2013. Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. *Water Research*, 47, 6812-6828.

- BOEHM, A. B., FUHRMAN, J. A., MRSE, R. D. & GRANT, S. B. 2003. Tiered approach for identification of a human fecal pollution source at a recreational beach: case study at Avalon Bay, Catalina Island, California. *Environmental Science and Technology*, 37, 673-80.
- BYAPPANAHALLI, M. N., NEVERS, M. B., KORAJKIC, A., STALEY, Z. R. & HARWOOD, V. J. 2012. *Enterococci* in the environment. *Microbiology and Molecular Biology Reviews*, 76, 685-706.
- CAI, L., JU, F. & ZHANG, T. 2014. Tracking human sewage microbiome in a municipal wastewater treatment plant. *Applied Microbiology and Biotechnology*, 98, 3317-3326.
- CAO, Y., VAN DE WERFHORST, L. C., DUBINSKY, E., BADGLEY, B., SADOWSKY, M., GRIFFITH, J. F., ANDERSEN, G. & HOLDEN, P. 2013. Evaluation of molecular community analysis methods for discerning fecal sources and human waste. *Water Research*, 47, 6862-72.
- CHERN, E. C., BRENNER, K., WYMER, L. & HAUGLAND, R. A. 2014. Influence of wastewater disinfection on densities of culturable fecal indicator bacteria and genetic markers. *Water Health*, 12, 410-7.
- CHEUNG, W. H., CHANG, K. C., HUNG, R. P. & KLEEVENS, J. W. 1990. Health effects of beach water pollution in Hong Kong. *Epidemiology and Infection*, 105, 139-162.
- CHUA, L. H. C., SHANAHAN, P., EDMOND, Y. M. L., SHUY, E. B., THOMPSON, J., DIXON, C. C., KERIGAN, K. K., NSHIMYIMANA, J. P., YEAGER, J. M., LEE, L. J. & POR, Y. L. 2010. Dry Weather bacteria monitoring and variation with land use for Kranji Reservoir catchment, Singapore. *17th Congress of the Asia and Pacific Division of the International Association of Hydraulic Engineering and Research*. Auckland, New Zealand.
- DESMARAIS, T. R., SOLO-GABRIELE, H. M. & PALMER, C. J. 2002. Influence of soil on fecal indicator organisms in a tidally influenced subtropical environment. *Applied and Environmental Microbiology*, 68, 1165-72.
- DICK, L., STELZER, E., BERTKE, E., FONG, D. & STOECKEL, D. 2010. Relative decay of *Bacteroidales* microbial source tracking markers and cultivated *Escherichia coli* in freshwater microcosms. *Applied and Environmental Microbiology*, 76, 3255-62.
- FIELD, K., BERNHARD, A. & BRODEUR, T. 2003. Molecular approaches to microbiological monitoring: Fecal source detection. *Environmental monitoring and assessment*, 81, 313-326.

- FIELD, K. G. & SAMADPOUR, M. 2007. Fecal source tracking, the indicator paradigm, and managing water quality. *Water Research*, 41, 3517-38.
- FUJIOKA, R. 1999. Soil: The environmental source of *Escherichia coli* and *Enterococci* in Guam's streams. *Applied Microbiology Symposium Supplement*, 85, 83S-89S.
- GAWLER, A. H., BEECHER, J. E., BRANDÃO, J., CARROLL, N. M., FALCÃO, L., GOURMELON, M., MASTERSON, B., NUNES, B., PORTER, J. & RINCÉ, A. 2007. Validation of host-specific *Bacteroidales* 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. *Water Research*, 41, 3780.
- GIVEN, S., PENDLETON, L. H. & BOEHM, A. B. 2006. Regional public health cost estimates of contaminated coastal waters: a case study of gastroenteritis at southern California beaches. *Environmental Science and Technology*, 40, 4851-8.
- GREEN, H., SHANKS, O., SIVAGANESAN, M., HAUGLAND, R. & FIELD, K. 2011. Differential decay of human faecal *Bacteroides* in marine and freshwater. *Environmental Microbiology*, 13, 3235-49.
- GREGORY, J. B., LITAKER, R. W. & NOBLE, R. T. 2006. Rapid One-Step Quantitative Reverse Transcriptase PCR Assay with Competitive Internal Positive Control for Detection of Enteroviruses in Environmental Samples. *Applied and Environmental Microbiology*, 72, 3960-3967.
- HARDINA, C. M. 1991. Soil: The environmental source of *Escherichia coli* and *Enterococci* in Hawaii's streams. *Environmental Toxicology and Water Quality*, 6, 185-195.
- HARWOOD, V., STALEY, C., BADGLEY, B., BORGES, K. & KORAJKIC, A. 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews*, 38, 1-40.
- HARWOOD, V. J., LEVINE, A. D., SCOTT, T. M., CHIVUKULA, V., LUKASIK, J., FARRAH, S. R. & ROSE, J. B. 2005. Validity of the indicator organism paradigm for pathogen reduction in reclaimed water and public health protection. *Applied and Environmental Microbiology*, 71, 3163-70.
- HAUGLAND, R. A., VARMA, M., SIVAGANESAN, M., KELTY, C., PEED, L. & SHANKS, O. C. 2010. Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected *Bacteroidales* species and human fecal waste by qPCR. *Systematic and Applied Microbiology*, 33, 348-57.

- HAZEN, T. C. 1988. Fecal coliforms as indicators in tropical waters: A review. *Toxicity Assessment*, 3, 461-477.
- HORMAN, A., RIMHANEN-FINNE, R., MAUNULA, L., VON BONSDORFF, C. H., TORVELA, N., HEIKINHEIMO, A. & HANNINEN, M. L. 2004. *Campylobacter spp.*, *Giardia spp.*, *Cryptosporidium spp.*, noroviruses, and indicator organisms in surface water in southwestern Finland, 2000-2001. *Applied and Environmental Microbiology*, 70, 87-95.
- IBEKWE, A. M., LEDDY, M., MURINDA, S. & AZIZ, R. 2013. Potential Human Pathogenic Bacteria in a Mixed Urban Watershed as Revealed by Pyrosequencing. *PLoS ONE*, 8, e79490.
- JENKINS, M. W., TIWARI, S., LORENTE, M., GICHABA, C. M. & WUERTZ, S. 2009. Identifying human and livestock sources of fecal contamination in Kenya with host-specific *Bacteroidales* assays. *Water Research*, 43, 4956-66.
- KILDARE, B., LEUTENEGGER, C., MCSWAIN, B., BAMBIC, D., RAJAL, V. & WUERTZ, S. 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: A Bayesian approach. *Water Research*, 41, 15.
- KNAPPETT, P. S., ESCAMILLA, V., LAYTON, A., MCKAY, L. D., EMCH, M., WILLIAMS, D. E., HUQ, R., ALAM, J., FARHANA, L., MAILLOUX, B. J., FERGUSON, A., SAYLER, G. S., AHMED, K. M. & VAN GEEN, A. 2011. Impact of population and latrines on fecal contamination of ponds in rural Bangladesh. *Science of the Total Environment*, 409, 3174-82.
- LAYTON, B., CAO, Y., EBENTIER, D., HANLEY, K., BALLESTÉ, E., BRANDÃO, J., BYAPPANAHALLI, M., CONVERSE, R., FARNLEITNER, A., GENTRY SHIELDS, J., GIDLEY, M., GOURMELON, M., LEE, C., LEE, J., LOZACH, S., MADI, T., MEIJER, W., NOBLE, R., PEED, L., REISCHER, G., RODRIGUES, R., ROSE, J., SCHRIEWER, A., SINIGALLIANO, C., SRINIVASAN, S., STEWART, J., VAN DE WERFHORST, L. C., WANG, D., WHITMAN, R., WUERTZ, S., JAY, J., HOLDEN, P., BOEHM, A., SHANKS, O. & GRIFFITH, J. 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. *Water Research*, 47, 6897-6908.
- LEY, R. E., HAMADY, M., LOZUPONE, C., TURNBAUGH, P. J., RAMEY, R. R., BIRCHER, J. S., SCHLEGEL, M. L., TUCKER, T. A., SCHRENZEL, M. D., KNIGHT, R. & GORDON, J. I. 2008. Evolution of Mammals and Their Gut Microbes. *Science*, 320, 1647-1651.
- LIN, C. & TONG, Z. 2013. Detecting Human Bacterial Pathogens in Wastewater Treatment Plants by a High-Throughput Shotgun Sequencing Technique. *Environmental Science and Technology*, 47, 5433-5441.

- LIN, Y. & TONG, Z. 2011. Pathogenic Bacteria in Sewage Treatment Plants as Revealed by 454 Pyrosequencing. *Environmental Science and Technology*, 45, 7173-7179.
- LUTHER, K. & FUJIOKA, R. 2004. Usefulness of monitoring tropical streams for male-specific RNA coliphages. *Water and Health*, 2, 171-181.
- MATTIOLI, M. C., PICKERING, A. J., GILSDORF, R. J., DAVIS, J. & BOEHM, A. B. 2013. Hands and water as vectors of diarrheal pathogens in Bagamoyo, Tanzania. *Environmental Science and Technology*, 47, 355-63.
- MCFETERS, G. A., BISSONNETTE, G. K., JEZESKI, J. J., THOMSON, C. A. & STUART, D. G. 1974. Comparative survival of indicator bacteria and enteric pathogens in well water. *Applied Microbiology* 27, 823-9.
- MCLELLAN, S. L., HUSE, S. M., MUELLER-SPITZ, S. R., ANDREISHCHEVA, E. N. & SOGIN, M. L. 2010. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environmental Microbiology*, 12, 378-392.
- MCQUAIG, S. M., SCOTT, T. M., HARWOOD, V. J., FARRAH, S. R. & LUKASIK, J. O. 2006. Detection of Human-Derived Fecal Pollution in Environmental Waters by Use of a PCR-Based Human Polyomavirus Assay. *Applied and Environmental Microbiology*, 72, 7567-7574.
- MOLINA, M., HUNTER, S., CYTERSKI, M., PEED, L. A., KELTY, C. A., SIVAGANESAN, M., MOONEY, T., PRIETO, L. & SHANKS, O. C. 2014. Factors affecting the presence of human-associated and fecal indicator real-time quantitative PCR genetic markers in urban-impacted recreational beaches. *Water Research*, 64, 196-208.
- NSHIMYIMANA, J. P., EKKLESIA, E., SHANAHAN, P., CHUA, L. H. C. & THOMPSON, J. R. 2014. Distribution and abundance of human-specific *Bacteroides* and relation to traditional indicators in an urban tropical catchment. *Applied Microbiology*, 116, 1369-1383.
- NTU 2008. Water quality monitoring, modeling and management for Kranji Catchment/Reservoir system – Phases 1 and 2 / May 2004 to December 2007. Singapore: Division of Environmental and Water Resources Engineering, School of Civil and Environmental Engineering, Nanyang Technological University.
- ODAGIRI, M., SCHRIEWER, A., HANLEY, K., WUERTZ, S., MISRA, P. R., PANIGRAHI, P. & JENKINS, M. W. 2015. Validation of *Bacteroidales* quantitative PCR assays targeting human and animal fecal contamination in the public and domestic domains in India. *Science of the Total Environment*, 502, 462-470.

- PARKER, J. K., MCINTYRE, D. & NOBLE, R. T. 2010. Characterizing fecal contamination in stormwater runoff in coastal North Carolina, USA. *Water Research*, 44, 4186-4194.
- PICKERING, A. J., JULIAN, T. R., MAMUYA, S., BOEHM, A. B. & DAVIS, J. 2011. Bacterial hand contamination among Tanzanian mothers varies temporally and following household activities. *Tropical Medicine & International Health*, 16, 233-9.
- PICKERING, A. J., JULIAN, T. R., MARKS, S. J., MATTIOLI, M. C., BOEHM, A. B., SCHWAB, K. J. & DAVIS, J. 2012. Fecal contamination and diarrheal pathogens on surfaces and in soils among Tanzanian households with and without improved sanitation. *Environmental Science and Technology*, 46, 5736-43.
- REISCHER, G. H., EBDON, J. E., BAUER, J. M., SCHUSTER, N., AHMED, W., ASTROM, J., BLANCH, A. R., BLOSCHL, G., BYAMUKAMA, D., COAKLEY, T., FERGUSON, C., GOSHU, G., KO, G., DE RODA HUSMAN, A. M., MUSHI, D., POMA, R., PRADHAN, B., RAJAL, V., SCHADE, M. A., SOMMER, R., TAYLOR, H., TOTH, E. M., VRAJMASU, V., WUERTZ, S., MACH, R. L. & FARNLEITNER, A. H. 2013. Performance characteristics of qPCR assays targeting human- and ruminant-associated bacteroidetes for microbial source tracking across sixteen countries on six continents. *Environmental Science and Technology*, 47, 8548-56.
- REZAEINEJAD, S., VERGARA, G. G., WOO, C. H., LIM, T. T., SOBSEY, M. D. & GIN, K. Y. 2014. Surveillance of enteric viruses and coliphages in a tropical urban catchment. *Water Research*, 58, 122-31.
- RIVERA, S. C. 1988. Isolation of fecal coliforms from pristine sites in a tropical rain forest. *Applied and Environmental Microbiology*, 54, 513-517.
- SANTO DOMINGO, J. W., BAMBIC, D. G., EDGE, T. A. & WUERTZ, S. 2007. Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. *Water Research*, 41, 3539-52.
- SAUER, E., VANDEWALLE, J., BOOTSMA, M. & MCLELLAN, S. 2011. Detection of the human specific *Bacteroides* genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. *Water Research*, 45, 4081-4091.
- SCHRIEWER, A., GOODWIN, K. D., SINIGALLIANO, C. D., COX, A. M., WANLESS, D., BARTKOWIAK, J., EBENTIER, D. L., HANLEY, K. T., ERVIN, J., DEERING, L. A., SHANKS, O. C., PEED, L. A., MEIJER, W. G., GRIFFITH, J. F., SANTODOMINGO, J., JAY, J. A., HOLDEN, P. A. & WUERTZ, S. 2013. Performance evaluation of canine-associated *Bacteroidales* assays in a multi-laboratory comparison study. *Water Research*, 47, 6909-20.

- SCOTT, T., ROSE, J., JENKINS, T., FARRAH, S. & LUKASIK, J. 2002. Microbial source tracking: current methodology and future directions. *Applied and Environmental Microbiology*, 68, 5796-803.
- SERCU, B., VAN DE WERFHORST, L., MURRAY, J. & HOLDEN, P. 2009. Storm Drains are Sources of Human Fecal Pollution during Dry Weather in Three Urban Southern California Watersheds. *Environmental Science and Technology*, 43, 293-298.
- SHANKS, O., ATIKOVIC, E., BLACKWOOD, A. D., LU, J., NOBLE, R., DOMINGO, J., SEIFRING, S., SIVAGANESAN, M. & HAUGLAND, R. 2008. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology*, 74, 745-52.
- SHANKS, O., KELTY, C., SIVAGANESAN, M., VARMA, M. & HAUGLAND, R. 2009a. Quantitative PCR for genetic markers of human fecal pollution. *Applied and Environmental Microbiology*, 75, 5507-13.
- SHANKS, O., WHITE, K., KELTY, C., SIVAGANESAN, M., BLANNON, J., MECKES, M., VARMA, M. & HAUGLAND, R. 2010. Performance of PCR-based assays targeting *Bacteroidales* genetic markers of human fecal pollution in sewage and fecal samples. *Environmental Science and Technology*, 44, 6281-8.
- SHANKS, O. C., KELTY, C. A., SIVAGANESAN, M., VARMA, M. & HAUGLAND, R. A. 2009b. Quantitative PCR for genetic markers of human fecal pollution. *Applied And Environmental Microbiology*, 75, 5507-13.
- SIMPSON, J., SANTO DOMINGO, J. & REASONER, D. 2002. Microbial source tracking: state of the science. *Environmental Science and Technology*, 36, 5279-88.
- SOLO-GABRIELE, H. M., WOLFERT, M. A., DESMARAIS, T. R. & PALMER, C. J. 2000. Sources of *Escherichia coli* in a coastal subtropical environment. *Applied and Environmental Microbiology*, 66, 230-7.
- STEWART, J., BOEHM, A., DUBINSKY, E., FONG, T.-T., GOODWIN, K., GRIFFITH, J., NOBLE, R., SHANKS, O., VIJAYAVEL, K. & WEISBERG, S. 2013. Recommendations following a multi-laboratory comparison of microbial source tracking methods. *Water Research*, 47, 6829-38.
- STEWART, J., GAST, R., FUJIOKA, R., SOLO GABRIELE, H., MESCHKE, J. S., AMARAL ZETTLER, L., DEL CASTILLO, E., POLZ, M., COLLIER, T., STROM, M., SINIGALLIANO, C., MOELLER, P. & HOLLAND, A. F. 2008. The coastal environment and human health: microbial indicators, pathogens, sentinels and reservoirs. *Environmental Health*, 7, S3.

- TAMBALO, D., FREMAUX, B., BOA, T. & YOST, C. 2012. Persistence of host-associated *Bacteroidales* gene markers and their quantitative detection in an urban and agricultural mixed prairie watershed. *Water Research*, 46, 2891-904.
- USEPA 1986. Ambient water quality criteria for bacteria--1986.
- USEPA 2002. EPA approves latest edition of water quality testing guide. (Forum). *Plant Engineering*, 56, 9.
- USEPA 2012a. Method 1611: *Enterococci* in Water by TaqMan Quantitative Polymerase Chain Reaction (qPCR) Assay. Washington DC: USEPA.
- USEPA 2012b. Recreational Water Quality Criteria. Washington DC: USEPA.
- USEPA 2013. Method 1609: *Enterococci* in water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) with Internal Amplification control (IAC) assay. In: AGENCY, U. S. E. P. (ed.). Washington: U.S. Environmental Protection Agency.
- WANG, D., SILKIE, S. S., NELSON, K. L. & WUERTZ, S. 2010. Estimating true human and animal host source contribution in quantitative microbial source tracking using the Monte Carlo method. *Water Research*, 44, 4760-75.
- YAMAHARA, K., LAYTON, B., SANTORO, A. & BOEHM, A. 2007. Beach sands along the California coast are diffuse sources of fecal bacteria to coastal waters. *Environmental Science and Technology*, 41, 4515-21.

Chapter 2. Literature Review

2.1 Microbial water pollution

Microorganisms include bacteria, viruses and protozoa and may contribute to the degradation of water quality. They exist in soil, water and air habitats as well as humans and animals and have been found to cause diseases in humans (Aw and Rose, 2012), fish (Ramírez-Toro et al., 2013), and other animals (Crump et al., 2002). When microbial water pollution is from a known source, e.g., a wastewater treatment plant, it is recognized as “point-source pollution” (Jamwal et al., 2011) and it is “non-point source pollution” when it is from an unknown source, e.g., stormwater (Abdelzaher et al., 2010).

2.2 Microbial water pollution in urban tropical waterways

2.2.1 Urban tropical waterways

Water availability, fertile soils, warm temperatures and limited seasonal variability have made the tropical region attractive to humans and many varieties of fauna and flora (Wohl et al., 2012). Currently the tropics hold most of the world’s human population. Southeast Asia is one of the most populated area in the tropical regions with 8% of the world population and over 48% of world urban population (Jones, 2013). The natural environment of this tropical zone is threatened by high population density and rapid urban development. Urban human development modifications are frequent due to the mismatch between population growth and available resources. The most notable anthropogenic activities include the disruption and destruction of land-cover

vegetation, deforestation, and introduction of ornamental trees (Giambelluca et al., 2003, Ziegler et al., 2007, Chappell et al., 2004, Walsh et al., 2011). All these changes have affected regional hydrology and caused increases in river flow, reduced the role of the forests in evapotranspiration, and led to changes in rainfall events and seasonality (Negri, 2004, Hayhoe et al., 2011).

The impact of anthropogenic activities on an urban watershed can be examined under two major categories: (1) the immediate impact on streams and (2) indirect impacts on receiving waters. First, the immediate impact of anthropogenic activities to urban waterways has been associated with the deterioration of water quality based on physical and biological characterization. The rapid degradation of water quality at the level of urban streams has often been attributed to the expansion of the impervious area within the watershed. Second, the indirect impact of urbanization on receiving waters was demonstrated by examining stormwater runoff characteristics (CWP, 2003). The runoff was found to affect the status of physical and biological indicators within the receiving waters due to the high volume and wide variety of pollutant loads across the urban watershed. The effect of drastic water quality changes in receiving waters such as urban estuaries is alarming. These estuaries are impacted primarily by nitrogen levels, fecal coliform bacteria, heavy metals, and hydrocarbons found mainly in the sediments (Vernberg, 1992, Kucklick et al., 1997, Sanger et al., 1999, Lerberg, 2000). The decline in water quality in urban tropical waterways is alarming and poses health risks to humans using impaired recreational waterbodies or beaches.

2.2.2 Microbial water pollutions in urban waterways

Water quality impairment of urban waterways is mainly caused by microbial water pollution loads and physical alterations. The physical alterations are a result of an ever increasing urban population (Tsatsaros et al., 2013) that triggers the intensification of infrastructure development and landscaping. These intensive alteration activities extend impervious surfaces that contribute to flooding and deteriorations of urban waterways, reservoirs and estuaries (CWP, 2003, Tsatsaros et al., 2013). Microbial water pollution loads related to flooding or stormwater runoff are among the major causes of water quality deterioration at urban beaches and in waterbodies worldwide (Prüss, 1998, Horman et al., 2004, Noble et al., 2006, Haile et al., 1999). Apart from microbial sources reported in Chapter 1, other major sources of microbial pollution in urban areas include humans, domestic pets, and wildlife.

Even with the progress made in infrastructure development, microbiology and biotechnology, the magnitude of problems associated with microbial water pollution in urban areas is still increasing. This is reflected through economic losses, the number of impaired recreational beaches and the incidence of gastroenteritis diseases among recreational swimmers (Haile et al., 1999, Noble et al., 2006). For example, the economic impact of fecally contaminated coastal waters is estimated as an annual financial loss of \$51 million dollars in Orange and Los Angeles counties in Southern California, USA (Given et al., 2006).

2.2.3 Microbial water pollution in Southeast Asia

Southeast Asia is divided into two geographic regions: Indochina and Maritime Southeast Asia. The Indochina region includes five countries: Cambodia, Laos, Myanmar, Thailand, and Vietnam. Maritime Southeast Asia includes seven countries: Brunei, Malaysia, East Timor, Indonesia, Philippines, and Singapore (UN 2013). These countries are undergoing development led by different economy and poverty level trends. Singapore and Brunei are the most advanced economically and have provided basic urban water sanitary infrastructure that include flood and sewage management systems. The remaining countries struggle to provide basic sanitary infrastructure and access to clean drinking water. Poverty in rural and urban slums combined with poor sanitary practices such as open defecation have led to surface and groundwater microbial pollution in these countries (Rajgire, 2013).

Generally, the urban areas of the majority of Southeast Asia countries, with the exception of Singapore and Brunei, are under high pressure to remediate environmental sanitation and improve water quality. Despite the commitment toward improvement, leaking sewer pipes, latrines, and discharge of untreated water in urban waterways are still common sources of urban microbial water pollution. Currently, estimates indicate that at least 13 million tons of human feces, 122 million m³ of urine, and 13 billion m³ of sewage are discharged in waterways of Southeast Asia and India (Rajgire, 2013). For example, the urban areas of the islands of Indonesia are among the areas most affected by heavy microbial pollution of water sources. Urban and rural areas of the island of Java, one of the most populated islands in Indonesia, have shallow groundwater heavily contaminated by human feces. Such poor sanitary practices contribute to the frequency

of water-related infectious diseases, especially in Indonesia and Vietnam (Willetts et al., 2009).

Microbial water pollution challenges for most of the Southeast Asian countries are not only related to the practice of open defecation and poor sanitary systems, but also to the capacity of monitoring and identifying point and non-point microbial water pollution sources. Few water quality studies have been carried out in this region and they have demonstrated an alarming freshwater deterioration to the point of suggesting that some waterways cannot be used for any primary recreational activities or be even treated by conventional techniques for clean water supply. The water quality case study of Bangpakong River in Eastern Thailand demonstrated that wet weather ameliorated the water quality, while dry weather deteriorated it (Bordalo, 2001). Such a phenomenon could be associated with dilution of pollution in the river by heavy rainfall during the wet season and an increase in the accumulation of contaminants during the dry seasons. The conclusions of this study suggested that the Bangpakong River could only be treated by advanced technologies for drinking water purposes and used exclusively for secondary recreational activities during the wet season (Bordalo, 2001). Microbial water pollution is still challenging in Southeast Asia. However, different water quality management, monitoring and microbial source tracking solutions are being pioneered in Singapore and they could have an impact at the local and regional scale.

2.2.4 Microbial water pollution in Singapore

Singapore is known worldwide for operating a highly integrated water resource management system. Singapore, at 1.5° north of the Equator, experiences a humid tropical climate with rainfall estimated at 2400 mm per year (NEA, 2009). Such heavy rainfall events in an urban area with a high percentage of impervious space and concrete-lined drainages require an efficient flood management strategy. Apart from increasing the size of drainages in flood prone areas, Singapore is also initiating the use of detention and retention basins in residential and some recreational park areas. Detention basins are designed to temporarily keep stormwater after rainfall, while retention basins are used for holding water for long periods of time and assisting in the management of the extra quantity of runoff (PUB, 2011). The integrated management of stormwater is not the only water management strategy appreciated in Singapore. The Active, Beautiful, Clean (ABC) program of the Public Utility Board (PUB) is also among one of the special water resources management strategies. This program brings together local communities and authorities with a common goal of improving freshwater quality and recreational use. Currently, the ABC program is running approximately 23 projects that promote the recreational use of water reservoirs, canals and areas surrounding the waterbodies. The promoted recreational areas include Bedok Reservoir, MacRitchie Reservoir, Marina Bay Reservoir, and Kallang River (PUB, 2011).

Despite the efforts of implementing wide water resources management programs, various water quality research projects, and building large concrete-lined drainages for flood management, Singapore still experiences challenges associated with efficiently

tracking and monitoring microbial water pollution. The microbial water pollution baseline study of Singapore's waterways, including the Kranji Reservoir, demonstrated high levels of FIB in waterways of different land use categories such as residential areas and the aquaculture/farming areas of Kranji catchment (NTU, 2008, Chua et al., 2010).

Recent studies have focused on monitoring microbial water pollution around Singapore. Nshimiyimana and colleagues (2014) carried out a study in a mixed-land use watershed of Kranji catchment in the northwest of Singapore that demonstrated high levels of FIB and the human-associated marker HF183 in the aquaculture areas. In addition, the study showed that potential sources of microbial water pollution in this area were wastewater septic tanks at different farms (Chicken, fish, and vegetables). Other microbial water pollution studies in Singapore assessed the abundance of human enteric viruses in wastewaters (Aw and Gin, 2010) and the prevalence and diversity of waterborne pathogenic viruses in freshwater (Aw et al., 2009). The latter study demonstrated the presence of adenoviruses, astroviruses and both norovirus genogroups I (GI) and II (GII) in all collected raw sewage samples and the effluent of secondary treatment. They also showed that detected enteric viruses groups were of a notable diversity suggesting the need for a comprehensive public health risk study. After these observations in raw sewage and secondary treatment effluent, Aw and Gin (2011) continued their study on Singapore freshwater sources. The findings demonstrated that human enteric viruses were common in 84 freshwater samples collected (83.3% detection) (Aw and Gin, 2011). Although enteric human viruses were found to be diverse and abundant, it is not clear what health risks they represent in Singapore's environment compared to other pathogens associated with human feces.

All these studies demonstrated that microbial water pollution is still problematic and that there is a need to validate indicators for microbial source tracking in Singapore. More studies, such as those exploring persistence, potential growth, and decay rates of all currently used bacteria and pathogen alternative indicators are needed. In addition, there is a need to understand the interactions with microbial communities in tropical waterways.

2.3 Human health risks related to pathogens harbored by microbial pollution

Microbial water pollution may refer to pathogens of multiple origins, but only those from human feces and other animals such as livestock and pets are associated with human health risks (Ashbolt et al., 2001, Dorevitch et al., 2010). Studies have shown that sewage contains multiple types of pathogens (McLellan et al., 2010, Lin and Tong, 2013, Cai et al., 2014) that are capable of causing various human diseases commonly known as waterborne diseases (Feachem et al., 1983). However, since all pathogens are not living freely in the water and can be symbiotically involved with other aquatic organisms, different disease groups are used to express different types of water related diseases. These groups are waterborne diseases (e.g., vibrio cholera), water-based diseases; water-washed diseases; and water-related diseases (Feachem et al., 1977, Feachem et al., 1983, White et al., 1972). Waterborne diseases (Table 1) include cholera; typhoid; amoebic and bacillary dysentery; and diarrhea, which can be caused by *Cryptosporidium*, *Giardia*, *Shigella*, *E. coli*-O157:H7 and norovirus (Hlavsa et al., 2011). Waterborne diseases also include diseases caused by chemical pollutants such as arsenic; fluoride; nitrates from fertilizers; and carcinogenic pesticides (Kjellstrom, 1986, Khan et al., 1997). Water-based diseases are those caused by parasites living

symbiotically or as intermediate organisms in contaminated water such as Schistosomiasis and Dracunculiasis. Water-washed diseases are those caused by poor personal hygiene and skin and eye contact with fecally contaminated water such as scabies, trachoma, typhus, lice and others. Water-related diseases are those caused by insect vectors, especially mosquitoes that breed or feed in contaminated water. The public using water for recreational activities has to be made aware of all these groups of water related diseases for human health risk control. For the purpose of this research project, this section will only focus on the problematic of waterborne diseases directly associated with sewage and which can cause illness through bathing, swimming and other recreational activities or consumption of seafood contaminated by sewage pathogens.

The findings of John Snow, an English physician of 1800s, led to linking a cholera outbreak to sewage leakage in Soho, London in 1854, thus increasing awareness of the role of water fecal contamination in the spread of waterborne diseases (Summers, 1989). In addition, strategies of human health protection were established. These include water quality management, which encompassed the detection of microorganisms of human gastrointestinal flora and that of animal feces (e.g. coliforms and Enterococci) (Ashbolt et al., 2001). Currently, methods of detection of pathogens in water have evolved to include proxies that are pathogen related (Moore et al., 1993) or even to analyze the entire microbial community to understand their interactions with pathogens (McLellan et al., 2010, Ibekwe et al., 2013, Aw and Rose, 2012). This evolution has been triggered by the increase in global disasters and the need to provide early warnings that could reduce human health risks (Aw and Rose, 2012).

Despite the progress in the development of water pathogen detection for human health risk control, waterborne diseases are still common. Their prevalence could be associated with different factors that may include reduced pace in standardizing pathogen detection techniques. Current sequencing-based molecular tools show promising future applications in rapid simultaneous detection of pathogen. However, advanced knowledge for operating them is still a requirement and the technology is improving. More efforts in application and optimization are still needed to reach standardization or full integration with the currently used protocols (Aw and Rose, 2012).

Table 1. Most common waterborne diseases (Hlavsa et al., 2011, Aw and Rose, 2012)

Microbes	Diseases	Pathogens	Symptoms	Causes	Incubation period
Bacteria	Campylobacteriosis	<i>Campylobacter jejuni</i>	Diarrhea or dysentery syndrome, mostly including cramps, fever and pain	Fecal contaminated water and food (e.g. chicken and unpasteurized milk)	2 to 10 days
	Cholera	<i>Vibrio cholerae</i>	water diarrhea, vomiting, and leg cramps	contaminated water, rivers and coastal waters	2 hours to 5 days
	<i>E coli</i> 0157:H7	<i>Escherichia coli</i>	Diarrhea (may be bloody), abdominal pain, nausea, vomiting, fever and hemolytic-uremic syndrome	Fecal contaminated water and food (e.g. undercooked group beef, cheeses, unpasteurized milk or juice, cider alfalfa sprouts)	1 to 8 days
	Legionellosis	<i>Legionella pneumophila</i>	Fever, chills, pneumonia, anorexia, muscle aches, diarrhea and vomiting	Fecal contaminated water	2 to 10 days
	Salmonellosis	<i>Salmonella enterica</i>	Abdominal pain, headache, fever, nausea, diarrhea, chills and cramps	Fecal contaminated water and food (e.g. poultry, eggs, meat and meat products, milk, smoked fish, protein food and juices)	1 to 3 days
	Vibrio infection	<i>Vibrio parahaemolyticus</i> , <i>Vibrio vulnificus</i>	Nausea, vomiting, headache and also dysentery-like symptoms in some patients	Fecal contaminated water and seafood (e.g., raw shellfish and oysters)	1 to 7 or more days

Table 1: Summary of the most common waterborne diseases (Aw and Rose 2012, Hlavsa et al. 2011)

Microbes	Diseases	Pathogens	Symptoms	Causes	Incubation period
Viruses	Adenovirus infection	<i>Adenoviridae virus</i>	Vary depending on which part of the body is infected	Fecal contaminated water	5 to 8 days
	Hepatitis A	<i>Hepatitis A virus</i>	fever, fatigue, stomach pain, nausea, dark urine and jaundice	Fecal contaminated water and food (e.g. ready-to-eat-foods, fruit and juice, milk products, shellfish, salads, vegetables and sandwiches)	28 days
	Viral gastroenteritis	<i>Calicivirus</i>	Diarrhea, vomiting, nausea, cramps, headache, muscle aches, tiredness, and slight fever	Fecal water contamination and food/ seafood (including ready-to-eat food such as salads, sandwiches and bread)	24 to 48 hours
Parasites	Amebiasis	<i>Entamoeba histolytica</i>	Diarrhea, stomach pain, and stomach cramping	Fecal contaminated water (e.g. ingesting infected pool or water supply)	2 to 4 weeks
	Cryptosporidiosis	<i>Cryptosporidium</i>	stomach rams, dehydration, nausea, vomiting, fever, and weight loss	Fecal contaminations (e.g. chlorinated pool)	2 to 10 days
	Giardiasis	<i>Giardia lamblia</i>	Diarrhea, excess gas, stomach or abdominal cramps, and upset stomach or nausea	Fecal contaminated water (e.g. swallowing recreational water contaminated with Giardia)	1 to 2 weeks

2.4 Molecular methods applied to detect, quantify and determine sources of microbial pollution

Microbial source tracking using molecular-based techniques in freshwater emerged with the second millennia (Hagedorn et al., 1999, Harwood et al., 2000, Bernhard and Field, 2000b, Bernhard and Field, 2000a, Wuertz et al., 2011, Harwood et al., 2014). This progress was mainly influenced by two reasons: (1) an attempt to collect concrete evidences that links human health risks to recreational water use and (2) efficient identification of FIB sources in waterways (Meays et al., 2004). The implementation process of this strategy is still ongoing and uses mainly library-independent analysis (Santo Domingo et al., 2007, Wuertz et al., 2011). The Library-independent methods for bacteria include cultivation-independent and cultivation dependent techniques. The cultivation-cultivation dependent techniques depend on the use of PCR (e.g. targeting specific bacteria such as *Bacteroidales* and *Rhodococcus*) and qPCR-based (e.g. targeting 16S rRNA gene or functional genes in specific bacteria such *Bacteroidales* and targeting specific DNA fragments such as cell surface proteins through metagenomic) (Wuertz et al., 2011, Harwood et al., 2014). While, cultivation-depend techniques include enumeration of specific or diagnostic colonies, hybridization of individual isolates, and community profiling (Wuertz et al., 2011, Cao et al., 2011b). The targets of the library-independent technique of interest for studies discussed in this thesis were host-associated *Bacteroidales* (Bernhard and Field, 2000b, Bernhard and Field, 2000a, Kildare et al., 2007, Shanks et al., 2008, Shanks et al., 2009, Shanks et al., 2007, Shanks et al., 2010a, Shanks et al., 2010b).

Figure 1: Timeline (2000 – 2017) for the major milestone in microbial source tracking by *Bacteroidales*-associated markers. All referenced studies are discussed in Thesis chapter 2

The host-associated markers classified under the library-independent technique have increased the development of alternative human-associated or animal-associated contamination detection markers. Bernhard and Field's (2000) study was the first to develop alternative human- and animal-associated markers based on obligate anaerobic *Bacteroides-Prevotella*. Polymerase chain reaction (PCR) primers were developed to target general *Bacteroidales* and host-associated *Bacteroidales* of human, ruminant and pig related species. The *Bacteroides-Prevotella* group was used for the development of these molecular biology-based techniques due to the taxon abundance in human feces and the ability to adapt to the host environment and living conditions (Bernhard and Field, 2000b, Gourmelon et al., 2007, Wuertz et al., 2011).

Bernhard and Field (2000) developed two human-specific markers, HF183F and HF134F. However, HF183F was the only one with high sensitivity and specificity to human feces. Additional human-associated quantification polymerase chain reactions (qPCR) assays were developed and targeted the 16S rRNA gene of *Bacteroidales* (Seurinck et al., 2005, Layton et al., 2006, Okabe et al., 2007, Shanks et al., 2009). These qPCR based assays had an overlapping targeted region of 600bp on 16S rRNA of *B. dorei* (Harwood et al., 2000). A recent cross-laboratory study of performance of human fecal anaerobe-associate PCR-based assays suggested that HF183 Taqman-based assay was more effective than other assays in California area (Sidhu et al., 2013, Boehm et al., 2013). H183-Taqman was selected after being compared to different assays that included BacH (Reischer et al., 2007), BacHum (Kildare et al., 2007), BsteriF1 (Shanks et al., 2010b), BtH (*B. theta*) (Yampara Iquise et al., 2008), gyeB (Lee and Lee, 2008), H183 endpoint (Bernhard and Field, 2000b), HF183 SYBR (Seurinck et al., 2005) and HumM2 (Shanks et al., 2009). Despite the high performance of the

HF183-Taqman assay, a combination of at least two assays (e.g. HF183 Taqman and BtH) was recommended to achieve better results in source tracking (Field and Samadpour, 2007, Stewart et al., 2013).

The application of host-associated *Bacteroidales* is being adapted by numerous countries worldwide and findings of geographical variation are being addressed by local or regional validation studies (Gawler et al., 2007, Edge et al., 2010, Neave et al., 2014, Harwood et al., 2009, Layton et al., 2013, Reischer et al., 2013, Odagiri et al., 2015, Matsuki et al., 2002). Tropical regions has seen a number of host-associated *Bacteroidales* markers validation and source tracking studies (Odagiri et al., 2015, Jenkins et al., 2009, Ahmed et al., 2010, Knappett et al., 2011, Pickering et al., 2011, Pickering et al., 2012, Mattioli et al., 2013, Harris et al., 2016, Chua et al., 2010, Nshimiyimana et al., 2014), but more studies are still needed as the vast area of the tropical regions is still unstudied.

The increase in the application of host-associated *Bacteroidales* markers is providing more insights on microbial source tracking in various environments. However, it should be noted that while designing new host-associated *Bacteroidales* markers to cover a new region, basic requirements for library independent technique proposed by Wuertz et al. (2011) should apply. These requirements suggest (1) the bacterial target should only be present in the source group (e.g. animal species, ruminant, or class of *Mammalia*) of interest and absent in other source groups; (2) bacterial target should be in similar numbers in feces of all subgroups and highly abundant in source groups (e.g. concentrations of the target could be higher or comparable to those of FIB); and (3) in the presence of significant fecal pollution sources, the occurrence of the target in other

source groups present should be known. In addition, survival and fate of the target in the environment should be evaluated (Wuertz et al., 2011).

2.5 Survival and decay of source identifiers: Can we discriminate current from aged fecal pollutions?

A number of studies on survival and persistence of host-associated *Bacteroidales* have been performed under temperate environmental conditions. They involved observing the growth or die-off in mesocosm or microcosms of marine or freshwater spiked with sewage or human feces or pure cultures of bacteria from *Bacteroidales* species (Kreader, 1998, Seurinck et al., 2005, Okabe et al., 2007, Walters and Field, 2009, Dick et al., 2010, Sokolova et al., 2012, Bae and Wuertz, 2009b, Bae and Wuertz, 2012, Bae and Wuertz, 2014, Bae and Wuertz, 2015). As the host-associated *Bacteroidales* markers being targeted are obligate anaerobes, there is a growing interest in understanding factors contributing to their decay and persistence in sediments. On one hand, such interest is associated with the fact that being anaerobic reduces their regrowth potential as recent discharge in surface water and inability of determining recent fecal contaminations (Walters and Field, 2006, Bae and Wuertz, 2009b). On the other hand, *Bacteroidales spp.* may survive longer in sediments under anaerobic condition with the potential of release into the overlying water under turbulent conditions (Le Fevre and Lewis, 2003, Craig et al., 2004, Ferguson et al., 2005). In addition, sediments could provide favorable environmental conditions to *Bacteroidales* based on nutrient availability, reduced predation, protection from sunlight inactivation, and biofilm formation (Craig et al., 2004, Davies et al., 1995, Lee et al., 2006, Hartz et al., 2008, Feng et al., 2010, Balzer et al., 2010).

The environmental factors that have been studied extensively under temperate conditions include temperature, sunlight, and salinity. The most studied of all these environmental factors is temperature (Balleste and Blanch, 2010, Walters and Field, 2009, Anderson et al., 2005, Okabe and Shimazu, 2007, Schulz and Childers, 2011, Kobayashi et al., 2013, Bae and Wuertz, 2012, Bae and Wuertz, 2009b, Kreader, 1998). Higher temperature ($> 25^{\circ}\text{C}$) led to increased die-off and nucleic acids degradation for *Bacteroides* and their related markers (Walters et al., 2009, Kreader, 1998, Okabe and Shimazu, 2007, Seurinck et al., 2005, Savichtcheva et al., 2005). The sunlight was reported challenging to study as it is related to the increase of oxygen in water, which in turn has an impact on the aerobic bacteria that could increase or decrease the survival and persistence of *Bacteroides* (Walters and Field, 2009, Bae and Wuertz, 2009b). Salinity did not have a major impact on die-off of *Bacteroides* (Okabe and Shimazu, 2007, Bae and Wuertz, 2015).

Beyond the environmental factors, studies have also explored the role of predation on the survival and persistence of *Bacteroidales*, and findings have shown that they have a considerable impact (Bell et al., 2009, Kreader, 1998, Okabe and Shimazu, 2007, Kobayashi et al., 2013). The effects of temperature and predators on the growth or decay rate of *Bacteroidales* were the first parameters to be studied under temperate climate. This first study of the effect of abiotic and biotic parameters examined the survival and persistence of *Bacteroides distasonis* using end-point PCR assays and Ohio River water as a microcosm water at different temperature levels and presence or absence of predators (Kreader, 1998). Regarding predation, the absence of eukaryotic predators was associated with the increase of persistence of *B. distasonis* DNA for a week at 24°C . This finding demonstrated that bacterial grazing by protozoa has also a

major effect on the survival of bacteria. The findings of this first study were also confirmed by other similar studies in different geographic locations under the temperate climate (Seurinck et al., 2005, Balleste and Blanch, 2010). Seurinck and colleagues (2005) confirmed similar effect of temperature on the human-specific marker (HF183). High temperature (28°C) increased the degradation of HF183 marker to 2 days in river freshwater, while the lower temperature (4°C) increased the persistence at 24 days (Seurinck et al., 2005). This investigation confirmed the findings on the role of temperature and predation in regulating persistence and survival of host-associated markers (Balleste and Blanch, 2010).

Persistence and survival of host-associated *Bacteroidales* (HPyV-qPCR, HF183 and *M. smithii* – end-point PCR), FIB, and pathogens (Adenoviruses-end-point PCR) were examined and compared in sewage. McQuaig and colleagues (2009) studied these indicators using sewage for a duration of 28 days at 25°C and 35°C. The results showed that FIB were persistent at elevated concentration for a prolonged time at 25°C, but showed a decreasing pattern at 35°C. On the other hand, HPyV did not decline significantly for 21 days, while HF183 marker was persistent for only 14 days at 25°C and was detected for the entire duration of the investigation at 35°C. In addition, Adenovirus and *M. smithii* qPCR signals persisted for the entire duration of the experiment and at both 25°C and 35°C temperatures (McQuaig et al., 2009).

Apart from challenges associated with survival and persistence of *Bacteroidales* spp. under different environmental conditions, current molecular applied detection and quantification techniques such as qPCR cannot discern intact from impaired cells since DNA signal was found persistent after cell death (Walters and Field, 2009). This

challenge led to the application of a sample treatment such as the use of Propidium monoazide (PMA) before qPCR. This PMA-based technique demonstrated promising results in differentiating DNA from intact to that of impaired cells (Bae and Wuertz, 2012, Bae and Wuertz, 2009a, Bae and Wuertz, 2014, Bae and Wuertz, 2009b). This mechanism is achieved during the qPCR by the inhibition of the amplification of DNA from damaged cells (Nocker et al., 2007, Bae and Wuertz, 2009b). Chemically, PMA penetrates the membrane of impaired cells and binds to their DNA during the light exposure and reduces the possibility of amplification. Bae and Wuertz (2009b) demonstrated that the combination of qPCR and PMA could allow the detection of intact *Bacteroidales* cells in wastewater and stormwater with a concentration of 1,000 mg l⁻¹ of solids. This property of PMA has increased its application in discriminating viable and dead fecal *Bacteroidales*; understanding decay rates of host-specific fecal *Bacteroidales* in seawater; the comparison of survival of *Bacteroidales* cells to that of other indicator species such as *Enterococcus spp.*, *campylobacter jejuni* and adenovirus; and decay kinetics of host-associated *Bacteroidales* markers in intact cells under continuous-flow (Bae and Wuertz, 2012, Bae and Wuertz, 2009a, Bae and Wuertz, 2014, Bae and Wuertz, 2009b). These studies contributed to the understanding of factors driving survival and persistence of *Bacteroidales* cells and that of qPCR signal from DNA of intact or impaired cells. In addition, the results provided evidence to the development of methods that could discriminate recent from aged fecal contaminations.

This review of the survival and persistence of host-associated *Bacteroidales* showed that a number of studies have been completed under temperate climate and covered major environmental factors that could affect die-off and survival under a variety of

conditions. In addition, techniques of discerning intact from impaired cells have also been developed. However, similar studies are still needed under tropical conditions where the temperature, sunlight, and other environmental conditions favor the growth and survival of FIB. In addition, some environmental factors like predation are still not well understood and can affect the ability to distinguish aged from recent contaminants especially in areas with high abundance of predators.

2.6 New sequencing-based tool to study pathogens harbored by microbial pollution in urban areas

The goal of this approach was to develop new tools relying on the analysis and understanding of the bacterial community composition and structure for source tracking and identification of pathogen-like bacteria (Cao et al., 2011b). The remarkable variation of the gut microbiome environment from host to host or animal to human due to differences in dietary preferences has also increased the interest in applying microbial community studies to source tracking (Ley et al., 2008, Shanks et al., 2011). Sewage microbial communities differ strongly from those of pristine water (McLellan et al., 2010). The effects of fecal material on water quality include changing community composition and the introduction of different fecally derived chemicals (Cao et al., 2011a, Cao et al., 2011b). This generates opportunities for source tracking by differentiating sewage/feces-affected from autochthonous microbial communities (Unno et al., 2010, Cao et al., 2011a, Dubinsky et al., 2012).

Cao and colleagues (2013) examined three different methods to understand their usefulness in applying bacterial community analysis to identify sources of microbial

pollution. The methods examined included fingerprinting analysis (terminal restriction fragment length polymorphism (TRFLP); phylogenic microarray (PhyloChip), and next-generation sequencing (Illumina) (Cao et al., 2013). They were tested using various mixed fecal samples (e.g., sewage, septage and feces from humans, dogs, pigs, deer, horses, cows, chickens, gulls, pigeons, and geese) to simulate microbial water pollution.. The methods were able to identify 90% of the microbial sources (e.g., specific animal) in tested samples. Of all animal fecal samples and samples from other sources (e.g. sewage or septage) included in mixed samples, only few were identified with 100% sensitivity and specificity. These included the differentiation of septage from sewage and feces from human, deer or horse (Cao et al., 2013). This comparison emphasized the usefulness of microbial community analysis in identifying the sources of microbial water pollution and demonstrated that sequencing technology (Illumina next-generation sequencing) had similar capabilities as other technologies used for the same task.

References: Chapter 2

- ABDELZAHER, A. M., WRIGHT, M. E., ORTEGA, C., SOLO-GABRIELE, H. M., MILLER, G., ELMIR, S., NEWMAN, X., SHIH, P., BONILLA, J. A., BONILLA, T. D., PALMER, C. J., SCOTT, T., LUKASIK, J., HARWOOD, V. J., MCQUAIG, S., SINIGALLIANO, C., GIDLEY, M., PLANO, L. R., ZHU, X., WANG, J. D. & FLEMING, L. E. 2010. Presence of pathogens and indicator microbes at a non-point source subtropical recreational marine beach. *Applied and Environmental Microbiology*, 76, 724-32.
- AHMED, W., YUSUF, R., HASAN, I., GOONETILLEKE, A. & 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, 838-845.
- ANDERSON, M. L., WHITLOCK, J. E. & HARWOOD, V. J. 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. *Applied and Environmental Microbiology*, 71, 8.
- ASHBOLT, N. J., WOK, G. & M, S. 2001. Indicators of microbial water quality. *Water Quality Guidelines: Guidelines Standards and Health*. London: IWA Publishing
- AW, T. G. & GIN, K. Y. 2011. Prevalence and genetic diversity of waterborne pathogenic viruses in surface waters of tropical urban catchments. *Applied and Environmental Microbiology*, 110, 11.
- AW, T. G. & GIN, K. Y. H. 2010. Environmental surveillance and molecular characterization of human enteric viruses in tropical urban wastewaters. *Applied and Environmental Microbiology*, 109, 716-730.
- AW, T. G., GIN, K. Y. H., EAN-OON, L. L., CHEN, E. X. & WOO, C. H. 2009. Prevalence and Genotypes of Human Noroviruses in Tropical Urban Surface Waters and Clinical Samples in Singapore. *Applied and Environmental Microbiology*, 75, 4984-4992.
- AW, T. G. & ROSE, J. B. 2012. Detection of pathogens in water: from phylochips to qPCR to pyrosequencing. *Current Opinion in Biotechnology*, 23, 422-430.
- BAE, S. & WUERTZ, S. 2009a. Discrimination of viable and dead fecal *Bacteroidales* bacteria by quantitative PCR with propidium monoazide. *Applied and Environmental Microbiology*, 75, 2940-4.

- BAE, S. & WUERTZ, S. 2009b. Rapid decay of host-specific fecal *Bacteroidales* cells in seawater as measured by quantitative PCR with propidium monoazide. *Water Research*, 43, 4850-9.
- BAE, S. & WUERTZ, S. 2012. Survival of host-associated *Bacteroidales* cells and their relationship with *Enterococcus spp.*, *Campylobacter jejuni*, *Salmonella enterica* serovar Typhimurium, and adenovirus in freshwater microcosms as measured by propidium monoazide-quantitative PCR. *Applied and Environmental Microbiology*, 78, 922-32.
- BAE, S. & WUERTZ, S. 2014. Differential decay kinetics of host-associated *Bacteroidales* markers in intact cells under continuous-flow freshwater and seawater conditions as measured by PMA-qPCR. In: ENGINEERING, D. O. C. A. E. (ed.). UC Davis.
- BAE, S. & WUERTZ, S. 2015. Decay of host-associated *Bacteroidales* cells and DNA in continuous-flow freshwater and seawater microcosms of identical experimental design and temperature as measured by PMA-qPCR and qPCR. *Water Research* 70, 205-13.
- BALLESTE, E. & BLANCH, A. R. 2010. Persistence of *Bacteroides* Species Populations in a River as Measured by Molecular and Culture Techniques. *Applied and Environmental Microbiology*, 76, 7608-7616.
- BALZER, M., WITT, N., FLEMMING, H. C. & WINGENDER, J. 2010. Faecal indicator bacteria in river biofilms. *Water Science and Technology*, 61, 1105-11.
- BELL, A., LAYTON, A., MCKAY, L., WILLIAMS, D., GENTRY, R. & SAYLER, G. 2009. Factors Influencing the Persistence of Fecal in Stream Water. *Environmental Quality*, 38, 1224-1232.
- BERNHARD, A. E. & 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, 1587-1594.
- 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, 4571-4574.
- BOEHM, A., VAN DE WERFHORST, L. C., GRIFFITH, J., HOLDEN, P., JAY, J., SHANKS, O., WANG, D. & WEISBERG, S. 2013. Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. *Water Research*, 47, 6812-6828.

- BORDALO, A. A. 2001. Water quality and uses of the Bangpakong River (Eastern Thailand). *Water Research*, 35, 3635-3642.
- CAI, L., JU, F. & ZHANG, T. 2014. Tracking human sewage microbiome in a municipal wastewater treatment plant. *Applied Microbiology and Biotechnology*, 98, 3317-3326.
- CAO, Y., VAN DE WERFHORST, L. C., DUBINSKY, E., BADGLEY, B., SADOWSKY, M., GRIFFITH, J. F., ANDERSEN, G. & HOLDEN, P. 2013. Evaluation of molecular community analysis methods for discerning fecal sources and human waste. *Water Research*, 47, 6862-72.
- CAO, Y., VAN DE WERFHORST, L. C., SERCU, B., MURRAY, J. L. & HOLDEN, P. A. 2011a. Application of an integrated community analysis approach for microbial source tracking in a coastal creek. *Environmental Science and Technology*, 45, 7195-201.
- CAO, Y., WU, C. H., ANDERSEN, G. L. & HOLDEN, P. A. 2011b. Community analysis-based methods. In: HAGEDORN, C., BLANCH, A. R. & HARWOOD, V. J. (eds.) *Microbial Source Tracking: Methods, Applications, and Case Studies*. New York: Springer.
- CHAPPELL, N., DOUGLAS, I., HANAPI, J. & TYCH, W. 2004. Sources of suspended sediment within a tropical catchment recovering from selective logging. *Hydrological Processes*, 18, 685-701.
- CHUA, L. H. C., SHANAHAN, P., EDMOND, Y. M. L., SHUY, E. B., THOMPSON, J., DIXON, C. C., KERIGAN, K. K., NSHIMYIMANA, J. P., YEAGER, J. M., LEE, L. J. & POR, Y. L. 2010. Dry Weather bacteria monitoring and variation with land use for Kranji Reservoir catchment, Singapore. *17th Congress of the Asia and Pacific Division of the International Association of Hydraulic Engineering and Research*. Auckland, New Zealand.
- CRAIG, D. L., FALLOWFIELD, H. J. & CROMAR, N. J. 2004. Use of microcosms to determine persistence of *Escherichia coli* in recreational coastal water and sediment and validation with in situ measurements. *Applied Microbiology*, 96, 922-30.
- CRUMP, J. A., GRIFFIN, P. M. & ANGULO, F. J. 2002. Bacterial contamination of animal feed and Its relationship to human foodborne illness. *Clinical Infectious Diseases*, 35, 859-865.
- CWP 2003. Impacts of impervious cover on aquatic systems. *Watershed Protection Research Monograph*. Ellicott City: Center for Watershed Protection.

- DAVIES, C. M., LONG, J. A., DONALD, M. & ASHBOLT, N. J. 1995. Survival of fecal microorganisms in marine and freshwater sediments. *Applied Environmental Microbiology*, 61, 1888-96.
- DICK, L., STELZER, E., BERTKE, E., FONG, D. & STOECKEL, D. 2010. Relative decay of *Bacteroidales* microbial source tracking markers and cultivated *Escherichia coli* in freshwater microcosms. *Applied and Environmental Microbiology*, 76, 3255-62.
- DOREVITCH, S., ASHBOLT, N. J., FERGUSON, C. M., FUJIOKA, R., MCGEE, C. D., SOLLER, J. A. & WHITMAN, R. L. 2010. Meeting report: knowledge and gaps in developing microbial criteria for inland recreational waters. *Environmental Health Perspect*, 118, 871-6.
- DUBINSKY, E. A., ESMAILI, L., HULLS, J. R., CAO, Y., GRIFFITH, J. F. & ANDERSEN, G. L. 2012. Application of phylogenetic microarray analysis to discriminate sources of fecal pollution. *Environmental Science and Technology*, 46, 4340-7.
- EDGE, T. A., HILL, S., SETO, P. & MARSALEK, J. 2010. Library-dependent and library-independent microbial source tracking to identify spatial variation in faecal contamination sources along a Lake Ontario beach (Ontario, Canada). *Water Science And Technology: A Journal Of The International Association On Water Pollution Research*, 62, 719-727.
- FEACHEM, R., BRADLEY, D., H, G. & MARA, D. 1983. *Sanitation and disease; health aspects of excreta and wastewater management*, Chichester, John Wiley & Sons.
- FEACHEM, R., MCGARRY, M. & MARA, D. 1977. *Water, Waste and Health in Hot Climates*, Chichester, John Wiley & Sons.
- FENG, F., GOTO, D. & YAN, T. 2010. Effects of autochthonous microbial community on the die-off of fecal indicators in tropical beach sand. *FEMS Microbiology Ecology*, 74, 214-25.
- FERGUSON, D. M., MOORE, D. F., GETRICH, M. A. & ZHOWANDAI, M. H. 2005. Enumeration and speciation of *Enterococci* found in marine and intertidal sediments and coastal water in southern California. *Applied Microbiology*, 99, 598-608.
- FIELD, K. G. & SAMADPOUR, M. 2007. Fecal source tracking, the indicator paradigm, and managing water quality. *Water Research*, 41, 3517-38.
- GAWLER, A. H., BEECHER, J. E., BRANDÃO, J., CARROLL, N. M., FALCÃO, L., GOURMELON, M., MASTERSON, B., NUNES, B., PORTER, J. & RINCÉ, A. 2007. Validation of host-specific *Bacteroidales* 16S rRNA genes as

markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. *Water Research*, 41, 3780.

GIAMBELLUCA, T., ZIEGLER, A., NULLET, M., TRUONG, D. & TRAN, L. 2003. Transpiration in a small tropical forest patch. *Agricultural and Forest Meteorology*, 117, 1-22.

GIVEN, S., PENDLETON, L. H. & BOEHM, A. B. 2006. Regional public health cost estimates of contaminated coastal waters: a case study of gastroenteritis at southern California beaches. *Environmental Science and Technology*, 40, 4851-8.

GOURMELON, M., CAPRAIS, M. P., SEGURA, R., LE MENNEC, C., LOZACH, S., PIRIOU, J. Y. & RINCE, A. 2007. Evaluation of two library-independent microbial source tracking methods to identify sources of fecal contamination in French estuaries. *Applied and Environmental Microbiology*, 73, 4857-66.

HAGEDORN, C., ROBINSON, S. L., FILTZ, J. R., GRUBBS, S. M., ANGIER, T. A. & RENEAU, R. B. 1999. Determining sources of fecal pollution in a rural Virginia watershed with antibiotic resistance patterns in fecal streptococci. *Applied and Environmental Microbiology*, 65, 5522-31.

HAILE, R. W., WITTE, J. S., GOLD, M., CRESSEY, R., MCGEE, C., MILLIKAN, R. C., GLASSER, A., HARAWA, N., ERVIN, C., HARMON, P., HARPER, J., DERMAND, J., ALAMILLO, J., BARRETT, K., NIDES, M. & WANG, G. 1999. The health effects of swimming in ocean water contaminated by storm drain runoff. *Epidemiology*, 10, 355-63.

HARRIS, A. R., PICKERING, A. J., HARRIS, M., DOZA, S., ISLAM, M. S., UNICOMB, L., LUBY, S., DAVIS, J. & BOEHM, A. B. 2016. Ruminants Contribute Fecal Contamination to the Urban Household Environment in Dhaka, Bangladesh. *Environmental Science and Technology*, 50, 4642-9.

HARTZ, A., CUVELIER, M., NOWOSIELSKI, K., BONILLA, T. D., GREEN, M., ESIÖBU, N., MCCORQUODALE, D. S. & ROGERSON, A. 2008. Survival potential of *Escherichia coli* and *Enterococci* in subtropical beach sand: implications for water quality managers. *Environmental Quality*, 37, 898-905.

HARWOOD, V., STALEY, C., BADGLEY, B., BORGES, K. & KORAJKIC, A. 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews*, 38, 1-40.

HARWOOD, V. J., BROWNELL, M., WANG, S., LEPO, J., ELLENDER, R. D., AJIDAHUN, A., HELLEIN, K. N., KENNEDY, E., YE, X. & FLOOD, C. 2009. Validation and field testing of library-independent microbial source tracking methods in the Gulf of Mexico. *Water Research*, 43, 4812-9.

- HARWOOD, V. J., WHITLOCK, J. & WITHINGTON, V. 2000. Classification of antibiotic resistance patterns of indicator bacteria by discriminant analysis: use in predicting the source of fecal contamination in subtropical waters. *Applied and Environmental Microbiology*, 66, 3698-704.
- HAYHOE, S., NEILL, C., PORDER, S., MCHORNEY, R., LEFEBVRE, P., COE, M., ELSENBEEER, H. & KRUSCHE, A. 2011. Conversion to soy on the Amazonian agricultural frontier increases streamflow without affecting stormflow dynamics. *Global Change Biology*, 17, 1821-1833.
- HLAVSA, M. C., ROBERTS, V. A., ANDERSON, A. R., HILL, V. R., KAHLER, A. M., ORR, M., GARRISON, L. E., HICKS, L. A., NEWTON, A., HILBORN, E. D., WADE, T. J., BEACH, M. J. & YODER, J. S. 2011. Surveillance for waterborne disease outbreaks and other health events associated with recreational water--United States 2007-2008. *Morbidity and Mortality Weekly Report (MMWR)*. Atlanta: Center for Disease Control.
- HORMAN, A., RIMHANEN-FINNE, R., MAUNULA, L., VON BONSDORFF, C. H., TORVELA, N., HEIKINHEIMO, A. & HANNINEN, M. L. 2004. *Campylobacter* spp., *Giardia* spp., *Cryptosporidium* spp., noroviruses, and indicator organisms in surface water in southwestern Finland, 2000-2001. *Applied and Environmental Microbiology*, 70, 87-95.
- IBEKWE, A. M., LEDDY, M., MURINDA, S. & AZIZ, R. 2013. Potential Human Pathogenic Bacteria in a Mixed Urban Watershed as Revealed by Pyrosequencing. *PLoS ONE*, 8, e79490.
- JAMWAL, P., MITTAL, A. K. & MOUCHEL, J.-M. 2011. Point and non-point microbial source pollution: A case study of Delhi. *Physics and Chemistry of the Earth, Parts A/B/C*, 36, 490-499.
- JENKINS, M. W., TIWARI, S., LORENTE, M., GICHABA, C. M. & WUERTZ, S. 2009. Identifying human and livestock sources of fecal contamination in Kenya with host-specific *Bacteroidales* assays. *Water Research*, 43, 4956-66.
- JONES, G. W. 2013. The population of Southeast Asia. *Working paper*. Singapore: Asia Research Institute.
- KHAN, A., AKHTAR, S., AHMAD, M., SAYED, S., HADI, S., KHAN, M., JALIL, M., AHMED, R. & FARUQUEE, M. 1997. Arsenic contamination in groundwater and its effect on human health with particular reference to Bangladesh. *Preventive and Social Medicine*, 16, 8.
- KILDARE, B., LEUTENEGGER, C., MCSWAIN, B., BAMBIC, D., RAJAL, V. & WUERTZ, S. 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: A Bayesian approach. *Water Research*, 41, 15.

- KJELLSTROM, T. 1986. *"Itai-Itai Disease."* In *Cadmium and Health*, Boca Raton, FL, CRC Press.
- KNAPPETT, P. S., ESCAMILLA, V., LAYTON, A., MCKAY, L. D., EMCH, M., WILLIAMS, D. E., HUQ, R., ALAM, J., FARHANA, L., MAILLOUX, B. J., FERGUSON, A., SAYLER, G. S., AHMED, K. M. & VAN GEEN, A. 2011. Impact of population and latrines on fecal contamination of ponds in rural Bangladesh. *Science of the Total Environment*, 409, 3174-82.
- KOBAYASHI, A., SANO, D. & OKABE, S. 2013. Effects of temperature and predator on the persistence of host-specific *Bacteroides-Prevotella* genetic markers in water. *Water Science and Technology*, 67, 838-45.
- KREADER, C. A. 1998. Persistence of PCR-detectable *Bacteroides distasonis* from human feces in river water. *Applied and Environmental Microbiology*, 64, 4103-4105.
- KUCKLICK, J., SIVERTSEN, S., SANDERS, M. & SCOTT, G. 1997. Factors influencing polycyclic aromatic hydrocarbon distributions in South Carolina estuarine sediments. *Experimental Marine Biology and Ecology*, 213, 13-29.
- LAYTON, A., MCKAY, L., WILLIAMS, D., GARRETT, V., GENTRY, R. & SAYLER, G. 2006. Development of *Bacteroides* 16S rRNA gene TaqMan-based real-time PCR assays for estimation of total, human, and bovine fecal pollution in water. *Applied and Environmental Microbiology*, 72, 4214-24.
- LAYTON, B., CAO, Y., EBENTIER, D., HANLEY, K., BALLESTÉ, E., BRANDÃO, J., BYAPPANAHALLI, M., CONVERSE, R., FARNLEITNER, A., GENTRY SHIELDS, J., GIDLEY, M., GOURMELON, M., LEE, C., LEE, J., LOZACH, S., MADI, T., MEIJER, W., NOBLE, R., PEED, L., REISCHER, G., RODRIGUES, R., ROSE, J., SCHRIEWER, A., SINIGALLIANO, C., SRINIVASAN, S., STEWART, J., VAN DE WERFHORST, L. C., WANG, D., WHITMAN, R., WUERTZ, S., JAY, J., HOLDEN, P., BOEHM, A., SHANKS, O. & GRIFFITH, J. 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. *Water Research*, 47, 6897-6908.
- LE FEVRE, N. M. & LEWIS, G. D. 2003. The role of resuspension in *Enterococci* distribution in water at an urban beach. *Water Science and Technology*, 47, 205-10.
- LEE, C. M., LIN, T. Y., LIN, C. C., KOHBODI, G. A., BHATT, A., LEE, R. & JAY, J. A. 2006. Persistence of fecal indicator bacteria in Santa Monica Bay beach sediments. *Water Research*, 40, 2593-602.
- LEE, G. & LEE, C. 2008. Molecular detection and characterization of human enteroviruses in Korean surface water. *Microbiology*, 46, 319-24.

- LERBERG, S. B. 2000. Responses of tidal creek macrobenthic communities to the effects of watershed development. *Estuaries*, 23, 838-853.
- LEY, R. E., HAMADY, M., LOZUPONE, C., TURNBAUGH, P. J., RAMEY, R. R., BIRCHER, J. S., SCHLEGEL, M. L., TUCKER, T. A., SCHRENZEL, M. D., KNIGHT, R. & GORDON, J. I. 2008. Evolution of Mammals and Their Gut Microbes. *Science*, 320, 1647-1651.
- LIN, C. & TONG, Z. 2013. Detecting Human Bacterial Pathogens in Wastewater Treatment Plants by a High-Throughput Shotgun Sequencing Technique. *Environmental Science and Technology*, 47, 5433-5441.
- MATSUKI, T., WATANABE, K., FUJIMOTO, J., MIYAMOTO, Y., TAKADA, T., MATSUMOTO, K., OYAIZU, H. & TANAKA, R. 2002. Development of 16S rRNA-gene-targeted group-specific primers for the detection and identification of predominant bacteria in human feces. *Applied and Environmental Microbiology*, 68, 5445-51.
- MATTIOLI, M. C., PICKERING, A. J., GILSDORF, R. J., DAVIS, J. & BOEHM, A. B. 2013. Hands and water as vectors of diarrheal pathogens in Bagamoyo, Tanzania. *Environmental Science and Technology*, 47, 355-63.
- MCLELLAN, S. L., HUSE, S. M., MUELLER-SPITZ, S. R., ANDREISHCHEVA, E. N. & SOGIN, M. L. 2010. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environmental Microbiology*, 12, 378-392.
- MCQUAIG, S. M., SCOTT, T. M., LUKASIK, J. O., PAUL, J. H. & HARWOOD, V. J. 2009. Quantification of Human Polyomaviruses JC Virus and BK Virus by TaqMan Quantitative PCR and Comparison to Other Water Quality Indicators in Water and Fecal Samples. *Applied and Environmental Microbiology*, 75, 3379-3388.
- MEAYS, C., BROERSMA, K., NORDIN, R. & MAZUMDER, A. 2004. Source tracking fecal bacteria in water: a critical review of current methods. *Environmental Management*, 73, 71-9.
- MOORE, A. C., HERWALDT, B. L., CRAUN, G. F., CALDERON, R. L., HIGHSMITH, A. K. & JURANEK, D. D. 1993. Surveillance for waterborne disease outbreaks--United States, 1991-1992. *Morbidity and Mortality Weekly Report (MMWR) CDC Surveillance Summaries*. Center for Disease Control.
- NEA 2009. Weatherwise Singapore. *Meteorological Services Division*. Singapore: National Environmental Agency.

- NEAVE, M., LUTER, H., PADOVAN, A., TOWNSEND, S., SCHOBEN, X. & GIBB, K. 2014. Multiple approaches to microbial source tracking in tropical northern Australia. *Microbiology Open*, 3, 860-874.
- NEGRI, A. J. 2004. The impact of Amazonian deforestation on dry season rainfall. *Journal of Climate*, 17, 1306-1319.
- NOBLE, R. T., GRIFFITH, J. F., BLACKWOOD, A. D., FUHRMAN, J. A., GREGORY, J. B., HERNANDEZ, X., LIANG, X., BERA, A. A. & SCHIFF, K. 2006. Multitiered approach using quantitative PCR to track sources of fecal pollution affecting Santa Monica Bay, California. *Applied and Environmental Microbiology*, 72, 1604-12.
- NOCKER, A., SOSSA-FERNANDEZ, P., BURR, M. D. & CAMPER, A. K. 2007. Use of propidium monoazide for live/dead distinction in microbial ecology. *Applied and Environmental Microbiology*, 73, 5111-7.
- NSHIMYIMANA, J. P., EKKLESIA, E., SHANAHAN, P., CHUA, L. H. C. & THOMPSON, J. R. 2014. Distribution and abundance of human-specific *Bacteroides* and relation to traditional indicators in an urban tropical catchment. *Applied Microbiology*, 116, 1369-1383.
- NTU 2008. Water quality monitoring, modeling and management for Kranji Catchment/Reservoir system – Phases 1 and 2 / May 2004 to December 2007. Singapore: Division of Environmental and Water Resources Engineering, School of Civil and Environmental Engineering, Nanyang Technological University.
- ODAGIRI, M., SCHRIEWER, A., HANLEY, K., WUERTZ, S., MISRA, P. R., PANIGRAHI, P. & JENKINS, M. W. 2015. Validation of *Bacteroidales* quantitative PCR assays targeting human and animal fecal contamination in the public and domestic domains in India. *Science of the Total Environment*, 502, 462-470.
- OKABE, S., OKAYAMA, N., SAVICHTCHEVA, O. & ITO, T. 2007. Quantification of host-specific *Bacteroides-Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in freshwater. *Applied Microbiology and Biotechnology*, 74, 890-901.
- OKABE, S. & SHIMAZU, Y. 2007. Persistence of host-specific *Bacteroides-Prevotella* 16S rRNA genetic markers in environmental waters: effects of temperature and salinity. *Applied Microbiology and Biotechnology*, 76, 935-44.
- PICKERING, A. J., JULIAN, T. R., MAMUYA, S., BOEHM, A. B. & DAVIS, J. 2011. Bacterial hand contamination among Tanzanian mothers varies temporally and following household activities. *Tropical Medicine & International Health*, 16, 233-9.

- PICKERING, A. J., JULIAN, T. R., MARKS, S. J., MATTIOLI, M. C., BOEHM, A. B., SCHWAB, K. J. & DAVIS, J. 2012. Fecal contamination and diarrheal pathogens on surfaces and in soils among Tanzanian households with and without improved sanitation. *Environmental Science and Technology*, 46, 5736-43.
- PRÜSS, A. 1998. Review of epidemiological studies on health effects from exposure to recreational water. *International Journal of Epidemiology*, 27, 1-9.
- PUB 2011. Active Beautiful Clean waters -- Design guidelines. *Second Edition*. Singapore: Public Utility Board.
- RAJGIRE, A. V. 2013. Open defecation: A prominent source of pollution in drinking water in villages *Life Sciences Biotechnology and Pharma Research*, 2, 11.
- RAMÍREZ-TORO, G. I., BALLESTER, K., MINNIGH, H. & LEVINE, J. F. Microbial contaminants in shellfish harvesting areas in Puerto Rico. WaterMicro2013 Symposium: 17th International Symposium on Health-Related Water Microbiology 15-20th September, 2013 2013 Florianopolis Island, Santa Catarina, Brazil. International Water Association.
- REISCHER, G. H., EBDON, J. E., BAUER, J. M., SCHUSTER, N., AHMED, W., ÅSTRÖM, J., BLANCH, A. R., BLÖSCHL, G., BYAMUKAMA, D., COAKLEY, T., FERGUSON, C., GOSHU, G., KO, G., DE RODA HUSMAN, A. M., MUSHI, D., POMA, R., PRADHAN, B., RAJAL, V., SCHADE, M. A., SOMMER, R., TAYLOR, H., TOTH, E. M., VRAJMASU, V., WUERTZ, S., MACH, R. L. & FARNLEITNER, A. H. 2013. Performance Characteristics of qPCR Assays Targeting Human- and Ruminant-Associated Bacteroidetes for Microbial Source Tracking across Sixteen Countries on Six Continents. *Environmental Science and Technology*, 47, 8548-8556.
- REISCHER, G. H., KASPER, D. C., STEINBORN, R., FARNLEITNER, A. H. & MACH, R. L. 2007. A quantitative real-time PCR assay for the highly sensitive and specific detection of human faecal influence in spring water from a large alpine catchment area. *Letters in Applied Microbiology*, 44, 351-6.
- SANGER, D. M., *, M. S., D., A. F. H. & SCOTT, G. I. 1999. Tidal Creek and Salt Marsh Sediments in South Carolina Coastal Estuaries: I. Distribution of Trace Metals. *Archives of Environmental Contamination and Toxicology*, 37, 445-457.
- SANTO DOMINGO, J. W., BAMBIC, D. G., EDGE, T. A. & WUERTZ, S. 2007. Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. *Water Research*, 41, 3539-52.
- SAVICHTCHEVA, O., OKAYAMA, N., ITO, T. & OKABE, S. 2005. Application of a direct fluorescence-based live/dead staining combined with fluorescence in

situ hybridization for assessment of survival rate of *Bacteroides* spp. in drinking water. *Biotechnology and Bioengineering*, 92, 356-363.

- SCHULZ, C. J. & CHILDERS, G. W. 2011. Fecal *Bacteroidales* diversity and decay in response to variations in temperature and salinity. *Applied and Environmental Microbiology*, 77, 2563-2572.
- SEURINCK, S., DEFOIRDT, T., VERSTRAETE, W. & SICILIANO, S. 2005. Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environmental Microbiology*, 7, 249-59.
- SHANKS, O., ATIKOVIC, E., BLACKWOOD, A. D., LU, J., NOBLE, R., DOMINGO, J., SEIFRING, S., SIVAGANESAN, M. & HAUGLAND, R. 2008. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology*, 74, 745-52.
- SHANKS, O., KELTY, C., SIVAGANESAN, M., VARMA, M. & HAUGLAND, R. 2009. Quantitative PCR for genetic markers of human fecal pollution. *Applied and Environmental Microbiology*, 75, 5507-13.
- SHANKS, O., SANTO DOMINGO, J., LU, J., KELTY, C. & GRAHAM, J. 2007. Identification of bacterial DNA markers for the detection of human fecal pollution in water. *Applied and Environmental Microbiology*, 73, 2416-22.
- SHANKS, O., WHITE, K., KELTY, C., HAYES, S., SIVAGANESAN, M., JENKINS, M., VARMA, M. & HAUGLAND, R. 2010a. Performance assessment PCR-based assays targeting *Bacteroidales* genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology*, 76, 1359-66.
- SHANKS, O., WHITE, K., KELTY, C., SIVAGANESAN, M., BLANNON, J., MECKES, M., VARMA, M. & HAUGLAND, R. 2010b. Performance of PCR-based assays targeting *Bacteroidales* genetic markers of human fecal pollution in sewage and fecal samples. *Environmental Science and Technology*, 44, 6281-8.
- SHANKS, O. C., KELTY, C. A., ARCHIBEQUE, S., JENKINS, M., NEWTON, R. J., MCLELLAN, S. L., HUSE, S. M. & SOGIN, M. L. 2011. Community structures of fecal bacteria in cattle from different animal feeding operations. *Applied and Environmental Microbiology*, 77, 2992-3001.
- SIDHU, J. P. S., AHMED, W., GERNJAK, W., ARYAL, R., MCCARTHY, D., PALMER, A., KOLOTELO, P. & TOZE, S. 2013. Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. *Science of the Total Environment*, 463-464, 488-496.

- SOKOLOVA, E., ASTROM, J., PETTERSSON, T. J. R., BERGSTEDT, O., HERMANSSON, M. & ÅSTRÖM, J. 2012. Decay of *Bacteroidales* genetic markers in relation to traditional fecal indicators for water quality modeling of drinking water sources. *Environmental Science and Technology*, 46, 892-900.
- STEWART, J., BOEHM, A., DUBINSKY, E., FONG, T.-T., GOODWIN, K., GRIFFITH, J., NOBLE, R., SHANKS, O., VIJAYAVEL, K. & WEISBERG, S. 2013. Recommendations following a multi-laboratory comparison of microbial source tracking methods. *Water Research*, 47, 6829-38.
- SUMMERS, J. 1989. *Soho: a History of London's most Colourful Neighborhood* 117., London, Bloomsbury.
- TSATSAROS, J. H., BRODIE, J. E., BOHNET, I. C. & VALENTINE, P. 2013. Water quality degradation of coastal waterways in the wet tropics, Australia. *Water Air and Soil Pollutants*, 224, 22.
- UNNO, T., JANG, J., HAN, D., KIM, J. H., SADOWSKY, M. J., KIM, O. S., CHUN, J. & HUR, H. G. 2010. Use of barcoded pyrosequencing and shared OTUs to determine sources of fecal bacteria in watersheds. *Environmental Science and Technology*, 44, 7777-82.
- VERNBERG, F. J. 1992. Impact of urbanization on high-salinity estuaries in the southeastern United States. *Netherlands Journal of Sea Research*, 30, 239-248.
- WALSH, R. P. D., WALSH, R. P. D., BIDIN, K., BLAKE, W. H., CHAPPELL, N. A., CLARKE, M. A., DOUGLAS, I., GHAZALI, R., SAYER, A. M., SUHAIMI, J., TYCH, W. & ANNAMMALA, K. V. 2011. Long-term responses of rainforest erosional systems at different spatial scales to selective logging and climatic change. *Philosophical transactions - Royal Society. Biological sciences*, 366, 3340-3353.
- WALTERS, S. & FIELD, K. 2006. Persistence and growth of fecal *Bacteroidales* assessed by bromodeoxyuridine immunocapture. *Applied and Environmental Microbiology*, 72, 4532-9.
- WALTERS, S. & FIELD, K. 2009. Survival and persistence of human and ruminant-specific faecal *Bacteroidales* in freshwater microcosms. *Environmental Microbiology*, 11, 1410-1421.
- WALTERS, S., YAMAHARA, K. & BOEHM, A. 2009. Persistence of nucleic acid markers of health-relevant organisms in seawater microcosms: Implications for their use in assessing risk in recreational waters. *Water Research*, 43, 4929-4939.
- WHITE, G., BRADLEY, D. & WHITE, A. 1972. *Drawers of Water*, Chicago, University of Chicago Press.

- WILLETTS, J., WICKEN, J. & ROBINSON, A. 2009. Meeting the sanitation and water challenge in South-East Asia and the Pacific. *Sanitation and Water Conference*. Melbourne: International water center.
- WOHL, E., BARROS, A., BRUNSELL, N., CHAPPELL, N., COE, M., GIAMBELLUCA, T., GOLDSMITH, S., HARMON, R., JUVIK, J., MCDONNELL, J. & OGDEN, F. 2012. The hydrology of the humid tropics. *Nature Climate Change*, 2, 655-662.
- WUERTZ, S., WANG, D., REISCHER, G. H. & FARNLEITNER, A. H. 2011. Chapter 3: Library-Independent Bacterial Source Tracking Methods. *In*: HAGEDORN, C., BLANCH, A. R. & HARWOOD, V. J. (eds.) *Microbial Source Tracking: Methods, Applications, and Case Studies*. London Springer.
- YAMPARA IQUISE, H., ZHENG, G., JONES, J. E. & CARSON, C. A. 2008. Use of a *Bacteroides thetaiotaomicron*-specific alpha-1-6, mannanase quantitative PCR to detect human faecal pollution in water. *Applied Microbiology*, 105, 1686-93.
- ZIEGLER, A., GIAMBELLUCA, T., PLONDKE, D., LEISZ, S., TRAN, L., FOX, J., NULLET, M., VOGLER, J., MINH TROUNG, D. & TRAN DUC, V. 2007. Hydrological consequences of landscape fragmentation in mountainous northern Vietnam: Buffering of Hortonian overland flow. *Journal of hydrology*, 337, 52-67.

Chapter 3.

Variation of Bacterial Communities with Water Quality in an Urban Tropical Catchment

Abstract

A major challenge for assessment of water quality in tropical environments is the natural occurrence and potential growth of Fecal Indicator Bacteria (FIB). To gain a better understanding of the relationship between measured levels of FIB and the distribution of sewage-associated bacteria including potential pathogens in the tropics this study compared the abundance of the FIB *E. coli* and the *Bacteroidales* HF183 marker with bacterial community structure determined by next-generation amplicon sequencing. Water was sampled twice over 6 months from 18 sites within a tropical urban catchment and reservoir, followed by DNA extraction, and sequencing targeting the V3-V4 region of the 16S rRNA gene. Multivariate statistical analyses indicated that bacterial community composition (BCC) varied between reservoir and catchment, within catchment land-uses, and with *E. coli* MPN. Beta-regression indicated that the proportion of sequences from sewage-associated taxa (SAT) or pathogen-like sequences (PLS) were predicted most significantly by measured levels of *E. coli* (log MPN/100ml) ($\chi^2 > 8.7$; $p < 0.003$). In addition, SAT were significantly predicted by log HF183 levels ($\chi^2 = 13.1$; $p \leq 0.0003$) while PLS were not. Our study suggests measurements of *E. coli* MPN could be useful in predicting samples enriched in sewage-associated and pathogen-like bacteria in tropical environments despite the potential for non-conservative behavior.

3.1 Introduction

Human sewage discharge may contain water-borne pathogens that pose disease risk (McQuaig et al., 2006a, McLellan et al., 2010, McLellan and Eren, 2014, Lisle and Rose, 1995, Prüss, 1998, Zoetendal et al., 2006, Cai et al., 2014, Lin and Tong, 2011). Fecal Indicator Bacteria (FIB) occur at high levels in sewage, and thus quantification of FIB in water samples is widely employed as a proxy for sewage-associated health risks (McLellan and Eren, 2014, Harwood et al., 2014). However, studies in tropical environments have demonstrated that FIB may have prolonged persistence and possible growth in warm high-nutrient waters (Carrillo, 1985, Jimenez, 1989, Winfield and Groisman, 2003, Rochelle-Newall et al., 2015). In addition, the abundance of FIB does not enable discrimination of human from other animal fecal sources (Bernhard et al., 2003, Byappanahalli et al., 2012a) and FIB have been linked to indigenous environmental sources such as soil, sediments, native fauna, water, and plants (Byappanahalli et al., 2012b, Fujioka, 2001, Rivera, 1988, Yamahara et al., 2007). Newly developed molecular-based host-associated *Bacteroidales* assays are used for discriminating human fecal contamination sources from those of other animals (Savichtcheva et al., 2007, Savichtcheva and Okabe, 2006, Ley et al., 2002, Gregory et al., 2006, McQuaig et al., 2009). A recent study in the urban environment of Singapore suggested combining FIB and host-associated *Bacteroidales* genetic markers to better identify and detect sewage and risk from human sewage-associated pathogens in tropical environments (Liang et al., 2015). Although the use of genetic markers is promising for fecal microbial source tracking, most assays have been established under temperate climates (Harwood et al., 2014, Layton et al., 2013, Bae and Wuertz, 2012)

and there is still a need to validate and develop assays for application under tropical conditions.

Next-generation sequencing (NGS) has proven useful as a tool for detection of multiple pathogen-like bacteria within a sample based on the diversity of conserved biomarkers such as the 16S rRNA gene (Cao et al., 2011a, Cao et al., 2011b, Cao et al., 2013, Unno et al., 2010, Ibekwe et al., 2013, Shanks et al., 2011, Salipante et al., 2013). Previous studies have utilized NGS for characterizing bacterial communities and identifying pathogen-like bacteria in sewage and wastewater influent (McLellan et al., 2010), across treatment stages in wastewater treatment plants (Lin and Tong, 2011, Lin and Tong, 2013, Cai et al., 2014), and in the environment (Ibekwe et al., 2013). While previous surveys of the 16S rRNA gene have lacked the phylogenetic resolution necessary to distinguish sequences at the level of species, new longer-read technologies allow sequences >400bp which enable more confident taxonomic assignment. Microbial diversity recovered by NGS has not yet been compared to water quality predicted by the abundance of FIB and source tracking markers, especially in tropical environments.

The tropical urban island of Singapore has an advanced and reputable water and wastewater management infrastructure where stormwater is collected through engineered drains and stored in reservoirs. Studies of surface water quality in Singapore have noted elevated levels of FIB (*E. coli*, total coliforms, and enterococci) at different catchment sites (e.g. in horticultural and residential areas) during dry weather (i.e. defined as >48 hours after a rain event) (NTU, 2008, Nshimiyimana et al., 2014, Chua et al., 2010, Ekklesia et al., 2015). The goal of this study was to apply Illumina^{TM} MiSeq 16S rRNA gene amplicon sequencing to characterize the bacterial community

composition (BCC) and the proportion of sequences related to sewage-associated taxa (SAT) and human pathogens (i.e. pathogen-like sequences, PLS) in water samples collected from an urban reservoir and catchment in Singapore. Our findings show variation in BCC, SAT, and PLS with water quality indicators (*E. coli* and HF183). This information will be useful for evaluating the utility of FIB to predict potential exposure to human pathogens in tropical water samples and to identify specific bacterial targets that may be relevant for further monitoring using quantitative methods.

3.2. Methodology

3.2.1 Study design and sample collection

3.2.1.1 Sampling and Site Characterization

Water samples from 18 sites in an urban reservoir and catchment in the northwest of Singapore Island were collected during dry weather in January and June 2009 (NTU, 2008). The catchment covers 61 square kilometers with mixed land use (NTU, 2008) where the residential areas (R) (19%) are distinguished by a high-density population in high-rise buildings and farming areas (F) (5%) are characterized by horticultural and agricultural activities including small scale production of flowers, vegetables, ornamental fish, and chicken eggs (NTU, 2008). The undeveloped area (U), the largest of the land-use categories covering 76% of the catchment, is maintained as limited-access land dominated by native vegetation. Stormwater and sewage are transported via separate conveyance systems. An underground system conveys sewage to wastewater treatment plants, while stormwater and surface runoff are drained by open concrete-lined channels that discharge into rivers and reservoirs. Some of the farming areas are served by on-site sewage and wastewater treatment systems, while the rest of the

farming areas and residential area is served by the underground sewerage system. The climate is tropical with temperature ranging from highs of 29 to 31°C during the day to lows of 23 to 24°C at night (NTU, 2008). Catchment samples were collected from open concrete-lined channels conveying water to the reservoir. Reservoir water samples were collected at four stations approximately 800m and 1,200m apart to provide spatial coverage of the reservoir surface.

3.2.1.2 DNA extraction and quantification of HF183, *E. coli* and Total coliforms

Particulates from water samples were concentrated onto 0.2-µm-pore-size cartridge filters (Millipore, Billerica, MA, USA), subjected to extraction of environmental DNA, and quantification of the *Bacteroidales* HF183 marker by qPCR as previously described (Nshimiyimana et al., 2014). Enumeration of FIB (*E. coli*, Total Coliforms) was performed on water samples by the most-probable-number (MPN) method (IDEXX Laboratories, Inc., Westbrook, ME, USA) (Nshimiyimana et al., 2014).

3.2.2 Illumina sequencing

3.2.2.1 Library preparation

Environmental DNA was used as template for PCRs for Illumina library preparation targeting the V3 to V4 16S rRNA region as described in Preheim et al. (Preheim et al., 2013) with modification of primers (Table 1). Briefly, the 16S rRNA gene was amplified using Taq polymerase (New England BioLabs® Inc, Ipswich, MA, USA) in 20 µl of reaction volume containing 100 µM each of primers 357F (Dalwai et al., 2007) and 806R (Li et al., 2012), 10 mM dNTPs, 50 mM MgCl₂, and bovine serum albumin (BSA). To avoid cycling templates past the mid-log phase and to normalize template

concentrations, samples were subjected to Real-time PCR to determine the optimum PCR cycles for library construction (15 to 27 cycles). Illumina adaptors and barcodes were added as previously described (Preheim et al., 2013) (Table 1). Barcoded PCR products at the predicted size of 550-650 bp were gel purified (QIAQuick Gel extraction kit, QIAGEN®, Valencia, CA, USA) and sequenced using the Illumina MiSeq platform at the MIT BioMicro Center (Cambridge, MA, USA). All DNA sequences generated in this study have been deposited in Genbank KX967493 to KX976459 are currently being assigned, this will be updated during revision).

Table 1. List of primers used for qPCR and library preparation for Illumina sequencing

Primers	References
16S rRNA PCR	(Li et al., 2012)
• Primer 357F CTCCTACGGGAGGCAGCAG • Primer 806R GGACTACHVGGGTWTCTAAT	and (Dalwai et al., 2007)
Universal Overhang PCR Primers	
• Primer 357F – PE1 ACACGACGCTCTTCCGATCTYRYRCTCCTACGGGA GGCAGCAG • Primer 806r – PE1 CGGCATTCCTGCTGAACCGCTCTTCCGATCTGGAC TACHVGGGTWTCTAAT	(Dalwai et al., 2007) and (Li et al., 2012)
Adaptor/Barcode PCR	
• Primer 357F_PE2 AATGATACGGCGACCACCGAGATCTACA • Primer 806R_PE2 CAAGCAGAAGACGGCATACTGAGATXXXXXXCGGT CTCGGCATTCCTGCTGAACCGCTCT	
Note: XXXXXX: 6bp of Barcode	
QPCR:	(Bernhard and Field, 2000) and (Seurinck et al., 2005)
• HF183 marker ATCATGAGTTCACATGTCCG • Bac242R Bacteroidales 16S rRNA TACCCCGCCTACTATCTAATG	

3.2.2.2 Next generation 16S rRNA gene sequencing

Base-calling and quality filtering were implemented by Illumina MiSeq software to generate FASTQ files containing sequences and quality scores. Resulting FASTQ files were demultiplexed based on barcode sequence and were processed through the UPARSE pipeline for additional quality control and identification of operational taxonomic units (OTUs) at 97% cluster width (Edgar, 2013). Overlapping regions of each paired-end sequence were merged to create a single read. Sequences were then quality filtered by adjusted Q score, globally trimmed to 400bp (sequences shorter than 400bp were discarded), and were de-replicated. Following OTU clustering all singleton sequences were discarded per recommended settings and chimeric sequences were identified using UCHIME (Edgar, 2013). OTUs were taxonomically classified based on representative sequences (cluster centroids) from kingdom to species using Silva ARB software (Quast et al., 2013) with 60 % assignment cut off. Trimmed and filtered sequences were mapped to OTUs to create a matrix of sequence abundance.

3.2.3 Identification of pathogen-like sequences (PLS) and sewage-associated taxa (SAT)

OTUs were screened to identify genera and species corresponding to human etiological agents as indicated by the US National Institute of Health (US NIH) (NHI, 2013), the Pathosystems Resources Integration Center (PATRIC) in collaboration with the National Institute of Allergy and Infectious Diseases (NIAID) (Wattam et al., 2014), and a database of emerging infectious diseases (Jones et al., 2008). In addition, all OTUs assigned to pathogen-bearing genera were screened for species-level relatedness to potential bacterial pathogens obtained from clinical specimens associated with

human disease using BlastN with the criterion of $\geq 99\%$ sequence identity where the best-hit sequence was confirmed via BLAST distance-based clustering. Sewage-associated taxa (SAT) were identified as OTUs that were identified as a genus that has previously been determined to associate with sewage by McClelland and co-workers (McLellan et al., 2010) and those OTUs shared by two municipal sewage samples from Singapore with annotations indicating they were derived from sewage or the human gut (Supplemental Table 3).

3.2.3 Data Analysis

Multivariate analysis of bacterial community composition (BCC) and the diversity of pathogen-like sequences (PLS) and sewage-associated taxa (SAT) was conducted in PERMANOVA+ for Plymouth Routines In Multivariate Ecological Research (PRIMER) V7 (Clarke and Gorley, 2006). Principal Coordinate Analysis (PCO) and Multidimensional Scaling (MDS) of Bray-Curtis similarity were used to identify samples with similar community composition. Bacterial community members governing variation at phyla and OTU levels were examined using Pearson correlation to PCO axes and visualized using the vector overlay function. Permutational multivariate analysis of variance (PERMANOVA) was used to explore how BCC, SAT, and PLS varied with land use or sampling dates and was implemented for OTUs (BCC, PLSg, and PLSs) or genera (SAT). Bray-Curtis indexes of BCC from the same or different sites on different dates were compared by two-tail t-test with unequal variance. The Draftsman plot tool was used to visualize multi-collinearity among Total Coliform MPN, *E. coli* MPN, and HF183 following a $\log_{10}$ transform and variables with Pearson correlation coefficient of $R < 0.5$ were retained for further analysis. The relationship between bacterial community composition (BCC) and $\log$ *E. coli* MPN or $\log$ HF183

GE/100ml was determined using the BIONEVE best selection procedure routine with AIC (Akaike information criterion) as the selection criterion based on 999 permutations in PERMANOVA+. The variation in BCC, SAT, or PLS explained by the abundance of the log HF183 marker or log *E. coli* MPN was determined by distance-based linear modeling (DistLM) routine implementing the AIC selection criteria and Best procedure (Anderson et al., 2008). The Marginal test of DistML was applied to determine the extent of BCC, SAT, and PLS variations explained by log *E. coli* MPN/100ml or log HF183 GE/100ml. Similarity Percentages (SIMPER) were used to identify OTUs (BCC and PLS species level) and genera (SAT and PLS genus level) contributing to the similarity or dissimilarity of bacterial groups sampled in the catchment and reservoir. SIMPER was calculated by decomposing average Bray-Curtis dissimilarity between all pairs of samples into percentage contributions from each OTU or genus.

RStudio (RStudio, 2012) software (R version 3.2.4) was used to generate a rarefaction graph through the permute, lattice, and vegan packages (Gotelli and Colwell, 2001, Buddle et al., 2005, Work et al., 2010, Jacobs, 2011). The distribution of log₁₀-transformed indicator bacteria (*E. coli* and total coliforms) and HF183 marker levels across samples was examined by Pearson's correlation and hierarchical clustering using Ward's method on standardized data (JMP Pro v.12). The extent to which log *E. coli* MPN, log HF183, land use, and sample date accounted for variability in the proportion of PLS or SAT sequences observed in the dataset was modeled using Beta-regression implemented via maximum likelihood in JMP Pro v.12 (SAT Institute Inc., Cary, NC, USA). The statistical significance of individual predictors was assessed via the Wald Chi Squared test.

3.3. Results

3.3.1 Bacterial Community Composition (BCC) in an Urban Reservoir and Catchment

A total of 3,810,864 paired-end Illumina MiSeq reads were quality filtered and overlapping paired ends were merged into 1,189,972 sequences ranging from 17,986 to 67,583 sequences per sample (also referred to herein as “reads”). Sequences were mapped onto 8,967 OTUs using the UPARSE pipeline (Table 2). All OTUs were classified according to phylum and 96.4%, 94.0%, 89.3%, and 67.0% were classified to class, order, family, and genus, respectively. Overall, sequences from the Proteobacteria phyla dominated most samples (57% of sequences) followed by Bacteroidetes (16%), Cyanobacteria (9%), Firmicutes (6%), and Actinobacteria (4%) (Figure 1A).

Rarefaction analysis of OTU richness indicated that, despite > 17,000 reads per sample, most sites were not sampled to saturation suggesting undiscovered diversity (Figure 1B). Samples collected from the same site on two different dates tended to harbor bacterial communities that were most similar to each other based on Bray Curtis similarity indexes (0-100) (intra-site similarity = 34 ± 8 versus inter-site similarity = 15.3 ± 0.6 ; $p = 0.002$ (Supplemental Table 1).

3.3.1.1 Bacterial Community Composition (BCC) Variation with Reservoir and Catchment Land Use

To examine the hypothesis that bacterial communities would vary among reservoir and catchment and within catchment land uses, principal coordinate analysis was applied.

Bacterial phyla in samples grouped into three clusters by PCO where all reservoir samples clustered away from samples collected from the catchment and samples R12_7, F8_1, F8_7, F5_1, and F5_7 formed a separate cluster (Figure 2A). Clustering was supported by non-metric multidimensional scaling (NMDS) (data not shown). Spearman correlation of individual phyla to PCO axes indicates that the cluster containing reservoir samples was enriched in sequences from Cyanobacteria, Planctomycetes, Chlorobi, Bacteroidetes, and Chloroflexi ($R=0.65$ to 0.9 to PCO1), while the cluster containing most of the residential sites was enriched in Proteobacteria ($R=0.50$ to PCO2) and Firmicutes ($R=-0.98$ to PCO1) were enriched in the horticultural sites, particularly F8_7 and F10_7 (Figure 2A).

Two-way PERMANOVA revealed that the distribution of bacterial OTUs varied significantly between reservoir and catchment ($p=0.001$, $F=8.7$) (Figure 2A-B) and among catchment land uses ($p=0.009$, $F=2$) but did not vary between months of sample collection ($p=0.16$, $F=1.2$) (Supplemental Table 2). SIMPER analysis indicated enrichment of reservoir samples in OTUs including *Flexibacter* OTU0036 and OTU0026 and Chitinophagaceae spp. OTU0031 while catchment samples were enriched in *Arcobacter* OTU0007, *Hydrogenophaga* OTU0032, and OTUs *12up* OTU0006 (Supplemental Table 5). Within the catchment, residential sites were enriched in sequences from *Novispirillum* OTU0011, Cyanobacterial spp. OTU0002, and *12up* OTU0006, while horticultural sites contained more sequences from *Tolumonas* OTU0014, *Prevotella* OTU2638, and Aeromonadaceae spp. OTU5002 (Supplemental Table 5).

Table 2. Distribution of bacterial OTUs and concentrations of water quality indicators in sampled land uses

Criteria	Land uses and sample collected per each category				Total sequences
	Horticultural (n=14)	Residential (n=14)	Undeveloped (n=2)	Reservoir (n=6)	
Geometric mean of human marker HF183 (Copies/ng genomic DNA)	57.0	9.0	3	9.3	
Geometric mean of human marker HF183 (GE/100ml)	8.0x10 ⁴	2.3x10 ³	7.0x10 ³	7.7x10 ³	
Geometric mean of <i>E. coli</i> (CFU/100ml)	1.7x10 ⁴	1.4x10 ²	1.1x10 ²	1.0	
Bacteria (All OTUs)	395,940	482,476	85,314	226,242	1,189,972
Pathogen-like Sequences (PLS) genus	41,334	31,065	1,890	1,398	75,687
Pathogen-like Sequences (PLS) species	15,145	11,032	780	857	27,814
Sewage-Associated Taxa (SAT) level (McLellan et al. 2010)	23,974	5,558	352	203	30,087
Total OTUs	Number of all OTUs from sequencing: 9,205; 8,967 OTUs of the Domain of Bacteria were used in this study; 238 OTUs not included were taxonomically unassigned or belonged to Archaea or Eukaryote				

3.3.1.2 Relationship between BCC, Fecal Indicator Bacteria (FIB), and HF183

To examine the hypothesis that sites with elevated FIB would harbor distinct bacterial communities, the relationship between BCC, FIB (*E. coli*, total coliform), and HF183 was examined in the reservoir and catchment samples. Total coliform was highly co-

linear with *E. coli* ($R=0.84$), thus *E. coli* was used to represent both in subsequent analysis.

A combination of log *E. coli* MPN and log HF183 GE/100ml showed a moderate correlation to BCC (BIOENV spearman correlation $R=0.48$), while their DistLM (Distance Linear Model) revealed a combined cumulative variance of 25.3% ($R^2=0.25$). The marginal test of DistLM showed that *E. coli* explained more significant variation in BCC at the OTU level than HF183 (*E. coli*, variation of 12%, $p=0.001$ compared to HF183, variation of 5.5%, $p=0.016$) (Table 3).

3.3.2 Relationship between Human *Bacteroides* Marker (HF183) and Proportion of *Bacteroides* and *Bacteroides dorei* Sequences.

Since the 16S rRNA gene sequence for the human-associated *Bacteroidales* HF183 marker and the species of *Bacteroides dorei* (Bernhard and Field, 2000, Bakir et al., 2006) matched the sequence of OTU45 in this study we examined the extent to which sequence distribution matched QPCR quantification. Moderate correlation between the relative abundance of *Bacteroides dorei* OTU45 and QPCR-measured HF183 was observed (log HF183 copies/100ml; $R=0.57$, $p=0.0002$) (Figure 5). The proportion of reads corresponding to *Bacteroides* (genus level) or *Bacteroides dorei* (OTU45) was highest for horticultural sites F5, F8, and F10 and lowest in the reservoir consistent with trends previously observed for HF183 (Nshimiyimana et al., 2014). The proportion of sequences corresponding to *Bacteroides dorei* OTU45 was moderately correlated with *E. coli* ($R=0.55$, $p=0.0005$) (Figure 5).

Figure 1. (A) Highly represented phyla across samples collected from the reservoir and catchment. Sample codes indicate land use: R=Residential, U=Undeveloped, K=Reservoir, F=Horticultural/Farming, sample number, and collection date “_1” identifies samples collected January 2009, and the rest were collected in July 2009. (B) Rarefaction analysis of species richness in individual samples. Line color corresponds to land use: Red=Reservoir, Blue=Residential, Cyan=Undeveloped, Green=Horticultural, and Purple = Reference samples. Reference samples 114_Sw and 115_Sw were collected January 2010

3.3.3 Distribution and Variation of Sewage-Associated Taxa (SAT)

To examine the hypothesis that sites with elevated levels of FIB would harbor signatures of sewage-associated taxa (SAT), we identified 30,087 reads (n=15 genera) (Supplemental Table 5) that corresponded to OTUs shared by sewage samples analyzed as part of this study (n=2) or to bacterial groups proposed by McLellan et al. (2010) as associated with human fecal pollution. Sewage-associated sequences from this study shared substantial overlap with SAT described by McLellan et al. (2010) including the shared genera *Bifidobacterium*, *Bacteroides*, *Parabacteroides*, *Faecalibacterium*, *Roseburia*, *Ruminococcus*, *Akkermansia*, *Subdoligranulum*, *Papillibacter*, and *Sutterella* (Supplemental Table 3)

Two-way PERMANOVA of Bray-Curtis resemblances among different samples revealed that SAT composition varied with catchment land use ($p = 0.004$) (Figure 2C, Supplemental Table 2) but not collection month ($p=1$) or the interaction of land use and collection month ($p=0.5$) (Supplemental Table 2). SIMPER analysis demonstrated that the distinction between horticultural and residential land uses was driven by the SAT *Prevotella*, *Faecalibacterium*, and *Bifidobacterium* enriched in horticultural areas and *Papillibacter* enriched in residential areas (Supplemental Table 5). The composition of SAT was moderately correlated to measured levels of *E. coli* and weakly correlated with HF183 levels (BIOENV: *E. coli*, $R=0.55$ and HF183, $R=0.1$). DistLM analysis of *E. coli* and HF183 demonstrated a combined cumulative variance of 37.6% ($R^2=0.37$) across SAT composition. The DistLM marginal test showed that *E. coli* explained more variation in the composition of SAT than HF183 (*E. coli*: 22.3%, $p=0.001$ compared to HF183: 14.4%, $p=0.001$) (Table 3)..

Table 3. Contribution of *E. coli* or HF183 to variance in bacterial community structure, SAT; PLSg, and PLSs

Variables	Combined <i>E. coli</i> and HF183 Contributions				Individual <i>E. coli</i> and HF183 Contributions	
	R ²	RSS	Cumulative variation	Pseudo-F*	P (Value)*	% of Variation*
Bacterial Community Structure (OTUs)						
Log10 HF183 (GE/100ml)	0.25	8,539	25.3%	2.0	0.016	6.0
Log10 <i>E. coli</i> (MPN/100ml)				5.0	0.001	12.1
Sewage-associated taxa (SAT)						
Log10 HF183 (GE/100ml)	0.30	49,150	30%	5.4	0.001	14.4
Log10 <i>E. coli</i> (MPN/100ml)				10	0.001	23.3
Pathogen-like sequences at genus level (PLSg)						
Log10 HF183 (GE/100ml)	0.20	41,626	20.7%	2.0	0.065	5.0
Log10 <i>E. coli</i> (MPN/100ml)				6.2	0.001	15.3
Pathogen-like sequences at species level (PLSs)						
Log10 HF183 (GE/100ml)	0.24	20,623	24.5%	2.0	0.088	5.0
Log10 <i>E. coli</i> (MPN/100ml)				8.4	0.001	20.0

*Determined by Distance Linear Modeling (DistLM) Marginal test based on permutation procedures (9999 permutations)

Figure 2. Multivariate analysis of bacterial community composition based on Principal Coordinate Analysis (PCO) of Bray-Curtis resemblance between samples. Sample codes indicate land use: R=Residential, U=Undeveloped, K=Reservoir, and F=Horticultural/Farming. Bacterial communities are distinguished by (A) bacterial phyla (B) OTUs, (C) SAT OTUs, and (D) PLS OTUs in the catchment (horticultural, residential, and undeveloped) and reservoir sites. (A) Individual bacterial phyla contributing to variation were determined by Spearman correlation ($R > 0.65$) to the first two PCO axis and are represented by vectors.

SAT sequences were observed to constitute the greatest proportion of sequences in horticultural areas (1% to 16.5% of total reads; n=14) and the lowest proportion in the reservoir (<1% of total reads; n=6) (Table 2 and Figure 3A, 6A). The proportion of SAT was moderately correlated to measured levels of *E. coli* and HF183 ($R=0.62$ to 0.65 , $p=0.001$) (Figure 5). The proportion of sequences from SAT in relation to log *E. coli* MPN, log HF183, land use, and sample dates was modeled by Beta-regression. The model for SAT ($R^2 = 0.66$) indicated that log *E. coli* MPN was the strongest predictor variable for the proportion of sewage-associated taxa observed in the bacterial community ($\chi^2 = 14.0$, $p < 0.0002$) followed by log HF183 abundance ($\chi^2 = 13.1$, $p = 0.0003$) and collection month ($\chi^2 = 7.6$, $p = 0.0057$) while land use did not emerge as a significant predictor ($p > 0.2$) upon controlling for other variables.

3.3.4 Distribution and variation of Pathogen-like Sequences

To examine the hypothesis that samples with elevated levels of FIB would also harbor signatures of potential human pathogens, we classified sequences as pathogen-like based on named genus or species-level identity to known or emerging pathogens by BlastN. Out of 75,687 sequences, 6.3% were classified to 33 genera harboring known pathogens (PLSg) (Table 2) while 2.3% of sequences matched pathogens at the species level (PLSs). The most highly represented PLSg were *Acinetobacter* (38%), *Arcobacter* (22%), *Pseudomonas* (8.2%), *Aeromonas* (7.4%), and *Clostridium* (7%). Samples with the highest and lowest contribution from PLSg sequences were respectively F10_7 (11%) and K4_1 (0.2%). PCO analysis indicated that the composition of PLSg and PLSs in samples clustered distinctly with reservoir or catchment origin (Figure 2D and 4). The two-way PERMANOVA of Bray-Curtis resemblances among samples showed

that PLS composition varied with catchment land use and collection month ($p < 0.05$) (Supplemental Table 2).

SIMPER analysis indicated that genera *Arcobacter*, *Aeromonas*, and *Acinetobacter* were enriched in the catchment while the PLS genera *Legionella*, *Leptospira*, and *Rickettsia* were enriched in the reservoir (Supplemental Table 5). Within the catchment PLSs, OTUs enriched in horticultural areas were classified as *Arcobacter cryaerophilus* (OTU6394 and OTU0016), *Clostridium disporicum* (OTU382), and *Pseudomonas mendocina* (OTU25). OTUs identified as *Acinetobacter junii* (OTU0005), *Aeromonas tecta* (OTU0163), and *Yersinia enterocolitica* (OTU0440) were enriched in residential areas (Supplemental Table 4 and Supplemental Table 5 and Supplemental Figure 1). DistLM analysis of *E. coli* and HF183 revealed a combined cumulative variance of 21% ($R^2 = 0.20$) across PLS composition. The DistLM marginal test indicated that log *E. coli* MPN could explain significant variation in the composition of PLSg and PLSs (15.3% for PLSg and 20% for PLSs, $p < 0.002$) while log HF183 accounted for 5% of the variability ($p \leq 0.08$) (Table 3).

Sites with the highest proportion of PLS were identified in horticultural and residential areas (Table 2, Figure 4 and 6). The lowest proportion PLS were observed in the reservoir and undeveloped sites (Table 2, Figures 4 and 6). Significant correlations were observed between the abundance of *E. coli* and proportions of sequences corresponding to PLSg and PLSs ($R = 0.6$, $p = 0.0001$) (Figure 5). In contrast, HF183 abundance was not significantly correlated with PLSg and PLSs ($R = 0.02$ to 0.08 , $p = 0.07$ to 0.90). The proportion of PLSs and PLSg in relation to log *E. coli* MPN, log HF183, land use, and sample date was modeled by Beta-regression. The model for PLSs had a poor fit and

was not considered further ($R^2 < 0$). The model for PLSg ($R^2 = 0.57$) indicated that log *E. coli* MPN was the strongest predictor variable ($\chi^2 = 8.79$, $p = 0.003$) followed by sample date ($\chi^2 = 4.249$, $p = 0.0393$) while log HF183 abundance and land use did not emerge as significant predictors ($p > 0.1$)

Figure 3. Hierarchical clustering of the distribution of SAT (genera). Comparison includes proportion of PLSg, PLSs, E. coli, and HF183. Heat map indicate sequence abundance as a fraction of the total sequences per sample and are individually scaled. E. coli and HF183 concentrations are log10 transformed.

Figure 4. Hierarchical clustering of the distribution of eighteen genera for which OTU sequences match human pathogens at the species level (16S rRNA identity $\geq 99\%$). Comparison includes proportion of SAT, PLSSs, *E. coli*, and HF183. Heat map indicate sequence abundance as a fraction of the total sequences per sample and are individually scaled. *E. coli* and HF183 concentrations are log10 transformed

Figure 5. Draftsman plot of quantities considered in this study: Log HF183 (GE/100ml and Copies/ng), *E. coli* (MPN/100ml), and proportions of sequences corresponding to SAT, PLSg, PLSs, and *B. dorei* OTU45. Significant Pearson correlations ($p < 0.05$) are shown by the red asterisk (*)

Figure 6. The distribution of sequences from (A) sewage-associated taxa (SAT) and (B, C) pathogen-like sequences identified to genus and species level (PLSg and PLSs, respectively) as a percentage of the total bacterial community composition for different land-use categories sampled in this study. Box and whiskers correspond to the median and quartile distribution. Tukey-Kramer Honest Significant Difference Test (HSD) and ANOVA analysis demonstrated variation of % SAT, % PLSg, and % PLSs across catchment land uses and reservoir. Letters of HSD test (A and B) indicate that catchment land use not represented by same letters have significantly different SAT, PLSg and PLSs.

3.3.5 Discussion

This study examined whether the bacterial community composition and the distribution of sequences from sewage-associated taxa and pathogen-like bacteria vary in relation to indicators of water quality indicators (i.e. Total coliform, *E. coli*, HF183 marker) and land use in an urban tropical catchment in Singapore. I hypothesized that sites with elevated levels of FIB would be associated with sequences from sewage and potential pathogens. Indeed, we observed that catchment samples with elevated levels of FIB were more likely to have higher proportions of sewage-associated and pathogen-like sequences classified at the genus and species levels. In a previous study multiple horticultural sites were observed to harbor elevated levels of FIB and the HF183 marker (37). During field sampling, these horticultural sites were noted to receive discharges from poorly operating septic tank systems that combined sewage with other wastewater sources (Nshimiyimana et al., 2014). Samples collected from these sites revealed a greater proportion of SAT and pathogen-like sequences (PLS genus and PLS species) (sites F5, F6, F8, F9, and F10) (Figure 3 and 4 and Figure 6 A-C and Supplemental Figure 1). The recovery of sequences from bacteria related to agents of gastro-intestinal tract disease (i.e. *Shigella sonnei*, *Clostridium perfringens*, *Arcobacter* *cryaerophilus*, *Campylobacter helveticus*, and *Yersinia enterocolitica*; Supplemental Table 4) from sites with elevated FIB raises the possibility that sewage discharge may have been associated with introduction of these human pathogens to the catchment (Ashbolt, 2015, Arias et al., 2006) (Supplemental Figure 1, Supplemental Table 4).

Overall, the reservoir was found to have the lowest measured concentrations of FIB and HF183 marker and also the lowest proportion of sequences from SAT and PLS (Table

2, Figure 3, 4 and 6A-C and Supplemental Figure 1). *Legionella* spp. sequences were enriched in the reservoir relative to catchment samples consistent with their ecology as ubiquitous endemic freshwater bacteria that may parasitize freshwater amoeba (Muder et al., 2002, Whiley et al., 2014). Sequences from *Rickettsia* spp., thought to be obligate intracellular parasites, were also enriched in reservoir samples, potentially reflecting life-cycle associations with amphibians, reptiles, or aquatic eukaryotes including protozoa and aquatic arthropods, as has been recently suggested (Weinert et al., 2009, Andoh et al., 2015, Parola et al., 2005). These findings suggest that microbial water quality in the reservoir may be influenced by both pollution from the catchment and endemic pathogens. Low concentrations of FIB, PLS, and SAT in the reservoir despite their comparatively high level of detection in the catchment suggests that ecological and physical processes within the reservoir may be effective at removing FIB and potential pathogens introduced from catchment sources.

The DNA-based methods employed in this study provide cultivation-independent insight into diversity and distribution of microorganisms in environmental samples, however certain limitations must be kept in mind during data interpretation. First, DNA-based methods are limited in their ability to distinguish live from dead cells and cannot provide information on the strain-specific virulence of pathogens that may govern their actual risk to human health. Second, there are conflicting reports on the ability of amplicon sequence-based studies to provide quantitative information regarding the relative abundance of sequences (Tan et al., 2015) and PCR may introduce biases skewing the proportion of sequences. In this study, the same PCR protocol was applied across samples enabling cross comparison among the samples despite assumed biases and we note that a good correlation ($R = 0.6$, $P=0.0002$, Figure

5) was observed between the relative distribution of sequences from OTU45 (*Bacteroides dorei*) and HF183 marker that quantifies *Bacteroides dorei* DNA by qPCR (Figure 5).

Our study suggests that sites with lower predicted water quality based on high measured levels of *E. coli* MPN may have a greater likelihood of harboring sewage-associated and pathogen-like bacteria. Similarly, a recent study that screened for potential pathogens by pyrosequencing of the V1 and V2 hypervariable regions of the SSU rRNA gene along the Santa Ana River, CA, USA demonstrated that sites with higher levels of human impact were more likely to yield sequences related to human pathogens (Ibekwe et al., 2013). Many previous studies have linked high counts of FIB and detection of HF183 to the presence of sewage and pathogen contamination in water (Staley et al., 2012, McQuaig et al., 2006b, Harwood et al., 2014). Interestingly, while the abundance of the HF183 human-associated *Bacteroidales* marker was predictive for the proportion of sewage-associated sequences detected, our data do not indicate that HF183 is predictive for enrichment of human pathogen-like sequences, pointing to the possibility that the majority of PLS observed in this study do not have a human origin. Alternatively, as suggested by Nshimiyimana et al. (Nshimiyimana et al., 2014) the protocol for quantification of HF183 using Sybr Real Time PCR may have unknown cross-reactivity with closely related environmental sequences. For this reason ongoing monitoring studies now employ the more specific TaqMan HF183 assay (Green et al., 2014). Significant correlation between *E. coli* and both the composition and proportion of SAT and PLS (Table 3 and Figure 5) together with observation of elevated PLS, SAT, HF183, and FIB at sites with suspected sewage contamination suggest that *E. coli* concentrations may indeed be useful to predict samples enriched in sewage-associated

bacteria including potential human pathogens in tropical waters. Further work is needed to determine if threshold levels of *E. coli* associated with risks from sewage exposure in tropical areas may be higher than temperate environments after accounting for differences in environmental baseline levels (Hazen, 1988, Rivera, 1988).

3.3.6 Conclusion

We have shown that profiling of BCC, including SAT and PLS using next-generation sequencing affords insights into the bacterial community at sites characterized by different levels of water quality indicators. This information provides direct insight into the agents that may potentially mediate public health risk at sites with high FIB. At the present time, the relationship between BCC, percent pathogen-like sequences, FIB levels, and human health risks is unknown. We suggest that detection and profiling of pathogen-like bacteria based on sequencing approaches is a promising tool for assessment of microbial populations associated with impaired water quality.

References: Chapter 3

- ANDERSON, M. J., R.N., G. & CLARKE, K. R. 2008. *PERMANOVA+for PRIMER: Guide to software and statistical methods*, Plymouth, PRIMER-E Ltd.
- ANDOH, M., SAKATA, A., TAKANO, A., KAWABATA, H., FUJITA, H., UNE, Y., GOKA, K., KISHIMOTO, T. & ANDO, S. 2015. Detection of *Rickettsia* and *Ehrlichia* spp. in Ticks Associated with Exotic Reptiles and Amphibians Imported into Japan. *PLoS ONE*, 10, e0133700.
- ARIAS, C., SALA, M R., DOMI NGUEZ, A., BARTOLOME, R., BENAVENTE, A., VECIANA, P., PEDROL, A., HOYO, G. & THE OUTBREAK WORKING, G. 2006. Waterborne epidemic outbreak of *Shigella sonnei* gastroenteritis in Santa Maria de Palautordera, Catalonia, Spain. *Epidemiology and Infection*, 134, 598-604.
- ASHBOLT, N. J. 2015. Microbial Contamination of Drinking Water and Human Health from Community Water Systems. *Curr Environ Health Rep*, 2, 95-106.
- BAE, S. & WUERTZ, S. 2012. Survival of host-associated bacteroidales cells and their relationship with *Enterococcus* spp., *Campylobacter jejuni*, *Salmonella enterica* serovar *Typhimurium*, and adenovirus in freshwater microcosms as measured by propidium monoazide-quantitative PCR. *Applied and Environmental Microbiology*, 78, 922-32.
- BAKIR, M. A., SAKAMOTO, M., KITAHARA, M., MATSUMOTO, M. & BENNO, Y. 2006. *Bacteroides dorei* sp. nov., isolated from human faeces. *International Journal of Systematic and Evolutionary Microbiology*, 56, 4.
- BERNHARD, A., GOYARD, T., SIMONICH, M. & FIELD, K. 2003. Application of a rapid method for identifying fecal pollution sources in a multi-use estuary. *Water Research*, 37, 909-913.
- BERNHARD, A. E. & FIELD, K. G. 2000. A PCR Assay To Discriminate Human and Ruminant Feces on the Basis of Host Differences in *Bacteroides-Prevotella* Genes Encoding 16S rRNA. *Applied and Environmental Microbiology*, 66, 4571-4574.
- BUDDLE, C. M., BEGUIN, J., BOLDUC, E., MERCADO, A., SACKETT, T. E., SELBY, R. D., VARADY-SZABO, H. & ZERAN, R. M. 2005. The importance and use of taxon sampling curves for comparative biodiversity research with forest arthropod assemblages. *The Canadian Entomologist*, 137, 7.
- BYAPPANAHALLI, M. N., NEVERS, M. B., KORAJKIC, A., STALEY, Z. R. & HARWOOD, V. J. 2012a. *Enterococci* in the environment. *Microbiology and Molecular Biology Reviews*, 76, 685-706.

- BYAPPANAHALLI, M. N., ROLL, B. M. & FUJIOKA, R. S. 2012b. Evidence for occurrence, persistence, and growth potential of *Escherichia coli* and *enterococci* in Hawaii's soil environments. *Microbes and Environments*, 27, 164-70.
- CAI, L., JU, F. & ZHANG, T. 2014. Tracking human sewage microbiome in a municipal wastewater treatment plant. *Applied Microbiology and Biotechnology*, 98, 3317-3326.
- CAO, Y., VAN DE WERFHORST, L. C., DUBINSKY, E., BADGLEY, B., SADOWSKY, M., GRIFFITH, J. F., ANDERSEN, G. & HOLDEN, P. 2013. Evaluation of molecular community analysis methods for discerning fecal sources and human waste. *Water Research*, 47, 6862-72.
- CAO, Y., VAN DE WERFHORST, L. C., SERCU, B., MURRAY, J. L. & HOLDEN, P. A. 2011a. Application of an integrated community analysis approach for microbial source tracking in a coastal creek. *Environmental Science and Technology*, 45, 7195-201.
- CAO, Y., WU, C. H., ANDERSEN, G. L. & HOLDEN, P. A. 2011b. Community analysis-based methods. In: HAGEDORN, C., BLANCH, A. R. & HARWOOD, V. J. (eds.) *Microbial Source Tracking: Methods, Applications, and Case Studies*. New York: Springer.
- CARRILLO, M. 1985. Survival and enumeration of the fecal indicators *Bifidobacterium adolescentis* and *Escherichia coli* in a tropical rain forest watershed. *Applied and Environmental Microbiology*, 50, 468-476.
- CHUA, L. H. C., SHANAHAN, P., EDMOND, Y. M. L., SHUY, E. B., THOMPSON, J., DIXON, C. C., KERIGAN, K. K., NSHIMYIMANA, J. P., YEAGER, J. M., LEE, L. J. & POR, Y. L. 2010. Dry Weather bacteria monitoring and variation with land use for Kranji Reservoir catchment, Singapore. *17th Congress of the Asia and Pacific Division of the International Association of Hydraulic Engineering and Research*. Auckland, New Zealand.
- CLARKE, K. R. & GORLEY, R. N. 2006. *PRIMER v6: user manual*, Plymouth, PRIMER-E Ltd.
- DALWAI, F., SPRATT, D. A. & PRATTEN, J. 2007. Use of Quantitative PCR and Culture Methods To Characterize Bacterial Biofilms. *Clinical Microbiology*, 45, 4.
- EDGAR, R. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, 10, 996.

- EKKLESIA, E., SHANAHAN, P., CHUA, L. H. C. & EIKAAS, H. S. 2015. Temporal variation of faecal indicator bacteria in tropical urban storm drains. *Water Research*, 68, 10.
- FUJIOKA, R. S. 2001. Monitoring coastal marine waters for spore-forming bacteria of faecal and soil origin to determine point from non-point source pollution. *Water science and technology*, 44, 181-188.
- GOTELLI, N. J. & COLWELL, R. K. 2001. Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters*, 3, 12.
- GREEN, H. C., HAUGLAND, R. A., VARMA, M., MILLEN, H. T., BORCHARDT, M. A., FIELD, K. G., WALTERS, W. A., KNIGHT, R., SIVAGANESAN, M., KELTY, C. A. & SHANKS, O. C. 2014. Improved HF183 Quantitative Real-Time PCR Assay for Characterization of Human Fecal Pollution in Ambient Surface Water Samples. *Applied and Environmental Microbiology*, 80, 9.
- GREGORY, J. B., LITAKER, R. W. & NOBLE, R. T. 2006. Rapid One-Step Quantitative Reverse Transcriptase PCR Assay with Competitive Internal Positive Control for Detection of Enteroviruses in Environmental Samples. *Applied and Environmental Microbiology*, 72, 3960-3967.
- HARWOOD, V., STALEY, C., BADGLEY, B., BORGES, K. & KORAJKIC, A. 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews*, 38, 1-40.
- HAZEN, T. C. 1988. Fecal coliforms as indicators in tropical waters: A review. *Toxicity Assessment*, 3, 461-477.
- IBEKWE, A. M., LEDDY, M., MURINDA, S. & AZIZ, R. 2013. Potential Human Pathogenic Bacteria in a Mixed Urban Watershed as Revealed by Pyrosequencing. *PLoS ONE*, 8, e79490.
- JACOBS, J. 2011. *Individual Based Rarefaction using R-package* [Online]. The R Project for Statistical Computing. Available: <http://www.jennajacobs.org/R/rarefaction.html> [Accessed July 2014].
- JIMENEZ, L. 1989. Survival and activity of *Salmonella typhimurium* and *Escherichia coli* in tropical freshwater. *Journal of Applied Bacteriology*, 67, 61-69.
- JONES, K. E., PATEL, N. G., LEVY, M. A., STOREYGARD, A., BALK, D., GITTLEMAN, J. L. & DASZAK, P. 2008. Global trends in emerging infectious diseases. *Nature*, 451, 990-3.

- LAYTON, B., CAO, Y., EBENTIER, D., HANLEY, K., BALLESTÉ, E., BRANDÃO, J., BYAPPANAHALLI, M., CONVERSE, R., FARNLEITNER, A., GENTRY SHIELDS, J., GIDLEY, M., GOURMELON, M., LEE, C., LEE, J., LOZACH, S., MADI, T., MEIJER, W., NOBLE, R., PEED, L., REISCHER, G., RODRIGUES, R., ROSE, J., SCHRIEWER, A., SINIGALLIANO, C., SRINIVASAN, S., STEWART, J., VAN DE WERFHORST, L. C., WANG, D., WHITMAN, R., WUERTZ, S., JAY, J., HOLDEN, P., BOEHM, A., SHANKS, O. & GRIFFITH, J. 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. *Water Research*, 47, 6897-6908.
- LEY, V., HIGGINS, J. & FAYER, R. 2002. Bovine enteroviruses as indicators of fecal contamination. *Applied and Environmental Microbiology*, 68, 3455-61.
- LI, D., SHARP, J. O., SAIKALY, P. E., ALI, S., ALIDINA, M., ALARAWI, M. S., KELLER, S., HOPPE-JONES, C. & DREWES, J. E. 2012. Dissolved Organic Carbon Influences Microbial Community Composition and Diversity in Managed Aquifer Recharge Systems. *Applied and Environmental Microbiology*, 78, 9.
- LIANG, L., GOH, S. G., VERGARA, G. G., FANG, H. M., REZAEINEJAD, S., CHANG, S. Y., BAYEN, S., LEE, W. A., SOBSEY, M. D., ROSE, J. B. & GIN, K. Y. 2015. Alternative fecal indicators and their empirical relationships with enteric viruses, *Salmonella enterica*, and *Pseudomonas aeruginosa* in surface waters of a tropical urban catchment. *Applied Environmental Microbiology*, 81, 850-60.
- LIN, C. & TONG, Z. 2013. Detecting Human Bacterial Pathogens in Wastewater Treatment Plants by a High-Throughput Shotgun Sequencing Technique. *Environmental Science and Technology*, 47, 5433-5441.
- LIN, Y. & TONG, Z. 2011. Pathogenic Bacteria in Sewage Treatment Plants as Revealed by 454 Pyrosequencing. *Environmental Science and Technology*, 45, 7173-7179.
- LISLE, J. T. & ROSE, J. B. 1995. Cryptosporidium Contamination of Water in the USA and UK - A Minireview. *Aqua*, 44, 103-117.
- MCLELLAN, S. L. & EREN, A. M. 2014. Discovering new indicators of fecal pollution. *Trends in Microbiology*, 22, 9.
- MCLELLAN, S. L., HUSE, S. M., MUELLER-SPITZ, S. R., ANDREISHCHEVA, E. N. & SOGIN, M. L. 2010. Diversity and population structure of sewage-derived microorganisms in wastewater treatment plant influent. *Environmental Microbiology*, 12, 378-392.

- MCQUAIG, S., SCOTT, T., HARWOOD, V., FARRAH, S. & LUKASIK, J. 2006a. Detection of human-derived fecal pollution in environmental waters by use of a PCR-based human polyomavirus assay. *Applied and Environmental Microbiology*, 72, 7567-74.
- MCQUAIG, S. M., SCOTT, T. M., HARWOOD, V. J., FARRAH, S. R. & LUKASIK, J. O. 2006b. Detection of Human-Derived Fecal Pollution in Environmental Waters by Use of a PCR-Based Human Polyomavirus Assay. *Applied and Environmental Microbiology*, 72, 7567-7574.
- MCQUAIG, S. M., SCOTT, T. M., LUKASIK, J. O., PAUL, J. H. & HARWOOD, V. J. 2009. Quantification of Human Polyomaviruses JC Virus and BK Virus by TaqMan Quantitative PCR and Comparison to Other Water Quality Indicators in Water and Fecal Samples. *Applied and Environmental Microbiology*, 75, 3379-3388.
- MUDER, R. R., YU, V. L. & STRAUSBAUGH, L. J. 2002. Infection Due to *Legionella* Species Other Than *L. pneumophila*. *Clinical Infectious Diseases*, 35, 990.
- NHI 2013. Guidelines for Research Involving Recombinant or Synthetic Nucleic Acid Molecules: Classification of Human Etiologic Agents on the Basis of Hazard (Appendix B). United State National Health Institute.
- NSHIMYIMANA, J. P., EKKLESIA, E., SHANAHAN, P., CHUA, L. H. C. & THOMPSON, J. R. 2014. Distribution and abundance of human-specific *Bacteroides* and relation to traditional indicators in an urban tropical catchment. *Applied Microbiology*, 116, 1369-1383.
- NTU 2008. Water quality monitoring, modeling and management for Kranji Catchment/Reservoir system – Phases 1 and 2 / May 2004 to December 2007. Singapore: Division of Environmental and Water Resources Engineering, School of Civil and Environmental Engineering, Nanyang Technological University.
- PAROLA, P., PADDOCK, C. D. & RAOULT, D. 2005. Tick-borne rickettsioses around the world: emerging diseases challenging old concepts. *Clinical Microbiology Reviews*, 18, 719-56.
- PREHEIM, S. P., PERROTTA, A. R., MARTIN-PLATERO, A. M., GUPTA, A. & ALM, E. J. 2013. Distribution-Based clustering: using ecology to refine the operational taxonomic unit. *Applied and Environmental Microbiology*, 79, 10.
- PRÜSS, A. 1998. Review of epidemiological studies on health effects from exposure to recreational water. *International Journal of Epidemiology*, 27, 1-9.
- QUAST, C., PRUESSE, E., YILMAZ, P., GERKEN, J., SCHWEER, T., YARZA, P., PEPLIES, J., AND & GLOCKNER, F. O. 2013. The SILVA ribosomal RNA

- gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41, D590.
- RIVERA, S. C. 1988. Isolation of fecal coliforms from pristine sites in a tropical rain forest. *Applied and Environmental Microbiology*, 54, 513-517.
- ROCHELLE-NEWALL, E., NGUYEN, T. M., LE, T. P., SENGTAHEUANGHOUNG, O. & RIBOLZI, O. 2015. A short review of fecal indicator bacteria in tropical aquatic ecosystems: knowledge gaps and future directions. *Frontiers in Microbiology*, 6, 308.
- RSTUDIO 2012. Integrated development environment for R Version 0.96.122 ed. Boston, MA: RStudio.
- SALIPANTE, S. J., SENGUPTA, D. J., ROSENTHAL, C., COSTA, G., SPANGLER, J., SIMS, E. H., JACOBS, M. A., MILLER, S. I., HOOGESTRAAT, D. R., COOKSON, B. T., MCCOY, C., MATSEN, F. A., SHENDURE, J., LEE, C. C., HARKINS, T. T. & HOFFMAN, N. G. 2013. Rapid 16S rRNA next-generation sequencing of polymicrobial clinical samples for diagnosis of complex bacterial infections. *PLoS One*, 8, e65226.
- SAVICHTCHEVA, O. & OKABE, S. 2006. Alternative indicators of fecal pollution: relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. *Water Research*, 40, 2463-76.
- SAVICHTCHEVA, O., OKAYAMA, N. & OKABE, S. 2007. Relationships between *Bacteroides* 16S rRNA genetic markers and presence of bacterial enteric pathogens and conventional fecal indicators. *Water Research*, 41, 13.
- SEURINCK, S., DEFOIRDT, T., VERSTRAETE, W. & SICILIANO, S. 2005. Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environmental Microbiology*, 7, 249-59.
- SHANKS, O. C., KELTY, C. A., ARCHIBEQUE, S., JENKINS, M., NEWTON, R. J., MCLELLAN, S. L., HUSE, S. M. & SOGIN, M. L. 2011. Community structures of fecal bacteria in cattle from different animal feeding operations. *Applied and Environmental Microbiology*, 77, 2992-3001.
- STALEY, C., RECKHOW, K. H., LUKASIK, J. & HARWOOD, V. J. 2012. Assessment of sources of human pathogens and fecal contamination in a Florida freshwater lake. *Water Research*, 46, 5799-812.
- TAN, B., NG, C., NSHIMYIMANA, J. P., LOH, L. L., GIN, K. Y. & THOMPSON, J. R. 2015. Next-generation sequencing (NGS) for assessment of microbial water

- quality: current progress, challenges, and future opportunities. *Frontiers in Microbiology*, 6, 1027.
- UNNO, T., JANG, J., HAN, D., KIM, J. H., SADOWSKY, M. J., KIM, O. S., CHUN, J. & HUR, H. G. 2010. Use of barcoded pyrosequencing and shared OTUs to determine sources of fecal bacteria in watersheds. *Environmental Science and Technology*, 44, 7777-82.
- WATTAM, A. R., ABRAHAM, D., DALAY, O., DISZ, T. L., DRISCOLL, T., GABBARD, J. L., GILLESPIE, J. J., GOUGH, R., HIX, D., KENYON, R., MACHI, D., MAO, C., NORDBERG, E. K., OLSON, R., OVERBEEK, R., PUSCH, G. D., SHUKLA, M., SCHULMAN, J., STEVENS, R. L., SULLIVAN, D. E., VONSTEIN, V., WARREN, A., WILL, R., WILSON, M. J., YOO, H. S., ZHANG, C., ZHANG, Y. & SOBRAL, B. W. 2014. PATRIC, the bacterial bioinformatics database and analysis resource. *Nucleic Acids Research*, 42, D581-91.
- WEINERT, L. A., WERREN, J. H., AEBI, A., STONE, G. N. & JIGGINS, F. M. 2009. Evolution and diversity of *Rickettsia* bacteria. *BMC Biology*, 7, 6.
- WHILEY, H., KEEGAN, A., FALLOWFIELD, H. & ROSS, K. 2014. Uncertainties associated with assessing the public health risk from *Legionella*. *Frontiers in Microbiology*, 5, 501.
- WINFIELD, M. D. & GROISMAN, E. A. 2003. Role of nonhost environments in the lifestyles of *Salmonella* and *Escherichia coli*. *Applied Environmental Microbiology*, 69, 3687-94.
- WORK, T. T., JACOBS, J. M., SPENCE, J. R. & VOLNEY, W. J. 2010. Increased green-tree retention required to maintain ground beetle biodiversity in boreal mixedwood forests *Ecological Applications*, 20, 10.
- YAMAHARA, K., LAYTON, B., SANTORO, A. & BOEHM, A. 2007. Beach sands along the California coast are diffuse sources of fecal bacteria to coastal waters. *Environmental Science and Technology*, 41, 4515-21.
- ZOETENDAL, E. G., VAUGHAN, E. E. & DE VOS, W. M. 2006. A microbial world within us. *Molecular Microbiology*, 59, 1639-1650.

Chapter 4.

***Bacteroidales* Markers for Microbial Source Tracking in Southeast Asia**

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, 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, 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.

4.1 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, Bernhard and Field, 2000b, 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., 2015b, Ekklesia et al., 2015a). 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, was based on the DNA binding dye SYBR Green. HF183 SYBR (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 to have limited cross-reaction with animal feces compared to other human-associated *Bacteroidales* markers such as HF183 (Yampara-Iquise et al., 2008, Aslan and Rose, 2013).

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

The objectives of this study were (i) to find the best assay(s) for detecting and monitoring human and sewage fecal pollution and (ii) to test the efficacy of a high performing canine-associated *Bacteroidales* assay (BacCan) in Singapore. We evaluated the performance of seven host-associated *Bacteroidales* marker assays against 295 fecal samples from humans, sewage, and eight relevant domesticated and wild animals. Both sensitivity and specificity were compared to select the best marker for

microbial source tracking in the tropical urban environment of Singapore and Southeast Asia, a region hosting one of the fastest growing human populations in the world.

4.2 Materials and Methods

4.2.1 Study area

This study was carried out in the tropical urban environment of Singapore, an island 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.

4.2.2 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 quantity of droppings, and iv) movement patterns such as grazing in groups of more than three animals. To ensure that all animal fecal samples collected were fresh or had recently been dropped, dry droppings or those found after a rainfall event were not collected. In addition, sightings of animals in the vicinity of the sampling location and fresh animal footprints were other indicators used to identify fresh droppings during wild animal tracking. NParks staff provided animal behavior training in order to ensure that animals were not stressed by our sampling events. Domesticated animals were sampled approximately two hours after every morning feeding at the rescue house of the Society for the Prevention of Cruelty to Animals in Singapore. The rescue house was selected because it provided an opportunity to sample multiple types and individuals of healthy domesticated animals.

Each fecal sample was collected using a 15-ml Falcon® tube (Corning Inc., Tewksbury, MA USA) and a stool nucleic acid collection and preservation bottle (Norgen Biotek Corp., Thorold, Ontario, Canada) for humans, monkeys and wild boars. Collected fecal samples were frozen on site, stored and carried to the laboratory using a high capacity

liquid nitrogen freezer (Taylor-Wharton Cryogenics LLC, Minnetonka, MN USA). Sewage samples were collected by the grab method and one liter was filtered onto Sterivex 0.22- μ m pore size filters (Millipore, Billerica, MA, USA). Filter pieces were removed from the cartridge and aseptically sliced into eight to ten fragments using a flame-sterilized scalpel prior to analysis. Once at the lab, fecal and sewage samples were stored at -80°C until processing.

Human stool samples were collected using an anonymous sampling strategy carefully designed to protect the identity of the human volunteers. Every individual volunteer was given a 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°C 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.

4.2.3 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 -20⁰C degrees until further analysis. DNA concentration and quality were measured and assessed using Qubit™ 2.0 Fluorometer (Invitrogen™, Waltham, MA, USA) and NanoDrop® ND-1000 (Waltham, MA, USA), respectively. Three procedural controls of water in bottles were filtered using Sterivex (0.22 µm pore size) and DNA extracted using Power DNA extraction kit tested negative by qPCR for the human markers used in this study.

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

4.2.4 qPCR assays

All probe-based *Bacteroidales*-associated 16S rRNA gene marker assay mixes included 20 μ L of qPCR mixture containing 10 μ L of TaqMan® Fast Advanced Master Mix 2.0 (ThermoFisher Scientific, Waltham, MA USA), primers and probes of which added quantities were determined based on their final concentration in the mix (Table 1), 2 μ L of genomic DNA, and molecular quality water to reach the final volume. The 20- μ L reaction mixture for the SYBR Green-based human-associated HF183 assay (HF183_SYBR) contained 10 μ L of KAPA SYBR® FAST 2X Master Mix (KAPABIOSYSTEMS, Woburn, MA, USA), 10 μ M of each primer and 2 μ L of DNA template (Seurinck et al., 2005). All quantification reactions were performed using the LightCycler® 480 Real-Time PCR system and software v. 1.5.0 (Roche Applied Sciences, Indianapolis, IN, USA) for calculating crossing point (Cp) values. Melting temperature (Tm) analysis for the HF183_SYBR Green assay was performed as previously described (Nshimyimana et al., 2014, Seurinck et al., 2005). The qPCR protocol for probe-based assays followed TaqMan® Fast Advanced Master manufacturer's instructions. Briefly, reactions were subjected to a pre-incubation step of 50°C for 2 min and 95°C for 20 s, followed by 40 cycles of 95°C for 3 s, 60°C for 30 s, and 40°C for 30 s. The qPCR assay for HF183 SYBR Green followed KAPA SYBR® FAST 2X Master manufacturer's instructions. Briefly, reactions were subjected to a pre-incubation step of 95°C for 3 min, followed by 45 cycles of 95°C for 10 s, 53°C for 20 s and 72°C for 1 s. Each sample, whether using probe-based or SYBR Green technology, was analyzed in duplicate and Cp values were examined after amplification to verify consistency (i.e., Cp difference between duplicates was ≤ 1 Cp).

The overall Cp difference between duplicates for samples amplified in this study was averaged at 0.40 ± 0.06 .

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

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

To ensure that the 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. $\geq 65\%$ 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 (Nshimiyimana et al. 2014).

4.2.4 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).

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

Assay	Primer or probe ¹	Concentration in 20- μ l 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		
	BacHum160_F	400 nM	TGAGTTCACATGTCCGCATGA		
BacHum (TaqMan)	BacHum241_R	400 nM	CGTTACCCCGCCTACTATCTAATG	81 bp	Kildare et al. 2007
		80 nM	FAM-TCCGGTAGACGATGGGGATGCGTT-TAMRA		
	BacHum193_P				
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		
	BacCan545_F	400 nM	GGAGCGCAGACGGGTTTT		
BacCan (TaqMan)	BacCan690_R1	400 nM	CAATCGGAGTTCTTCGTGATATCTA	145 bp	Kildare et al. 2007
	BacCan690_R2	400 nM	AATCGGAGTTCCTCGTGATATCTA		
	BacCan656_P	80 nM	FAM-TGGTGTAGCGGTGAAA-MGB		
	BacUni520_F	400 nM	CGTTATCCGGATTTATTGGGTTTA		
BacUni (TaqMan)	BacUni690_R1	400 nM	CAATCGGAGTTCTTCGTGATATCTA	145 bp	Kildare et al. 2007
	BacUni-690_R2	400 nM	AATCGGAGTTCCTCGTGATATCTA		
	BacUni656_P	80 nM	FAM-TGGTGTAGCGGTGAAA-MGB		
	BtH4515901_F	900 nM	CATCGTTCGTCAGCAGTAACA		
B. theta (TaqMan)	BtH4515963_R	900 nM	CCAAGAAAAAGGGACAGTGG	63 bp	Yampara-Iquise et al. 2008
	BtH_P	250 nM	Roche Universal library probe #062		

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

4.2. 6 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/ μ l) of each assay's DNA plasmid standard containing the target sequence. Standard curves of C_p 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 C_p 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 (Nshimiyimana et al., 2014) (Table 2). For HF183 SYBR, amplification of the correct target was confirmed based on the criteria that sample melting temperatures (T_m) fall within two standard deviations of the mean T_m associated with qPCR standards at concentrations of 10^1 - 10^6 copies per qPCR ($78.9^\circ\text{C} \pm \text{SD } 0.15$).

4.2. 7 Evaluation of assay sensitivity and specificity

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 (concentration $> \text{LOD}$) 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.

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

Eq. 1

$$\text{Specificity} = \frac{TN(\text{True Negative})}{(TN+FP(\text{False Positive}))}$$

Eq. 2

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+FP+TN+FN)}$$

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.

4.2.8 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).

Table 2. Performance characteristics of tested Bacteroidales assays

Assay	Slope	Y-intercept	R²	Efficiency	LOD¹ (copies/μl)
HF183-SYBR	-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

¹ Limit of detection

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

Sample	Percentage or number of positive samples							n
	Human Assay					Canine Assay	Universal Assay	
	HF183-SYBR	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
Number of inhibited samples								
Inhibited samples	0	10	7	24	1	3	13	58

4. 3. Results

4.3.1 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-Taqman, and HF183-SYBR) 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, HF183 and BacHum, and 40% for BacH; cats: 10% B. theta; and chickens: 30% HF183-SYBR, 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 gene copies/g wet feces, while the range in cross-reacting chicken samples was from 3.8 to 5.2 log 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 gene copies/g wet feces) and the canine assay detected two cat samples, recording 6.2 and 6.5 log 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 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 wild animals fecal samples were very low ($< \text{LOD}$) (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.

4.3.1.1 Comparison of marker concentration 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 gene copies/ml; Supplemental Table 2), while levels of BacH were the lowest (5.7 ± 1.5 log 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 ± 0.97 , 6.0 ± 1.55 , and 6.0 ± 1.80). 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 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 plant (Supplemental Table 2). The canine marker was found at low concentrations in human stool and sewage compared to human markers ($< 3.9 \pm 0.6$ log gene copies/g wet feces) with the exception of two samples (> 4.0 log gene copies/g wet feces). The universal marker, BacUni, was detected at the highest concentration in all sewage (5.2 ± 0.3 sd log gene copies/ml), human (7.5 ± 0.8 sd log

copies/g of wet feces) and animal samples excluding birds (6.7 ± 1.6 SD log gene copies/g of wet feces) (Figure 1, 2, 3, and Supplemental Table 2).

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	Human	1.0	0.63	0.88	0.83
HF183		1.0	0.60	0.90	0.83
BacHum		1.0	0.65	0.91	0.85
BacH		1.0	0.51	0.90	0.81
B. theta		1.0	0.69	0.98	0.90
Universal					
BacUni	Mammal	1.0	1.0	1.0	1.0
Canine					
BacCan	Dog	-	0.80	0.97	0.96

4.3.2 Performance of host-associated *Bacteroidales* assays

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

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

4.3.2.1 Human associated assays

4.3.2.2 Universal assay

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

4.3.2.3 Canine-associated assay

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

Figure 1. Concentrations of Bacteroidales-associated genetic markers in human stool and sewage 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.

Figure 2. Concentrations of Bacteroidales-associated 16S rRNA 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 markers. The red discontinuous lines separate types of fecal samples tested.

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.

4.4. Discussion

4.4.1 Performance of Canine-associated and Universal *Bacteroidales* assays

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

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

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

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

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

This study validated the application of host-*Bacteroidales* markers for microbial source tracking in Southeast Asia. A recent study established decay models for *B. theta* in artificial fresh and seawater as a function of sunlight intensity and salinity with one log₁₀ reduction times (T₉₀) from 3.1 to 53 h (Liang et al. 2017). The lower concentration of *B. theta* in sewage in Singapore suggests that the marker, if detected in environmental samples, should be viewed as an indicator of recent contamination and could be complemented with another validated marker of greater abundance in sewage, such as BacHum. Although the use of molecular based methodology may require a high investment for regional developing countries, the outcome of applying these markers could have a high return associated with sewage leak mitigation and control of fecal non-point sources of pollution in the urban environment.

4.5. Conclusions

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

- The human-associated *Bacteroidales* assay B. theta, the canine-associated assay BacCan, and the universal *Bacteroidales* assay BacUni are validated assays for use in the tropical urban environment of Southeast Asia.
- Rabbit fecal samples cross-reacted with all human-associated assays except B. theta. This calls for caution when considering MST studies in the region and confirms the usefulness of the highest specificity of B. theta to discriminate human fecal contamination from other sources.
- Findings of this study should assist emerging densely populated cities in the Southeast Asia region with limited sewage management infrastructures to efficiently identify and track potential sources of human fecal pollution for mitigation of sewer leaks and protection of recreational waters.
- Fate and transport studies of these validated MST markers under tropical conditions in water and sediments are needed to correctly identify pollution events in waterways and water bodies on a temporal and spatial scale.
- Overall, we demonstrated the need for similar validation studies before applying any existing human-associated *Bacteroidales* markers in different geographical locations. Regional application of the findings of this study would benefit from optimization of a regional protocol under urban tropical environmental conditions.

References: Chapter 4

- AHMED, W., POWELL, D., GOONETILLEKE, A. & 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, 579-86.
- AHMED, W., STEWART, J., POWELL, D. & GARDNER, T. 2008b. Evaluation of *Bacteroides* markers for the detection of human faecal pollution. *Letters in Applied Microbiology*, 46, 237-42.
- AHMED, W., YUSUF, R., HASAN, I., GOONETILLEKE, A. & 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, 838-845.
- ASLAN, A. & 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, 51-6.
- BERNHARD, A. E. & 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, 1587-1594.
- 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, 4571-4574.
- BOEHM, A., VAN DE WERFHORST, L. C., GRIFFITH, J., HOLDEN, P., JAY, J., SHANKS, O., WANG, D. & WEISBERG, S. 2013. Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. *Water Research*, 47, 6812-6828.
- EKKLESIA, E., SHANAHAN, P., CHUA, L. H. & EIKAAS, H. S. 2015a. Associations of chemical tracers and faecal indicator bacteria in a tropical urban catchment. *Water Research*, 75, 270-81.
- EKKLESIA, E., SHANAHAN, P., CHUA, L. H. C. & EIKAAS, H. S. 2015b. Temporal variation of faecal indicator bacteria in tropical urban storm drains. *Water Research*, 68, 10.
- 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, 689.

- GAWLER, A. H., BEECHER, J. E., BRANDÃO, J., CARROLL, N. M., FALCÃO, L., GOURMELON, M., MASTERSON, B., NUNES, B., PORTER, J. & RINCÉ, A. 2007. Validation of host-specific *Bacteroidales* 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. *Water Research*, 41, 3780.
- GREEN, H., SHANKS, O., SIVAGANESAN, M., HAUGLAND, R. & FIELD, K. 2011. Differential decay of human faecal *Bacteroides* in marine and freshwater. *Environmental Microbiology*, 13, 3235-49.
- GREEN, H. C., DICK, L. K., GILPIN, B., SAMADPOUR, M. & FIELD, K. G. 2012. Genetic Markers for Rapid PCR-Based Identification of Gull, Canada Goose, Duck, and Chicken Fecal Contamination in Water. *Applied and Environmental Microbiology*, 78, 503-510.
- HARRIS, D. C. 1995. *4th Edition: Quantitative Chemical Analysis*. New York, W. H. Freeman and Company.
- HARWOOD, V. J., BROWNELL, M., WANG, S., LEPO, J., ELLENDER, R. D., AJIDAHUN, A., HELLEIN, K. N., KENNEDY, E., YE, X. & FLOOD, C. 2009. Validation and field testing of library-independent microbial source tracking methods in the Gulf of Mexico. *Water Research*, 43, 4812-9.
- HAUGLAND, R. A., VARMA, M., SIVAGANESAN, M., KELTY, C., PEED, L. & SHANKS, O. C. 2010. Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected *Bacteroidales* species and human fecal waste by qPCR. *Systematic and Applied Microbiology*, 33, 348-57.
- JENKINS, M. W., TIWARI, S., LORENTE, M., GICHABA, C. M. & WUERTZ, S. 2009. Identifying human and livestock sources of fecal contamination in Kenya with host-specific *Bacteroidales* assays. *Water Research*, 43, 4956-66.
- KILDARE, B., LEUTENEGGER, C., MCSWAIN, B., BAMBIC, D., RAJAL, V. & WUERTZ, S. 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: A Bayesian approach. *Water Research*, 41, 15.
- LAYTON, B., CAO, Y., EBENTIER, D., HANLEY, K., BALLESTÉ, E., BRANDÃO, J., BYAPPANAHALLI, M., CONVERSE, R., FARNLEITNER, A., GENTRY SHIELDS, J., GIDLEY, M., GOURMELON, M., LEE, C., LEE, J., LOZACH, S., MADI, T., MEIJER, W., NOBLE, R., PEED, L., REISCHER, G., RODRIGUES, R., ROSE, J., SCHRIEWER, A., SINIGALLIANO, C., SRINIVASAN, S., STEWART, J., VAN DE WERFHORST, L. C., WANG, D., WHITMAN, R., WUERTZ, S., JAY, J., HOLDEN, P., BOEHM, A., SHANKS, O. & GRIFFITH, J. 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. *Water Research*, 47, 6897-6908.

- LIANG, L., GOH, S. G., VERGARA, G. G. R. V., FANG, H. M., REZAEINEJAD, S., CHANG, S. Y., BAYEN, S., LEE, W. A., SOBSEY, M. D., ROSE, J. B. & GIN, K. Y. H. 2015. Alternative Fecal Indicators and Their Empirical Relationships with Enteric Viruses, *Salmonella enterica*, and *Pseudomonas aeruginosa* in Surface Waters of a Tropical Urban Catchment. *Applied and Environmental Microbiology*, 81, 850-860.
- MATTIOLI, M. C., PICKERING, A. J., GILSDORF, R. J., DAVIS, J. & BOEHM, A. B. 2013. Hands and water as vectors of diarrheal pathogens in Bagamoyo, Tanzania. *Environmental Science and Technology*, 47, 355-63.
- NEAVE, M., LUTER, H., PADOVAN, A., TOWNSEND, S., SCHOBEN, X. & GIBB, K. 2014. Multiple approaches to microbial source tracking in tropical northern Australia. *Microbiology Open*, 3, 860-874.
- NPARKS. 2016. *Parks and nature reserves* [Online]. Singapore Government: National Parks. Available: <https://www.nparks.gov.sg/gardens-parks-and-nature/parks-and-nature-reserves> [Accessed June 15 2016].
- NSHIMYIMANA, J. P., EKKLESIA, E., SHANAHAN, P., CHUA, L. H. C. & THOMPSON, J. R. 2014. Distribution and abundance of human-specific *Bacteroides* and relation to traditional indicators in an urban tropical catchment. *Applied Microbiology*, 116, 1369-1383.
- ODAGIRI, M., SCHRIEWER, A., HANLEY, K., WUERTZ, S., MISRA, P. R., PANIGRAHI, P. & JENKINS, M. W. 2015. Validation of *Bacteroidales* quantitative PCR assays targeting human and animal fecal contamination in the public and domestic domains in India. *Science of the Total Environment*, 502, 462-470.
- OKABE, S., OKAYAMA, N., SAVICHTCHEVA, O. & ITO, T. 2007. Quantification of host-specific *Bacteroides-Prevotella* 16S rRNA genetic markers for assessment of fecal pollution in freshwater. *Applied Microbiology and Biotechnology*, 74, 890-901.
- PICKERING, A. J., JULIAN, T. R., MAMUYA, S., BOEHM, A. B. & DAVIS, J. 2011. Bacterial hand contamination among Tanzanian mothers varies temporally and following household activities. *Tropical Medicine & International Health*, 16, 233-9.
- PICKERING, A. J., JULIAN, T. R., MARKS, S. J., MATTIOLI, M. C., BOEHM, A. B., SCHWAB, K. J. & DAVIS, J. 2012. Fecal contamination and diarrheal pathogens on surfaces and in soils among Tanzanian households with and without improved sanitation. *Environmental Science and Technology*, 46, 5736-43.

- REISCHER, G. H., EBDON, J. E., BAUER, J. M., SCHUSTER, N., AHMED, W., ÅSTRÖM, J., BLANCH, A. R., BLÖSCHL, G., BYAMUKAMA, D., COAKLEY, T., FERGUSON, C., GOSHU, G., KO, G., DE RODA HUSMAN, A. M., MUSHI, D., POMA, R., PRADHAN, B., RAJAL, V., SCHADE, M. A., SOMMER, R., TAYLOR, H., TOTH, E. M., VRAJMASU, V., WUERTZ, S., MACH, R. L. & FARNLEITNER, A. H. 2013. Performance Characteristics of qPCR Assays Targeting Human- and Ruminant-Associated Bacteroidetes for Microbial Source Tracking across Sixteen Countries on Six Continents. *Environmental Science and Technology*, 47, 8548-8556.
- SANTO DOMINGO, J. W., BAMBIC, D. G., EDGE, T. A. & WUERTZ, S. 2007. Quo vadis source tracking? Towards a strategic framework for environmental monitoring of fecal pollution. *Water Research*, 41, 3539-52.
- SCHRIEWER, A., GOODWIN, K. D., SINIGALLIANO, C. D., COX, A. M., WANLESS, D., BARTKOWIAK, J., EBENTIER, D. L., HANLEY, K. T., ERVIN, J., DEERING, L. A., SHANKS, O. C., PEED, L. A., MEIJER, W. G., GRIFFITH, J. F., SANTODOMINGO, J., JAY, J. A., HOLDEN, P. A. & WUERTZ, S. 2013. Performance evaluation of canine-associated *Bacteroidales* assays in a multi-laboratory comparison study. *Water Research*, 47, 6909-20.
- SEURINCK, S., DEFOIRDT, T., VERSTRAETE, W. & SICILIANO, S. 2005. Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environmental Microbiology*, 7, 249-59.
- SILKIE, S. S. & NELSON, K. L. 2009. Concentrations of host-specific and generic fecal markers measured by quantitative PCR in raw sewage and fresh animal feces. *Water Research*, 43, 4860-71.
- SUZUKI, T. A. & WOROBEY, M. 2014. Geographical variation of human gut microbial composition. *Biology Letters*, 10.
- TAMBALO, D. D., BOA, T., LILJEBJELKE, K. & YOST, C. K. 2012. Evaluation of two quantitative PCR assays using *Bacteroidales* and mitochondrial DNA markers for tracking dog fecal contamination in waterbodies. *Microbiological Methods*, 91, 459-467.
- WIDMER, K., VAN HA, N. T., VINTNANTHARAT, S., STHIANNOPKAO, S., WANGSAATMAJA, S., PRASETIATI, M. A., THANH, N. C., THEPNOO, K., SUTADIAN, A. D., THAO, H. T., FAPYANE, D., SAN, V., VITAL, P. & HUR, H. G. 2013. Prevalence of *Escherichia coli* in surface waters of Southeast Asian cities. *World Journal of Microbiology and Biotechnology*, 29, 2115-24.
- WILKES, G., BRASSARD, J., EDGE, T. A., GANNON, V., JOKINEN, C. C., JONES, T. H., MARTI, R., NEUMANN, N. F., RUECKER, N. J., SUNOHARA, M., TOPP, E. & LAPEN, D. R. 2013. Coherence among Different Microbial Source

Tracking Markers in a Small Agricultural Stream with or without Livestock Exclusion Practices. *Applied and Environmental Microbiology*, 79, 6207-6219.

YAMPARA-IQUISE, H., ZHENG, G., JONES, J. E. & CARSON, C. A. 2008. Use of a *Bacteroides thetaiotaomicron*-specific *alpha*-1-6, mannanase quantitative PCR to detect human faecal pollution in water. *Applied Microbiology*, 105, 1686-93.

YATSUNENKO, T., REY, F. E., MANARY, M. J., TREHAN, I., DOMINGUEZ-BELLO, M. G., CONTRERAS, M., MAGRIS, M., HIDALGO, G., BALDASSANO, R. N., ANOKHIN, A. P., HEATH, A. C., WARNER, B., REEDER, J., KUCZYNSKI, J., CAPORASO, J. G., LOZUPONE, C. A., LAUBER, C., CLEMENTE, J. C., KNIGHTS, D., KNIGHT, R. & GORDON, J. I. 2012. Human gut microbiome viewed across age and geography. *Nature*, 486, 222-227.

AB. 2011. *TaqMan Fast Advanced Master Mix on the Roche LightCycler 480 Real-Time PCR System* [Online]. Applied Biosystems by Life Technologies. Available:
https://tools.thermofisher.com/content/sfs/brochures/cms_087007.pdf
[Accessed November 20 2016].

GREEN, H. C., HAUGLAND, R. A., VARMA, M., MILLEN, H. T., BORCHARDT, M. A., FIELD, K. G., WALTERS, W. A., KNIGHT, R., SIVAGANESAN, M., KELTY, C. A. & SHANKS, O. C. 2014. Improved HF183 Quantitative Real-Time PCR Assay for Characterization of Human Fecal Pollution in Ambient Surface Water Samples. *Applied and Environmental Microbiology*, 80, 9.

VERHOUGSTRAETE, M. P., MARTIN, S. L., KENDALL, A. D., HYNDMAN, D. W. & ROSE, J. B. 2015. Linking fecal bacteria in rivers to landscape, geochemical, and hydrologic factors and sources at the basin scale. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 10419-10424.

Chapter 5.

Evaluation of digital droplet PCR (ddPCR) for improved Microbial Source Tracking

Abstract

This study addressed whether digital droplet PCR (ddPCR) could improve sensitivity and specificity of human-associated *Bacteroidales* genetic markers and their quantification in environmental and fecal composite samples. BacHum and B. theta, previously validated for microbial source tracking in Singapore and Southeast Asia, were tested in 180 samples and quantified by qPCR and ddPCR (n = 35 human stool, n = 70 domestic and wild animal feces, n = 20 sewage, n = 20 environmental and n = 35 composite samples). Quantification of BacHum by ddPCR increased specificity (from 0.70 to 0.90) and accuracy (from 0.80 to 0.94) relative to qPCR, while the B. theta marker performed similarly on both platforms (specificity = 0.98 for qPCR and ddPCR). DdPCR and qPCR quantification of environmental and fecal composite samples were highly correlated ($R > 0.87$, $p < 0.0001$, $n = 110$) where concentrations measured by ddPCR were consistently lower than those measured by qPCR, by a factor of 3.3 ± 1.6 for B. theta and by a factor of 14.4 ± 8.5 for BacHum. High differences between quantifications by qPCR and ddPCR for BacHum concentrations were associated with a systematic error in the qPCR standard curve, while similar errors were not observed in the quantification by ddPCR. Hence ddPCR was the preferred method by removing systematic errors associated with the standard curve and improving quantification of samples with low target concentrations (n=41 for BacHum & B. theta) that were not reliably assessed by qPCR. The BacHum and B. theta assays, once adapted for ddPCR, achieved similar or improved performance relative to previously

established qPCR assays. We conclude that ddPCR is a suitable tool for microbial source tracking; however, other factors such as cost-effectiveness should be considered.

5.1 Introduction

Several significant technological advances have improved the application of microbial source tracking (MST) to identify sources of fecal contamination in waterways and water bodies. Use of the polymerase chain reaction (PCR) to identify DNA-based markers in environmental samples enabled development of cultivation-independent methods for water quality assessment. PCR-based MST was improved by adoption of real-time quantitative PCR (qPCR) that enabled quantification of targeted fecal sources (Dick and Field, 2004, McQuaig et al., 2009, Seurinck et al., 2005, Bernhard and Field, 2000). This improvement led to the development of multiple QPCR assays to determine the relative contributions from humans and various animal sources of fecal contamination (Shanks et al., 2008, Shanks et al., 2009, Kildare et al., 2007). However, limitations of QPCR were associated with poor MST performance (Shanks et al., 2012, Layton et al., 2013, Stewart et al., 2013, Molina et al., 2014) notably, quantification errors associated with PCR inhibition (Noble et al., 2010, Cao et al., 2012), variability in standard curves, low reproducibility over time and between different laboratories (Shanks et al., 2012), performance variability associated with the use of reagents purchased from different vendors (Cao et al., 2013) or batches of products from the same vendor (Sivaganesan et al., 2011), and lack of consistent standard material (Cao et al., 2013). Recently, the USEPA approved the use of two qPCR methods as a standard method for water quality assessment (USEPA, 2013, USEPA, 2012a, USEPA, 2012b) highlighting the utility of the technology despite these concerns.

Introduction of a next-generation quantitative tool, the digital droplet PCR (Pinheiro et al., 2012, McDermott et al., 2013, Hindson et al., 2013) promises to address some limitations of qPCR and further improve the implementation of MST. The workflow for digital droplet PCR includes partitioning the reagent mix containing a DNA binding fluorescent dye into thousands of individual reactions droplets in an oil emulsion followed by amplification (McDermott et al., 2013). The target concentration is determined by counting the droplets that score a positive or a negative for target amplification based on fluorescence of the DNA-binding dye. The distribution of positive droplets is processed using Poisson statistics to generate concentrations of the target in each sample without the use of a standard curve (Hindson et al., 2011).

Digital droplet PCR has been reported to be superior to qPCR in terms of sensitivity, specificity, accuracy, reproducibility, absence of inhibitors effects and highest in precision for quantifying lower concentrations often at the level of the limit of detection (Pinheiro et al., 2012, McDermott et al., 2013, Hindson et al., 2013, Whale et al., 2013). These characteristics have led to the application of ddPCR in multiple fields including the medical field for virus detection and quantification (Hayden et al., 2013), in food testing to assess levels of contamination (Floren et al., 2015, Morisset et al., 2013) and in determining concentrations of waterborne viruses (Racki et al., 2014), bloom-forming cyanobacteria (Cao et al., 2015, Te et al., 2015), and the HF183 and *Enterococcus* markers for microbial source tracking (Cao et al., 2015, Te et al., 2015). Given its increasing application to water quality assessment, we sought to evaluate whether this technology was an appropriate platform for quantification of two *Bacteroidales* assays recently validated for MST in tropical environments.

Human-associated *Bacteroidales* 16S rRNA genetic markers were developed as MST markers to identify, quantify and monitor levels of human fecal pollutions by qPCR (Yampara-Iquise et al., 2008, Shanks et al., 2009, Shanks et al., 2010, Sauer et al., 2011, Green et al., 2014, Molina et al., 2014, Kildare et al., 2007). *Bacteroidales* were preferred due to their ubiquity in human and animal guts, and their anaerobic life conditions in the environment (Fogarty and Voytek, 2005). Validation of human-associated *Bacteroidales* 16S rRNA genetic markers requires testing markers against fecal samples from animals and humans to determine the sensitivity and cross-reactivity of the assays. Fecal samples, and environmental water samples, are often associated with high levels of inhibitors and low levels of targets (near level of limit of detection) leading to quantification errors by qPCR (Grgicak et al., 2010, Cao et al., 2012). In this study we assessed how ddPCR compared to qPCR by asking whether ddPCR improves the sensitivity, specificity, and accuracy of human-associated *Bacteroidales* markers (B. theta and BacHum) in fecal and sewage samples of mixed origin and by comparing the detection and quantification of human-associated *Bacteroidales* markers in environmental and composite samples.

5.2. Methods

5.2.1. Sample selection and preparation

A total of 180 genomic DNA (gDNA) samples were obtained consisting of genomic DNA (gDNA) from human and animal fecal samples (n=105) and sewage samples (n=20) previously used in a validation study of *Bacteroidales* markers for microbial source tracking in Southeast Asia (Nshimiyimana et al., 2016), environmental water samples collected in tropical urban residential area (n=20), and composite fecal samples (n=35). Animal samples included domesticated animals (cats n=10, rabbits n=10, dogs n=10 and chickens n=10); wild animals (wild boars n=10, monkeys n=10, and birds n=10); while 35 human stool samples were obtained from volunteers as described in Chapter 4. Environmental water samples were filtered onto a 0.22µm membrane prior to DNA extraction as described in (Chapter 3). The concentration of each DNA sample was measured by NanoDrop nd-1000 (NanoDrop Technologies, Wilmington, DE, USA) and was diluted to a concentration of 2 ng/µL for use throughout this study. Composite fecal samples (n=35) were prepared by combining human and animal fecal DNA samples under a blind mixing protocol followed by dilution to 2 ng/µl (Table 2 and S1). The expected detection of markers for each composite sample was determined based on the combined characteristics of the undiluted individual samples measured by real-time qPCR and reported in chapter 4.

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

Assay	Primer or probe ¹	Concentration in 20µl reaction mix	Oligonucleotide (5' to 3')	Amplicon size	Reference
	BacHum160_F	400 nM	TGAGTTCACATGTCCGCATGA		
BacHum	BacHum241_R	400 nM	CGTTACCCCGCCTACTATCTAATG	81 bp	Kildare et al. 2007
(TaqMan)	BacHum193_P	80 nM	FAM-TCCGGTAGACGATGGGGATGCGTT- TAMRA		
B. theta (TaqMan)	BtH4515901_F	900 nM	CATCGTTCGTCAGCAGTAACA		Yampara-
	BtH4515963_R	900 nM	CCAAGAAAAAGGGACAGTGG	63 bp	Iquise et
	BtH_P	250 nM	Roche Universal library probe #062		al. 2008

¹F: Forward; R: Reverse; P: Probe. FAM, 6-carboxyfluorescein; TAMRA, 6-carboxytetramethylrhodamine.

5.2.2 Design of *B. thetaiotaomicron*-specific alpha-1-6, mannanase (*B. theta*) plasmid DNA

A *B. theta* plasmid DNA was designed and used in this study as a qPCR standard. The TaqMan qPCR assay for *Bacteroides thetaiotaomicron*-specific alpha-1-6, mannanase (average concentration of *B. thetaiotaomicron* in 1.39×10^8 cells/g of human feces) (Yampara-Iquise et al., 2008) was performed using the StepOnePlus Real-Time PCR System (Applied Biosystems®, Foster City, CA, USA). Each 25 µl qPCR reaction mixture contained 100 pg of *B. thetaiotaomicron* (VPI 5482) nucleic acid extract, 1x TaqMan® Environmental Master Mix 2.0 (Applied Biosystems®, Foster City, CA, USA), 900 nM (each) of *B. theta*-F and *B. theta*-R primers, and 250 nM *B. theta*-P probe #62 (Roche, Mannheim, Germany). The thermal cycling conditions applied were 2 min at 50°C and 10 min at 95°C, followed by 45 cycles at 94°C for 15 s and 60°C for 1 min. The amplicons from the qPCR were run on 1.5% agarose gel and the 63-bp target bands were extracted using a QIAquick® Gel Extraction Kit (Qiagen, Valencia, CA, USA).

Purified amplicons were inserted into the PCR™4-TOPO vector by use of the TA Cloning® Kit for sequencing (Invitrogen, Carlsbad, CA, USA). Plasmid DNA containing the target *B. thetaiotaomicron* sequences was extracted using a QIAprep Spin MiniPrep Kit (Qiagen, Valencia, CA, USA). Sequences of inserted target product were analyzed by ABI Prism® 3730 Genetic Analyzer in the DNA Sequencing Facility at the University of California, Davis.

Construction of plasmid DNA containing the BacHum assay standard was previously described by our laboratory (Odagiri et al., 2015). DNA concentrations of plasmid standards for *B. theta* and BacHum were measured by Qubit 2.0 Fluorometer (ThermoFisher Scientific, Waltham, MA, USA) and converted to plasmid copies based on the molecular weight of plasmid and insert. Standards were diluted to 10^8 copies/ul in deionized water, and were stored as single-use aliquots at -80C.

5.2.3 qPCR and ddPCR assays

5.2.3.1 qPCR assays

QPCR for human-associated *Bacteroidales* gene marker assays were carried out using Bio-Rad CFX96 (Bio-Rad, Hercules, CA, USA) (qPCR). All assays were conducted in 20 μ L reactions containing 10 μ L of SsoAdanced™ Universal Probes Supermix 2X (Bio-Rad, Hercules, CA, USA). Primer and probe concentrations were respectively 400nM and 80nM for BacHum, and 900nM and 250nM for *B. theta* based on published assay conditions (Yampara-Iquise et al., 2008, Kildare et al., 2007, Odagiri et al., 2015) (Table 1). Genomic DNA was added at 4 ng per reaction and molecular quality water to reach the final volume. Briefly, qPCR reactions were subjected to enzyme activation at 95°C for 3 min, denaturation followed by 40 cycles of amplification consisting of 95°C for 15 seconds, annealing at 60°C for 1min. Replicates were compared to demonstrate similarity of crossing point (C_p) values such that the coefficient of variability was $\leq 1\%$ for high concentrations ($C_p < 10$) and $\leq 2\%$ for lower concentrations ($C_p > 33$).

5.2.3.2 DdPCR assays

Samples were quantified using QX200™ Droplet Digital™ PCR (ddPCR) (Bio-Rad, Hercules, CA, USA) following the manufactures recommendations. Briefly, all B. theta and BacHum assay reactions were prepared from 10µL of ddPCR™ Supermix for probes (NodUTP), primers and probes at final concentrations of 900nM and 250nM respectively, with 4 ng genomic DNA. Each sample was loaded onto a D8™ cartridge (Bio-Rad, Hercules, CA, USA) with 70µL of ddPCR™ droplet generator oil (Bio-Rad, Hercules, CA, USA) and subjected to droplet generation on a QX200™ Droplet generator machine (Bio-Rad, Hercules, CA, USA). 40µL of each resulting emulsion was loaded to a 96 well-plate and sealed using PX1™ Plate Sealer (Bio-Rad, Hercules, CA, USA). Briefly, ddPCR assays were subjected to a thermocycle with enzyme activations at 95°C for 10min; denaturation 94°C for 30 seconds, annealing temperature 60°C for 1min; enzyme deactivation 98°C for 10min followed by 40 cycles and a continuous hold at 4°C. Quantification results were read using ddPCR™ Droplet Reader Oil (Bio-Rad, Hercules, CA, USA); QX200™ Droplet Reader (Bio-Rad, Hercules, CA, USA) and QuantaSoft™ software (Bio-Rad, Inc., Hercules, CA, USA). Quality control for ddPCR was implemented as previously described (Cao et al., 2015, Huggett et al., 2013, Pinheiro et al., 2012). Briefly, the number of accepted droplets generated per each sample run by ddPCR in this study was $\geq 10,000$. Positive concentrations of samples quantified in this study had a dynamic range from 0.12% to 99.0% positive droplets.

5.2. 4 ddPCR and qPCR comparison criteria

We used three different criteria to compare performance of ddPCR and qPCR for marker quantification including: (1) depth of quantification, (2) sensitivity and specificity, and (3) quantification of composite and environmental samples.

Table 2. Composite samples scored positive against expected outcome determined based on previous qPCR of mixed samples.

Expected outcome of composite samples*	Number of ddPCR samples matching expected outcome			n
	BacHum/B. theta	B. theta	BacHum	
Negative/Negative	6	7	8	
Positive/Positive	10	10	10	
Positive/Negative	9	9	9	
Negative/Positive	8	6	8	

*: Based on inclusion of one or more samples previously determined to be positive for the marker and previous qPCR: performed on mixed samples using BacHum and B. theta assays

5.2.4.1 Depth of quantification

We assessed the depth of quantification using plasmid DNA of B. theta and BacHum. Calibration curves were established using ten-fold serial dilutions (10^1 to 10^7 Copies/ μ l) of each assay's DNA plasmid standard containing the target sequence. Calibration curves of C_p 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 C_p values were calculated based on propagation of error in the standard curve (Harris, 1995). The limit of detection (LOD) for individual assays was determined as the concentration significantly different from the no template control (NTC) at 95% confidential intervals as previously described (Nshimiyimana et al., 2014). High linearity ($R^2 \geq 0.99$) and amplification efficiency (90% - 94%) of the qPCR standards confirmed best performance and good laboratory practices throughout the

study (Figure S1A &1C) (Shanks et al., 2012). For ddPCR, the limit of detection was determined by amplification of replicate standards of increasing dilutions and determining the concentration threshold where 95% of PCR reactions above the threshold were positive (i.e. the LOD95 method) (Stewart et al., 2013, CODEX, 2010, Burd, 2010).

5.2.4.2 Sensitivity and Specificity

We analyzed 125 animal fecal and sewage samples to examine if sensitivity and specificity of human-associated *Bacteroidales* markers (B. theta and BacHum) would vary based on the quantification platform used (qPCR or ddPCR). Briefly, assay sensitivity was determined based on the number of target host samples (human stool and sewage samples) testing positive for the assayed marker (B. theta and BacHum), divided by the total number of host samples tested (Kildare et al., 2007). Specificity was calculated as the number of non-target host samples tested negative for the assayed marker, divided by the total number of non-target samples tested (cats, rabbits, dogs, wild boars, monkeys, chicken and birds) (Kildare et al., 2007). The accuracy of the assays was computed as the ratio of target and non-target host samples identified correctly by total number of samples tested (Odagiri et al., 2015).

5.2.5 Statistical analysis

QPCR standard curves (i.e. calibration curves) were calculated by linear regression of standard concentrations and crossing-point (Cp) values. Concentrations determined by qPCR and ddPCR were log-transformed prior to statistical analysis unless indicated otherwise. The repeatability of quantification by ddPCR and qPCR was compared by

calculating the coefficient of variation from sample replicates. Pearson correlation was used to compare human marker concentrations quantified by qPCR and ddPCR in composite, environmental, and serial diluted plasmid DNA samples. The concentration of human markers measured in the same sample by qPCR and ddPCR were compared using non-transformed data. ANOVA and Paired T-test of environmental samples was used to compare the performance of qPCR and ddPCR in quantifying levels of human markers in fecal polluted environmental water samples. Mean values of human and sewage were compared using ANOVA followed by Tukey's honest significant difference (HSD) post-hoc multiple-comparisons test to determine difference in quantification across assays in human and sewage samples using JMP Pro (SAS Institute Inc., Cary, NC USA).

Table 3. Comparison of qPCR and ddPCR limit of detections.

Expected (Based on the plasmid)		qPCR (log copies/reaction)						ddPCR (log copies/reaction)					
copies/ reaction	log copies/ reaction	BacHum	SD	% CV	B. theta	SD	% CV	BacHum	SD	% CV	B.theta	SD	% CV
10 ⁷	7.0	7.1	0.02	0.21	7.6	0.17	2.29	TNTC	ND	ND	TNTC	ND	ND
10 ⁶	6.0	6.8	0.16	2.30	6.0	0.04	0.59	TNTC	ND	ND	TNTC	ND	ND
10 ⁵	5.0	6.0	0.02	0.32	4.8	0.03	0.64	TNTC	ND	ND	TNTC	ND	ND
10 ⁴	4.0	5.2	0.02	0.44	3.9	0.02	0.56	3.6	0.05	1.49	4.0	0.04	1.09
10 ³	3.0	4.0	0.09	2.25	2.8	0.04	1.32	2.8	0.04	1.47	2.6	0.02	0.65
100.0	2.0	2.9	0.05	1.80	1.9	0.00	0.08	2.0	0.07	3.18	2.0	0.00	0.26
50.0	1.7	ND	ND	ND	ND	ND	ND	1.8	0.02	1.20	1.6	0.03	1.65
25.0	1.4	ND	ND	ND	ND	ND	ND	1.5	0.03	2.22	1.4	0.01	0.76
12.5	1.1	ND	ND	ND	ND	ND	ND	1.2	0.02	1.66	1.1	0.02	2.23
10.0	1	0.8	0.06	3.29	1.2	0.03	2.82	1.1	0.06	5.72	0.9	0.02	2.10
5.8	0.7	ND	ND	ND	ND	ND	ND	0.7***	0.03	3.53	ND	ND	ND
3.5	0.5	ND	ND	ND	ND	ND	ND	0.77	0.40	54.4	0.4*	0.01	2.40
1.5	0.2	ND	ND	ND	ND	ND	ND	0	0	0	0.2**	0.06	3.50

Notes: Abbreviations: ND: Not determined; TNTC: positive droplets were Too numerous to count;

All averages represent triplicate samples with asterisks denoting exceptions based on duplicate (*), n=6 (**), and n=9 (***) samples.

Italic font represent values obtained below the detection limits determined at 95% confidence: qPCR BacHum = 19.1 copies/reaction; qPCR B.theta = 4.6 copies/reaction; ddPCR BacHum = 5.8 copies/reaction; ddPCR B.theta = 1.5 copies/reaction.

Figure 1. Concentrations of human-*Bacteroidales* markers (*B. theta* and BacHum) in human stool (Figure 1A), sewage (Figure 1B), and animal fecal samples (Figure 1C and Figure 2D) as quantified using qPCR and Digital droplet PCR (ddPCR). Blue color shows concentrations of *B. theta* quantified using qPCR, the green color shows concentrations of *B. theta* quantified using ddPCR, purple color shows concentrations of BacHum quantified using qPCR, and the goldenrod color shows concentrations of BacHum quantified using ddPCR. Grey colored En dash sign in Figure 1C indicates the level of the limit of detection (LOD).

5.3 Results

5.3.1 Target standards

Dilution series of two plasmid-borne standards were tested by both real-time qPCR and ddPCR. Quantification of standards by ddPCR indicated a dynamic range of 1.5 to 10^4 copies/reaction for B. theta and 5.8 to 10^4 copies/reaction for BacHum (Table 3). For ddPCR, quantification of standards above 10^4 (i.e. 10^5 to 10^7) was not possible due to saturation of droplets with positive signal. Quantification of standards by qPCR indicated a dynamic range of 4.6 to $>10^7$ copies/reaction for B. theta and 19.1 to $>10^7$ copies/reaction for BacHum (Table 3). All ddPCR and qPCR data generated passed the MIQE guidelines (Bustin et al., 2009, Huggett et al., 2013) and the efficiency of qPCR amplification ranged from 0.9 to 0.94 falling within the range attributed to good laboratory performance of human fecal-associated assays (Shanks et al., 2012, Griffith and Weisberg, 2011) (Figure 2A and 2B). Comparison of standard concentrations calculated based on Qubit fluorometry followed by dilution, with values measured by qPCR or ddPCR indicated good agreement for three of the four assays (qPCR-BacHum, ddPCR-B. theta, and qPCR-B. theta) evidenced by high linearity with a slope of 1.0 for measured versus expected values ($R^2 = 0.99$) (Figure 2 A, B, D). In contrast, while the ddPCR-BacHum assay exhibited high linearity ($R^2 = 0.99$) a slope of 1.1 indicated that diluted standards (5.8 to 10^4 copies/reaction) had higher marker concentrations than expected (Figure 2C). This trend in the ddPCR result was more pronounced at high dilution (<100 copies/reaction) (Slope 1.2, $R^2 = 0.99$) (Table 3 and Figure 3) suggesting pipetting error during serial dilution may have compromised the BacHum standards. For comparative purposes the qPCR standards for BacHum were not re-calibrated following ddPCR analysis. Assay reproducibility was examined for qPCR and ddPCR.

Replicates of standards yielded closely similar values with coefficients of variation 0.08 to 3.3% for qPCR and 0.65 to 5.7% for ddPCR indicating high precision for both platforms. This trend was consistent in high dilution/low concentration samples (<100 copies/reaction) (Table 3, S1).

5.3.2 Human and animal fecal samples: Sensitivity, specificity and accuracy

Human-associated *Bacteroidales* assays, BacHum and B. theta, validated as the best markers for microbial source tracking in Singapore and Southeast Asia as described in Chapter 4, were used to compare performance of ddPCR and qPCR. The qPCR for BacHum showed 95% (19/20) detection in sewage and 100% (35/35) in human stool samples, while the quantification of the same sample by ddPCR demonstrated a detection of 97.1% (34/35) for human stool and 100% (20/20) for sewage samples (Table 4). B. theta detection in human stool samples was 94.3% (33/35) by qPCR and 85.7% (30/35) by qPCR. The modest reduction in the detection of human markers in human stool samples by ddPCR (BacHum: 100% to 97.1% and B. theta: 94% to 87.7%) was associated with variation in classification of samples quantified near the detection limits for both platforms.

Decreased detection of human marker (BacHum) in non-target samples by ddPCR yielded an improvement in specificity. The number of animal (mammals) samples cross-reacting with BacHum using the qPCR platform was 22/60 versus 7/60 for ddPCR. The use of ddPCR was associated with a reduction in the number of cross-reacting samples by 22% (Figure 1C and 1D, Table 4). The quantification of B. theta

in non-target samples using qPCR and ddPCR did not change; both platforms identified one cat sample as positive for cross reactivity (Figure 2, Table 4).

Table 4. Performance of human-associated *Bacteroidales* assays as revealed by qPCR and ddPCR quantification of human fecal samples and domesticated and wild animal fecal samples

Sample categories	Percentage of Positive Samples (%)				n
	qPCR		ddPCR		
	BacHum	B.theta	BacHum	B.theta	
Humans					
Sewage	95	100	100	100	20
Human Stool	100	94.3	97.1	85.7	35
Domesticated Animals					
Rabbits	60	0	50	0	10
Dogs	10	0	0	0	10
Cats	90	10	0	10	10
Chickens	60	0	20	0	10
Wild Animals					
Monkeys	0	0	0	0	10
Wild boards	0	0	0	0	10
Birds	0	0	0	0	10

The specificity and accuracy of the BacHum marker were improved by ddPCR. In this study the quantification of BacHum by qPCR delivered a specificity of 0.63 (Table 5) and an accuracy of 0.80 (Table 5). The use of ddPCR improved all performance metrics for BacHum with a specificity of 0.88 by ddPCR compared to 0.63 by qPCR (Table 5) and an accuracy of 0.93 by ddPCR compared to 0.80 by qPCR (Table 5).

DdPCR showed a reduced sensitivity for BacHum and B. theta in human stool samples. The quantification of BacHum in human stool demonstrated a reduction in sensitivity from 1.0 by qPCR to 0.97 by ddPCR, while a quantification of B. theta showed a reduction from 0.94 to 0.85 by ddPCR (Table 5). In addition, B. theta showed also a reduction in accuracy from 0.97 by qPCR to 0.95 by ddPCR (Table 5). The sensitivity to sewage and specificity of B. theta remained unchanged for qPCR and ddPCR.

5.3.3 Composite and environmental samples

5.3.3.1 Environmental samples

The concentrations of human markers (B. theta and BacHum) in twenty catchment water samples were quantified by qPCR and ddPCR (Figure 4C and 4D). For all samples, qPCR indicated significantly higher concentrations of human markers than ddPCR (Paired T. test BacHum: $p=0.00016$ and B. theta: $p=0.0021$). High correlations were observed between concentrations of human marker quantified by qPCR and ddPCR (B. theta: $R=0.87$ and BacHum: $R=0.93$, $p < 0.0001$) (B. theta: slope = 1.2, $R^2=0.98$) except BacHum (slope = 0.6, $R^2=0.85$) (Figure 3C and D).

5.3.3.2 Composite samples

Concentrations of human markers (B. theta and BacHum) in composite fecal samples determined by qPCR and ddPCR were highly correlated (B. theta: slope 1.3, $R^2 = 0.97$ and BacHum: slope = 0.8, $R^2=0.94$ and B. theta: $R=0.90$, $p<0.0001$ and BacHum: $R=0.91$, $p<0.0001$) (Figure 4A and 4B).

Figure 2. Measured concentrations of positive controls of human-*Bacteroidales* markers (B. theta and BacHum) quantified using qPCR and ddPCR compared to expected concentrations of plasmid standards based on DNA concentrations measured by Qubit Fluorometer. A slope of 1 and $R^2 \geq 0.99$ demonstrate a good fit between measured and expected concentrations. QPCR efficiencies of BacHum and B. theta assays are demonstrated (2A and 2B) ($\text{Efficiency} = 10^{-1/\text{slope}} - 1$). X-axis is the expected concentrations and Y-axis is the measure concentrations by qPCR (2B and 2D) or by ddPCR (2A and 2C). Measured and expected concentrations of BacHum ($0 - 10^7$ Copies/ μl) by qPCR (2A). Measured and expected concentrations of B. theta ($0 - 10^7$ Copies/ μl) by qPCR (2B). Measured and expected concentrations of BacHum ($0 - 10^4$ Copies/ μl) by ddPCR (2C). Measured and expected concentrations of B. theta ($0 - 10^4$ Copies/ μl) by ddPCR (2D).

Figure 3. Measured and expected concentrations of positive controls of human-*Bacteroidales* markers (B. theta and BacHum) quantified using ddPCR, and computed using plasmid DNA of each marker at concentrations below 100 copies per reaction. A slope of 1 and $R^2 \geq 0.99$ demonstrate a good fit between measured and expected concentrations. X-axis is the expected concentrations and Y-axis is the measure concentrations by ddPCR (A: B. theta and B. BacHum). Expected concentrations of B. theta (50, 25, 12.5, 10, 3.5, and 1.5 Copies/reaction) (S2A). Expected concentrations of BacHum (50, 12.5, 10, and 5.8 Copies/reaction) (S2B).

Figure 4. Comparison of concentrations of human-*Bacteroidales* markers (BacHum and *B. theta*) in composite (3A and 3B) and environmental samples (3C and 3D). Comparison of concentrations of *B. theta* quantified by qPCR and ddPCR in composite samples (3A). Comparison of concentrations of BacHum quantified by qPCR and ddPCR in composite samples (3B). Comparison of concentrations of *B. theta* quantified by ddPCR and qPCR in environmental samples (3C). Comparison of concentrations of BacHum quantified by ddPCR and qPCR in environmental samples (3D). Bisector 1:1line (C and D). All graphs include slope and R^2 values. All plotted values are above the limit of detection (C: $n > \text{LOD} = 5$ & D: $n > \text{LOD} = 9$ for each marker under qPCR or ddPCR). Environmental samples are expressed at Copies/ 100ml of sampled volume.

Comparison of the human markers detected by ddPCR in composite fecal samples with the predicted classification of the samples based on previous qPCR analysis of undiluted samples (Chapter 4) confirmed similar performance of ddPCR and qPCR in detecting human fecal contamination (Table 2). Overall, the results of ddPCR matched the results predicted based on the results presented in Chapter 4 at 94.3 % (33/35 composite samples) for *B. theta* and 91.4% (32/35 composite samples) for BacHum (Table 2). Similar levels of agreement between qPCR results from this study and the previous study (Chapter 4) were obtained (Table 2) with some exceptions. Dilution of DNA from human stool in this study may have reduced marker concentrations leading to non-detects in two samples (samples B30 and B35). Dilution is also likely to reduce effects of low-level PCR inhibition that may have remained undetected in samples based on criteria established in Chapter 4 and may explain why the samples mixed into composite samples B6 and B7 originally tested negative for both BacHum and *B. theta* but scored positive for both by ddPCR and QPCR in this study.

5.3.3.3 Comparison of qPCR and ddPCR performance across different sample types.

We examined the relationship between concentrations of human markers by qPCR and ddPCR using human stool, sewage, composite, environmental samples, and standards. Overall, concentrations determined by qPCR (copies/reaction) were 14.4 ± 8.5 and 3.3 ± 1.6 times higher than measured by ddPCR respectively for BacHum and *B. theta* (Table 6). The more divergent performance of the BacHum assay between ddPCR and qPCR platforms likely relates to the previously discussed observation that diluted standards used to form the qPCR standard curve may have accumulated pipetting error. Indeed, while *B. theta* standards over the interval 10 to 10^4 copies/reaction were

measured at nearly the same level by qPCR and ddPCR (1.7 ± 1.0) the BacHum standards were 12.0 ± 14.0 fold higher by qPCR (Table 3, Table 6). The difference between qPCR and ddPCR concentrations differed with sample type, but the effect was not significant due to high variation among individual samples within each type. The smallest difference was for *B. theta* quantified in sewage (2.4 ± 1.5 fold difference between platforms; Table 6). The largest difference between concentrations of BacHum was measured in composite samples where qPCR concentrations were 22.5 ± 7.8 times those measured by ddPCR (Table 6).

Table 5. Human assays performance metrics calculated using BacHum and *B. theta* qPCR and ddPCR concentrations

Assays					
Performance	Samples	qPCR		ddPCR	
Metrics		BacHum	<i>B.theta</i>	BacHum	<i>B.theta</i>
Sensitivity	Sewage	0.95	1.00	1.00	1.00
	Human Stool	1.00	0.94	0.97	0.85
Specificity	Mammals	0.63	0.98	0.88	0.98
Accuracy	Sewage, Human stool &				
	Mammals	0.80	0.97	0.93	0.95

Table 6. Ratios of concentrations determined by qPCR and ddPCR for B. theta and BacHum quantified in the same sample

Type of samples	Average		Median	Maximum	Minimum	Average		Median	Maximum	Minimum
	BacHum	SD	BacHum	BacHum	BacHum	B. theta	SD	B. theta	B. theta	B. theta
Human	11.1	7	9.3	26.4	2.1	3.4	1.8	3.3	9.4	0.8
Sewage	12.5	5.7	12.1	26.9	4.1	2.6	1.5	2.4	7.7	1.3
Composite samples	22.1	7.8	22.5	34.1	0.8	3.4	1.3	3.7	34.1	0.9
Environmental samples	14	8.5	12.7	33.4	2.9	ND	ND	ND	ND	ND
Standards	12	14	6.5	41.6	3.4	2	1	1.7	3.6	1.5
All Data	14.9	8.5	14.4	34.1	0.8	3.3	1.6	3.3	9.4	0.8

Notes: Ratios of qPCR to ddPCR (qPCR/ddPCR) were determined for concentrations measured in copies/reaction and average, median, max, min determined for each sample type. Abbreviation: -ND: not determined due to insufficient dataset N<3

5.4 Discussion

5.4.1 Quantification of Standards by ddPCR and qPCR

This study validated the use of ddPCR as a tool for quantification of human *Bacteroidales* markers (BacHum and *B. theta*) for microbial source tracking. Currently, it is still challenging to detect human fecal pollution sources at low concentrations in environmental samples by qPCR (Harwood et al., 2014). The latter has also been reported to be associated with quantification errors (Grgicak et al., 2010, Cao et al., 2012), while ddPCR was reported to have a high sensitivity, specificity and provides concentrations without the use of a standard curve (Hindson et al., 2011, Hindson et al., 2013, Hayden et al., 2013, Yang et al., 2014, Cao et al., 2015, Te et al., 2015). This study provided insights for quantification of the target in sewage, fecal, composite, and environmental samples by ddPCR and evaluated how ddPCR compares to qPCR which has seen application in different geographies for microbial sources tracking (Ahmed et al., 2008a, Ahmed et al., 2008b, Ahmed et al., 2010, Neave et al., 2014, Harwood et al., 2009, Shanks et al., 2010, Haugland et al., 2010, Odagiri et al., 2015, Gawler et al., 2007, Edge et al., 2010).

A high degree of linearity and amplification efficiency of standards were achieved in this study (Figure 2). BacHum and *B. theta* serial diluted standards quantified by qPCR demonstrated a wider linear dynamic range (BacHum: $19.1 - 10^7$ copies/reaction and *B. theta*: $4.6 - 10^7$ copies/reaction) than ddPCR (BacHum: $5.8 - 10^4$ copies/reaction and *B. theta*: $1.5 - 10^4$ copies/reaction) (Figure 2, Table 3). Similar results were reported in

other two studies with application to freshwater and marine water quality monitoring in California, USA (Cao et al., 2015) and Singapore (Te et al., 2015).

High repeatability was consistent with previous studies and findings of this study (Morisset et al., 2013, Whale et al., 2013, Cao et al., 2015). The reproducibility of *B. theta* and BacHum concentrations as measured by qPCR and ddPCR were comparable (Table 3) and the limit of detection for the human markers (BacHum and *B. theta*) for ddPCR ranged within the margin of the previous studies (Figure 2 and Table 3). The ddPCR capabilities of detecting lower concentrations with high reproducibility enhanced the specificity of BacHum in this study showing potential usefulness in microbial source tracking as most of targets are detected a lower concentration in environmental samples (Figure 3, Table 3). The upper detection limit (10^4 copies/ μ l) had a lower coefficient of variability ($\leq 5.7\%$) and reproducibility achieved at 10^4 and 10^3 ($<3.8\pm0.1$, 2.5%, Supplemental Figure S1).

5.4.2 Use of ddPCR for microbial sources tracking

This study indicated mixed results for the comparison of ddPCR and qPCR for microbial source tracking by ddPCR with improvements noted for the BacHum assay but few significant changes for the *B. theta* assay. First, for BacHum, ddPCR revealed a source of systematic error in our qPCR standards, despite observation that these standards were well within established criteria for good performance by qPCR (i.e. high R^2 and PCR efficiency). This systematic error, leading to an average of 12-fold difference between expected standard concentrations and those measured by ddPCR, was likely associated with pipetting error for serial dilutions, which would have gone undetected in a study without a ddPCR benchmark. In addition, use of ddPCR improved

performance of the BacHum assay relative to qPCR. We observed an increase in specificity of BacHum by ddPCR (Table 5) and a reduction of the number of non-target cross-reacting animals (22%, Table 4) previously classified as positive for the marker by qPCR (Pinheiro et al., 2012, McDermott et al., 2013, Hindson et al., 2013). Enumeration of the BacHum marker in environmental, human stool, sewage, and composite samples by qPCR indicated 14-fold higher concentrations than determined by ddPCR (Table 6). This difference can be explained in large part by the systematic bias uncovered in the standard curve, as discussed above.

In contrast, the *B. theta* assay demonstrated a more similar range of quantification by both qPCR and ddPCR (i.e. 3.3 fold difference between platforms for both standards and samples, Table 6). Regression of *B. theta* standard concentrations estimated by dilution and measured by ddPCR maintained a slope of 1.0 at high dilutions (Figure 2 B and D) suggesting that higher accuracy in the qPCR standard curve contributed to better agreement of the *B. theta* results from the two platforms. A tropical study quantifying *Microcystis* demonstrated that qPCR quantifications were 1.3-fold to 6.8 fold (average 2.8-folds) higher than ddPCR (Te et al., 2015). For the medical field studies, a reversed trend was observed where various markers measured by ddPCR tend to be higher than those measured by qPCR. In a study using clinical samples, ddPCR concentrations tended to be 7-fold higher than qPCR (Hindson et al., 2013). Furthermore, another clinical study reported that ddPCR concentrations were up to 30-fold higher than qPCR (Taylor et al., 2015). Such studies, like ours, were based on qPCR assays that met the MIQE (2009) standards raising a concerning possibility that significant errors in values quantified by qPCR may be more prevalent than realized in published qPCR studies.

The *B. theta* assay by ddPCR and qPCR demonstrated similar sensitivity to sewage samples and specificity against non-target samples (Table 4 and 5). Findings of a previous study in California, USA using human markers (HF183) also demonstrated similar sensitivity and specificity by both platforms, ddPCR and qPCR (Cao et al., 2015). Lower sensitivity of the *B. theta* and BacHum ddPCR assays for human stool samples was likely due to low target concentrations below detection in human samples quantified as positive by qPCR ($C_p > 33$ cycles).

The trend of high correlation between concentrations of targets measured by qPCR and ddPCR in environmental and mixed samples was previously reported in other qPCR/ddPCR comparison study in California, USA (Cao et al., 2015) and Belgium, Europe (Verhaegen et al., 2016). The study in California reported a high linearity and correlation in concentrations of human-associated fecal indicators (HF183 and *Enterococcus*) quantified by qPCR and ddPCR in freshwater and marine water samples (e.g. Correlation $R \geq 0.93$, $p < 0.001$) (Cao et al., 2015). The study in Belgium, Europe examined the performance of qPCR and ddPCR in quantifying Shiga toxin-producing *E. coli* in Bovine feces and reported high linearity and correlation in mixed samples (e.g. Slope=1.0; $R^2 \geq 0.87$; $p < 0.005$) (Verhaegen et al., 2016).

Digital droplet is currently viewed as a better amplification tool than qPCR, as it can quantify lower targets in samples (< 100 copies/ reaction) (Figure 3) as demonstrated by limit of detection in this study (ddPCR-*B. theta*: LOD = 1.5 copies/ reaction and ddPCR-BacHum: LOD= 5.8 copies/reaction, Table 3). Lower concentrations were only evaluated for ddPCR as they have been examined for qPCR by previous research and strategies of improvement were presented (Ellison et al., 2006). This ability to detect

low concentrations and improvement of sensitivity and specificity of BacHum by ddPCR is attractive for microbial source tracking (MST) work. The use of ddPCR for MST demonstrated a potential improvement in the validation of the best performing markers for different environments and geographies. However, the general performance with respect to sensitivity, specificity and accuracy, was closely similar to that of qPCR despite uncertainties associated with the qPCR standard curve. The application of ddPCR for MST should then be weighed with the inclusion of other factors such as cost-effectiveness, time needed for sample processing and the overall impact of the study.

5.5 Conclusions

We assessed how ddPCR compared to qPCR. We evaluated whether ddPCR could improve sensitivity and specificity, and depth of quantification of human-associated *Bacteroidales* markers; and examined whether ddPCR improves the detection of human-associated *Bacteroidales* markers (*B. theta* and BacHum) in environmental and composite samples under tropical conditions.

- We demonstrated that ddPCR was a suitable platform for applying human-*Bacteroidales* markers for microbial source tracking with enhanced specificity observed for BacHum.
- Digital droplet PCR improved the quantification of qPCR-BacHum as it provided a baseline of expected quantification and allowed to identify systematic errors although qPCR-BacHum had passed all quality control parameters

- High reproducibility of ddPCR and the absence of a standard curve and associated errors enables more accurate quantification and makes it suitable for long-term and/or multi-laboratory investigations
- Performance similarity between qPCR and ddPCR platforms for quantifying human markers in composite and environmental samples supported the continuous usefulness of qPCR for targets above LOD. In addition, the burden of the high cost associated with ddPCR could still be alleviated by the use of qPCR with the expectation of a satisfactory outcome

References

- AHMED, W., POWELL, D., GOONETILLEKE, A. & 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, 579-86.
- AHMED, W., STEWART, J., POWELL, D. & GARDNER, T. 2008b. Evaluation of Bacteroides markers for the detection of human faecal pollution. *Letters in Applied Microbiology*, 46, 237-42.
- AHMED, W., YUSUF, R., HASAN, I., GOONETILLEKE, A. & 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, 838-845.
- BERNHARD, A. E. & FIELD, K. G. 2000. A PCR Assay To Discriminate Human and Ruminant Feces on the Basis of Host Differences in Bacteroides-Prevotella Genes Encoding 16S rRNA. *Applied and Environmental Microbiology*, 66, 4571-4574.
- BOEHM, A., VAN DE WERFHORST, L. C., GRIFFITH, J., HOLDEN, P., JAY, J., SHANKS, O., WANG, D. & WEISBERG, S. 2013. Performance of forty-one microbial source tracking methods: A twenty-seven lab evaluation study. *Water Research*, 47, 6812-6828.
- BURD, E. M. 2010. Validation of laboratory-developed molecular assays for infectious diseases. *Clinical Microbiology Reviews*, 23, 6.
- BUSTIN, S. A., BENES, V., GARSON, J. A., HELLEMANS, J., HUGGETT, J., KUBISTA, M., MUELLER, R., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J. & WITWER, C. T. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, 55, 611-22.
- CAO, Y., GRIFFITH, J. F., DOREVITCH, S. & WEISBERG, S. B. 2012. Effectiveness of qPCR permutations, internal controls and dilution as means for minimizing the impact of inhibition while measuring Enterococcus in environmental waters. *Applied Microbiology*, 113, 66-75.
- CAO, Y., RAITH, M. R. & GRIFFITH, J. F. 2015. Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. *Water Research* 70, 337-49.
- CAO, Y., SIVAGANESAN, M., KINZELMAN, J., BLACKWOOD, A. D., NOBLE, R. T., HAUGLAND, R. A., GRIFFITH, J. F. & WEISBERG, S. B. 2013. Effect

- of platform, reference material, and quantification model on enumeration of *Enterococcus* by quantitative PCR methods. *Water Research*, 47, 233-41.
- CODEX 2010. Guidelines on performance criteria and validation of methods for detection, identification and quantification of specific DNA sequences and specific proteins in foods. Rome.
- DICK, L. K. & FIELD, K. G. 2004. Rapid estimation of numbers of fecal Bacteroidetes by use of a quantitative PCR assay for 16S rRNA genes. *Applied and Environmental Microbiology*, 70, 5695-5697.
- EDGE, T. A., HILL, S., SETO, P. & MARSALEK, J. 2010. Library-dependent and library-independent microbial source tracking to identify spatial variation in faecal contamination sources along a Lake Ontario beach (Ontario, Canada). *Water Science And Technology: A Journal Of The International Association On Water Pollution Research*, 62, 719-727.
- ELLISON, S. L. R., ENGLISH, C. A., BURNS, M. J. & KEER, J. T. 2006. Routes to improving the reliability of low level DNA analysis using real-time PCR. *BMC Biotechnology*, 6, 33-33.
- FLOREN, C., WIEDEMANN, I., BRENIG, B., SCHUTZ, E. & BECK, J. 2015. Species identification and quantification in meat and meat products using droplet digital PCR (ddPCR). *Food Chemistry*, 173, 1054-8.
- FOGARTY, L. R. & VOYTEK, M. A. 2005. Comparison of *bacteroides*-prevotella 16S rRNA genetic markers for fecal samples from different animal species. *Applied and Environmental Microbiology*, 71, 5999-6007.
- GAWLER, A. H., BEECHER, J. E., BRANDÃO, J., CARROLL, N. M., FALCÃO, L., GOURMELON, M., MASTERSON, B., NUNES, B., PORTER, J. & RINCÉ, A. 2007. Validation of host-specific Bacteroidales 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. *Water Research*, 41, 3780.
- GREEN, H. C., HAUGLAND, R. A., VARMA, M., MILLEN, H. T., BORCHARDT, M. A., FIELD, K. G., WALTERS, W. A., KNIGHT, R., SIVAGANESAN, M., KELTY, C. A. & SHANKS, O. C. 2014. Improved HF183 Quantitative Real-Time PCR Assay for Characterization of Human Fecal Pollution in Ambient Surface Water Samples. *Applied and Environmental Microbiology*, 80, 9.
- GRGICAK, C. M., URBAN, Z. M. & COTTON, R. W. 2010. Investigation of reproducibility and error associated with qPCR methods using Quantifiler(R) Duo DNA quantification kit. *Forensic Sciences*, 55, 1331-9.

- GRIFFITH, J. F. & WEISBERG, S. B. 2011. Challenges in implementing new technology for beach water quality monitoring: Lessons from California demonstration project. *Marine Technology Society* 45, 65-73.
- HARRIS, D. C. 1995. *4th Edition: Quantitative Chemical Analysis*. New York, W. H. Freeman and Company.
- HARWOOD, V., STALEY, C., BADGLEY, B., BORGES, K. & KORAJKIC, A. 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews*, 38, 1-40.
- HARWOOD, V. J., BROWNELL, M., WANG, S., LEPO, J., ELLENDER, R. D., AJIDAHUN, A., HELLEIN, K. N., KENNEDY, E., YE, X. & FLOOD, C. 2009. Validation and field testing of library-independent microbial source tracking methods in the Gulf of Mexico. *Water Research*, 43, 4812-9.
- HAUGLAND, R. A., VARMA, M., SIVAGANESAN, M., KELTY, C., PEED, L. & SHANKS, O. C. 2010. Evaluation of genetic markers from the 16S rRNA gene V2 region for use in quantitative detection of selected *Bacteroidales* species and human fecal waste by qPCR. *Systematic and Applied Microbiology*, 33, 348-57.
- HAYDEN, R. T., GU, Z., INGERSOLL, J., ABDUL-ALI, D., SHI, L., POUNDS, S. & CALIENDO, A. M. 2013. Comparison of droplet digital PCR to real-time PCR for quantitative detection of cytomegalovirus. *Clinical Microbiology*, 51, 540-6.
- HINDSON, B. J., NESS, K. D., MASQUELIER, D. A., BELGRADER, P., HEREDIA, N. J., MAKAREWICZ, A. J., BRIGHT, I. J., LUCERO, M. Y., HIDDESEN, A. L., LEGLER, T. C., KITANO, T. K., HODEL, M. R., PETERSEN, J. F., WYATT, P. W., STEENBLOCK, E. R., SHAH, P. H., BOUSSE, L. J., TROUP, C. B., MELLEN, J. C., WITTMANN, D. K., ERNDT, N. G., CAULEY, T. H., KOEHLER, R. T., SO, A. P., DUBE, S., ROSE, K. A., MONTESCLAROS, L., WANG, S., STUMBO, D. P., HODGES, S. P., ROMINE, S., MILANOVICH, F. P., WHITE, H. E., REGAN, J. F., KARLIN-NEUMANN, G. A., HINDSON, C. M., SAXONOV, S. & COLSTON, B. W. 2011. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Analytical Chemistry*, 83, 8604-8610.
- HINDSON, C. M., CHEVILLET, J. R., BRIGGS, H. A., GALLICHOTTE, E. N., RUF, I. K., HINDSON, B. J., VESSELLA, R. L. & TEWARI, M. 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nature Methods*, 10, 1003-5.
- HUGGETT, J. F., FOY, C. A., BENES, V., EMSLIE, K., GARSON, J. A., HAYNES, R., HELLEMANS, J., KUBISTA, M., MUELLER, R. D., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J., WITTWER, C. T. & BUSTIN, S. A. 2013. The digital MIQE guidelines: Minimum Information

for Publication of Quantitative Digital PCR Experiments. *Clinical Chemistry*, 59, 892-902.

- KILDARE, B., LEUTENEGGER, C., MCSWAIN, B., BAMBIC, D., RAJAL, V. & WUERTZ, S. 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal *Bacteroidales*: A Bayesian approach. *Water Research*, 41, 15.
- LAYTON, B., CAO, Y., EBENTIER, D., HANLEY, K., BALLESTÉ, E., BRANDÃO, J., BYAPPANAHALLI, M., CONVERSE, R., FARNLEITNER, A., GENTRY SHIELDS, J., GIDLEY, M., GOURMELON, M., LEE, C., LEE, J., LOZACH, S., MADI, T., MEIJER, W., NOBLE, R., PEED, L., REISCHER, G., RODRIGUES, R., ROSE, J., SCHRIEWER, A., SINIGALLIANO, C., SRINIVASAN, S., STEWART, J., VAN DE WERFHORST, L. C., WANG, D., WHITMAN, R., WUERTZ, S., JAY, J., HOLDEN, P., BOEHM, A., SHANKS, O. & GRIFFITH, J. 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. *Water Research*, 47, 6897-6908.
- MCDERMOTT, G., DO, D., LITTERST, C., MAAR, D., HINDSON, C., STEENBLOCK, E., LEGLER, T., JOUVENOT, Y., MARRS, S., BEMIS, A., SHAH, P., WONG, J., WANG, S., SALLY, D., JAVIER, L., DINIO, T., HAN, C., BRACKBILL, T., HODGES, S., LING, Y., KLITGORD, N., CARMAN, G., BERMAN, J., KOEHLER, R., HIDDESEN, A., WALSE, P., BOUSSE, L., TZONEV, S., HEFNER, E., HINDSON, B., CAULY, T., HAMBY, K., PATEL, V., REGAN, J., WYATT, P., KARLIN NEUMANN, G., STUMBO, D. & LOWE, A. 2013. Multiplexed target detection using DNA-binding dye chemistry in droplet digital PCR. *Analytical chemistry*, 85, 11619-27.
- MCQUAIG, S. M., SCOTT, T. M., LUKASIK, J. O., PAUL, J. H. & HARWOOD, V. J. 2009. Quantification of Human Polyomaviruses JC Virus and BK Virus by TaqMan Quantitative PCR and Comparison to Other Water Quality Indicators in Water and Fecal Samples. *Applied and Environmental Microbiology*, 75, 3379-3388.
- MOLINA, M., HUNTER, S., CYTERSKI, M., PEED, L. A., KELTY, C. A., SIVAGANESAN, M., MOONEY, T., PRIETO, L. & SHANKS, O. C. 2014. Factors affecting the presence of human-associated and fecal indicator real-time quantitative PCR genetic markers in urban-impacted recreational beaches. *Water Research*, 64, 196-208.
- MORISSET, D., STEBIH, D., MILAVEC, M., GRUDEN, K. & ZEL, J. 2013. Quantitative analysis of food and feed samples with droplet digital PCR. *PLoS One*, 8, e62583.
- NEAVE, M., LUTER, H., PADOVAN, A., TOWNSEND, S., SCHOBEN, X. & GIBB, K. 2014. Multiple approaches to microbial source tracking in tropical northern Australia. *Microbiology Open*, 3, 860-874.

- NOBLE, R. T., BLACKWOOD, A. D., GRIFFITH, J. F., MCGEE, C. D. & WEISBERG, S. B. 2010. Comparison of Rapid Quantitative PCR-Based and Conventional Culture-Based Methods for Enumeration of *Enterococcus* spp. and *Escherichia coli* in Recreational Waters. *Applied and Environmental Microbiology*, 76, 7437-7443.
- NSHIMYIMANA, J. P., CRUZ, M. C., THOMPSON, R. J. & WUERTZ, S. 2016. *Bacteroidales* Markers for Microbial Source Tracking in Southeast Asia. *Water Research*, 118, pp. 239-248.
- NSHIMYIMANA, J. P., EKKLESIA, E., SHANAHAN, P., CHUA, L. H. C. & THOMPSON, J. R. 2014. Distribution and abundance of human-specific *Bacteroides* and relation to traditional indicators in an urban tropical catchment. *Applied Microbiology*, 116, 1369-1383.
- ODAGIRI, M., SCHRIEWER, A., HANLEY, K., WUERTZ, S., MISRA, P. R., PANIGRAHI, P. & JENKINS, M. W. 2015. Validation of *Bacteroidales* quantitative PCR assays targeting human and animal fecal contamination in the public and domestic domains in India. *Science of the Total Environment*, 502, 462-470.
- PINHEIRO, L. B., COLEMAN, V. A., HINDSON, C. M., HERRMANN, J., HINDSON, B. J., BHAT, S. & EMSLIE, K. R. 2012. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Analytical Chemistry*, 84, 1003-11.
- RACKI, N., MORISSET, D., GUTIERREZ-AGUIRRE, I. & RAVNIKAR, M. 2014. One-step RT-droplet digital PCR: a breakthrough in the quantification of waterborne RNA viruses. *Analytical and Bioanalytical Chemistry*, 406, 661-7.
- SAUER, E., VANDEWALLE, J., BOOTSMA, M. & MCLELLAN, S. 2011. Detection of the human specific *Bacteroides* genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. *Water Research*, 45, 4081-4091.
- SEURINCK, S., DEFOIRDT, T., VERSTRAETE, W. & SICILIANO, S. 2005. Detection and quantification of the human-specific HF183 *Bacteroides* 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environmental Microbiology*, 7, 249-59.
- SHANKS, O., ATIKOVIC, E., BLACKWOOD, A. D., LU, J., NOBLE, R., DOMINGO, J., SEIFRING, S., SIVAGANESAN, M. & HAUGLAND, R. 2008. Quantitative PCR for detection and enumeration of genetic markers of bovine fecal pollution. *Applied and Environmental Microbiology*, 74, 745-52.

- SHANKS, O., KELTY, C., SIVAGANESAN, M., VARMA, M. & HAUGLAND, R. 2009. Quantitative PCR for genetic markers of human fecal pollution. *Applied and Environmental Microbiology*, 75, 5507-13.
- SHANKS, O., WHITE, K., KELTY, C., SIVAGANESAN, M., BLANNON, J., MECKES, M., VARMA, M. & HAUGLAND, R. 2010. Performance of PCR-based assays targeting *Bacteroidales* genetic markers of human fecal pollution in sewage and fecal samples. *Environmental Science and Technology*, 44, 6281-8.
- SHANKS, O. C., SIVAGANESAN, M., PEED, L., KELTY, C. A., BLACKWOOD, A. D., GREENE, M. R., NOBLE, R. T., BUSHON, R. N., STELZER, E. A., KINZELMAN, J., ANAN'EVA, T., SINIGALLIANO, C., WANLESS, D., GRIFFITH, J., CAO, Y., WEISBERG, S., HARWOOD, V. J., STALEY, C., OSHIMA, K. H., VARMA, M. & HAUGLAND, R. A. 2012. Interlaboratory comparison of real-time PCR protocols for quantification of general fecal indicator bacteria. *Environmental Science & Technology*, 46, 945-53.
- SIVAGANESAN, M., SIEFRING, S., VARMA, M. & HAUGLAND, R. A. 2011. MPN estimation of qPCR target sequence recoveries from whole cell calibrator samples. *Microbiological Methods*, 87, 343-9.
- STEWART, J., BOEHM, A., DUBINSKY, E., FONG, T.-T., GOODWIN, K., GRIFFITH, J., NOBLE, R., SHANKS, O., VIJAYAVEL, K. & WEISBERG, S. 2013. Recommendations following a multi-laboratory comparison of microbial source tracking methods. *Water Research*, 47, 6829-38.
- TAYLOR, S. C., CARBONNEAU, J., SHELTON, D. N. & BOIVIN, G. 2015. Optimization of Droplet Digital PCR from RNA and DNA extracts with direct comparison to RT-qPCR: Clinical implications for quantification of Oseltamivir-resistant subpopulations. *Virological Methods*, 224, 58-66.
- TE, S. H., CHEN, E. Y. & GIN, K. Y.-H. 2015. Comparison of Quantitative PCR and Droplet Digital PCR Multiplex Assays for Two Genera of Bloom-Forming Cyanobacteria, *Cylindrospermopsis* and *Microcystis*. *Applied and Environmental Microbiology*, 81, 5203-5211.
- USEPA 2012a. Method 1611: Enterococci in Water by TaqMan Quantitative Polymerase Chain Reaction (qPCR) Assay. Washington DC: USEPA.
- USEPA 2012b. Recreational Water Quality Criteria. Washington DC: USEPA.
- USEPA 2013. Method 1609: Enterococci in water by TaqMan® Quantitative Polymerase Chain Reaction (qPCR) with Internal Amplification control (IAC) assay. In: AGENCY, U. S. E. P. (ed.). Washington: U.S. Environmental Protection Agency.

- VERHAEGEN, B., DE REU, K., DE ZUTTER, L., VERSTRAETE, K., HEYNDRIKX, M. & VAN COILLIE, E. 2016. Comparison of Droplet Digital PCR and qPCR for the Quantification of Shiga Toxin-Producing *Escherichia coli* in Bovine Feces. *Toxins*, 8.
- WHALE, A. S., COWEN, S., FOY, C. A. & HUGGETT, J. F. 2013. Methods for applying accurate digital PCR analysis on low copy DNA samples. *PLoS One*, 8, e58177.
- YAMPARA-IQUISE, H., ZHENG, G., JONES, J. E. & CARSON, C. A. 2008. Use of a *Bacteroides thetaiotaomicron*-specific alpha-1-6, mannanase quantitative PCR to detect human faecal pollution in water. *Applied Microbiology*, 105, 1686-93.
- YANG, R., PAPARINI, A., MONIS, P. & RYAN, U. 2014. Comparison of next-generation droplet digital PCR (ddPCR) with quantitative PCR (qPCR) for enumeration of *Cryptosporidium* oocysts in faecal samples. *International Journal for Parasitology* 44, 1105-13.

Chapter 6

Decay of Human-associated *Bacteroidales* Viable Cells and DNA, and Fecal Indicator Bacteria in Tropical Freshwater Sediments.

Abstract

We examined the sediment-associated decay rates (under dark and anoxic conditions) of two previously validated human-associated microbial source tracking markers for anaerobic *Bacteroidales* (*B. thetaiotaomicron*-specific α -1-6, mannanase (B. theta) and BacHum). Additionally, we monitored the number of *E. coli*, *Enterococcus* and heterotrophic protozoa. Tropical freshwater sediments were used to create over 160 microcosms of one treatment and three controls (no-sewage, sewage nutrients (autoclaved sewage), and $< 5 \mu\text{m}$ pore size filtered supernatant), incubated in the dark under anoxic conditions at 30° C. Microcosms were sampled eight times in 20 d and analyzed for protozoa by DAPI staining and fluorescence microscopy, for FIB by Colilert and Enterolert, and for human-associated *Bacteroidales* (PMA treated and not PMA treated) by droplet digital PCR. Human markers and FIB demonstrated first order exponential decay. Viable *Bacteroidales* cells and DNA as measured by B. theta and BacHum survived the shortest in sediments ($T_{99} < 9 \text{ d}$), while FIB survived the longest ($T_{99} < 14.5 \text{ d}$). *Bacteroidales* viable cells exhibited slower decay rates in the water column ($T_{99} < 6 \text{ d}$) than did FIB ($T_{99} < 2.5 \text{ d}$). Protozoa were found to influence the survival of FIB in sediments, but not of human-associated gene markers. This work supports the use of human markers for predicting health risks associated with recreational uses of waterways and waterbodies in the tropics.

6.1 Introduction

Sediments at recreational beaches have been reported as a potential non-point source of fecal contamination due to prolonged survival of fecal contaminant and pathogens (Halliday and Gast, 2011). The annual worldwide estimation of sickness related to coastal bathing contaminated with wastewater has been estimated at 120 million cases of gastrointestinal illness and 50 million cases of respiratory diseases (Shuval, 2003, Halliday and Gast, 2011). These bathing-associated diseases are prevalent due to continuous fecal contamination of recreational beaches and a prolonged survival of fecal contamination in beach sediments (Halliday and Gast, 2011, Davies et al., 1995). Survival of fecal bacterial cells in sediments is enhanced by availability of nutrients accumulated from algae, debris and plankton (Halliday and Gast, 2011, Davies et al., 1995). In addition, benthic biofilms, the autochthonous bacterial community, sunlight or UV inactivation, and moderate temperature are known to contribute to the survival of fecal bacteria and pathogens in sediments (Davies et al., 1995, Priester et al., 2007, Feng et al., 2010, Nealson, 1997).

Host-associated *Bacteroidales* genetic markers are among the most commonly used microbial source tracking (MST) assays and have been applied in a variety of different geographical locations (Layton et al., 2013, Ahmed et al., 2010, Ahmed et al., 2009, Jenkins et al., 2009, Odagiri et al., 2015, Kildare et al., 2007, Yampara-Iquise et al., 2008, Toledo-Hernandez et al., 2013, Gawler et al., 2007, Schriewer et al., 2013, Green et al., 2012, Ahmed et al., 2008a, Ahmed et al., 2008b, Harwood et al., 2009, Reischer et al., 2013, Neave et al., 2014, Ervin et al., 2014, Okabe et al., 2007). While the effect of environmental factors such as temperature, salinity, and sunlight on the survival and

persistence of *Bacteroidales* cells and DNA, respectively, in the water column has been extensively studied (Bae and Wuertz, 2009b, Bae and Wuertz, 2014, Bae and Wuertz, 2015, Liang et al., 2017, Schulz and Childers, 2011, Green et al., 2011, Okabe and Shimazu, 2007, Kobayashi et al., 2013, Walters and Field, 2009, Bae and Wuertz, 2012) relatively little is known about their fate in sediments. A recent study examined the survival and persistence of host-associated *Bacteroidales* cells and DNA in comparison with *E. coli* and *Enterococcus* in freshwater water sediments at 6° C and 20° C (Kim and Wuertz, 2015). The authors concluded that sediments could act as a non-point source of fecal genetic markers. Despite these findings, sediments have received little attention as potential reservoirs of source tracking markers, especially in the tropical environment (Kim and Wuertz, 2015).

The role of protozoa on the decay of FIB (e.g. *E. coli*) in sediments and water column under aerobic and ambient conditions has been established in temperate and sub-tropical regions (Anderson et al. 2005, Wanjugi et al., 2016, Dick et al., 2010b, Korajkic et al., 2013a, Korajkic et al., 2013b). Studies suggested that the presence of protozoa plays a role in regulating the persistence of FIB (e.g. *E. coli*) in sediments. However, the effect of protozoa on human-associated *Bacteroidales* (anaerobic bacteria) is yet to be determined as life of protozoa under anoxic conditions has been established (Fenchel et al. 1990, Fenchel and Finlay, 1990 and 1995, Sacca et al. 2009.).

Validation of human-associated *Bacteroidales* markers (B. theta and BacHum) for microbial sources tracking in Singapore and Southeast Asia provided a set of markers for the detection, tracking, and monitoring sewage pollution in surface water (see Chapter 4). However, the decay of human-associated *Bacteroidales* markers in tropical

freshwater sediments is yet to be determined. The objectives of this study were (1) to examine sediment-associated decay rates of two previously validated human-associated microbial source tracking markers for *Bacteroidales*, *B. theta* and BacHum, and of the fecal indicator bacteria (FIB) (*E. coli* and *Enterococcus*) and (2) to evaluate the role of protozoa in decay.

6.2 Methodology

6.2.1 Experimental design

6.2.1.1 Microcosm Design

Freshwater and sediments were collected from a pristine tropical freshwater reservoir in Singapore (MacRitchie Reservoir, 1.3448°N, 103.8224°E) and used for preparing microcosms. Sediments were mixed with water (2 to 1 ratio) by shaking to ensure a maximum protozoa resuspension in the water phase. Mixed sediments were then settled for 10 min and the supernatant was removed and stored in clean containers (24 hrs at ambient temperature) (Anderson et al. 2005) until used for reseeded protozoa in the microcosms. Sediments were then autoclaved three times at 120 °C for 45 min each time to reach sterile conditions and ensure that all protozoa in the microcosms were derived from sediment supernatant. Raw sewage was filtered using a 5-µm cellulose acetate filter (Sartorius Stedim Biotech, Gottingen, Germany) to limit the introduction of additional protozoa during inoculation. Microcosms were then prepared using autoclaved sediments, sediment supernatant for the reseeded protozoa, water from the sampling site, and sewage for inoculation.

Microcosms designed for this study included one treatment and three controls (Table 1). (1) The treatment microcosm was designed to have grazing by the addition of sediment supernatant, reservoir water, autoclaved sediments and sewage. (2) A no-sewage control was designed to examine natural grazing in sediment supernatant used in this study and was achieved by mixing autoclaved sediments, sediment supernatant and reservoir water. (3) A sewage nutrients control (autoclaved sewage) was used to examine if nutrients contained in sewage could have effect on bacterial persistence in microcosms.

The nutrient control microcosm was created by combining: autoclaved sediments, autoclaved sewage, sediments supernatant and reservoir water. (4) The $< 5\mu\text{m}$ control was designed to examine if the size-based reduction of grazing in the system could have effect on the decay of FIB and human-associated markers (BacHum and B. theta). The $< 5\mu\text{m}$ control consisted of autoclaved sediments, filtered sewage and sediment supernatant filtered through a $5\text{-}\mu\text{m}$ cellulose acetate filter (Sartorius Stedim Biotech, Gottingen, Germany) (Table 1). The mixing ratios were 18 ml of filtered or autoclaved sewage (to achieve an initial concentration $\geq 10^4/100\text{ml}$), 60 ml of water column and 60 g of wet autoclaved sediment. All microcosms were wrapped in aluminum foil to ensure dark conditions before incubation. Microcosms were then incubated under dark and anoxic conditions at 30°C for 20 days (Table 1).

6.2.1.2 Microcosm sample collection

A total of 160 microcosms were prepared using 250-ml Falcon flask bottles with a canted neck (0.2 mm), plug seal cap and blue vented cap (Corning incorporated, Corning, New York, USA) (Sewage: 18mL , water column:60 mL, sediments: 60g or

2cm of depth). Five replicates were created and triplicate microcosms were sacrificed for each treatment or control on each sampling day (Sampling days: 0, 1, 2, 3, 5, 7, 10, 14 and 20 days). Falcon flask bottle microcosms wrapped with aluminum foil were gently collected from the incubator, unwrapped without disturbing or shaking them and then sampled.

The water column was sampled first followed by the sediment supernatant. To collect sediment supernatant samples, microcosms were shaken for 2 minutes for cell detachment and resuspension, settled for 10 minutes before collecting samples (Kim et al., 2014). Samples were collected from the water column and sediments supernatant for FIB enumeration (10 ml each sample), PMA-treated DNA for quantifying intact *Bacteroidales* cells (1 ml each sample), non-PMA treated DNA for total human *Bacteroidales* (10 ml each sample), and protozoa for microscopy count (1 ml each sample) (Table 1). Protozoa samples were fixed using 40% glutaraldehyde (1:10 dilution for 1-ml sample). Non-PMA treated water column and sediment supernatant samples were collected using a Falcon tube (Corning incorporated, Corning, New York, USA), then stored at -80°C. Before further processing for DNA extraction, samples were filtered using Millipore Sterivex™-GS 0.22 µm Filter Units (Millipore, Billerica, MA, USA). PMA-treated samples were store in -20 degree freezer until processing. Physicochemical parameters were measured using a miniDO₂T Logger (Precision Measurement Engineering Inc., California, USA) fixed onto the falcon flask and recorded the temperature, pH, dissolved Oxygen (DO), and oxygen levels (Pictures).

Table 1. Conditions and criteria for the design of sediments microcosms under anoxic conditions

Set up Criteria	Treatment (T)	Control 1: No sewage	Control 2: Nutrients (Autoclaved sewage)	Control 3: Filtered (< 5µm Control)
Temperature	30°C	30°C	30°C	30°C
Environment	Anoxic	Anoxic	Anoxic	Anoxic
Conditions	Dark	Dark	Dark	Dark
Grazing	Grazing	Natural unconditioned grazing in the matrix used for this study	Reduced grazing	Reduced grazing and understanding microbial community
	Autoclaved sediments	Autoclaved sediments	Autoclaved sediments	Autoclaved sediments
	Sediments supernatant	Sediments supernatant	Sediments supernatant	Filtered sediments supernatant (5µm)
Variables	Filtered sewage (5µm)		Autoclaved sewage	Settled filtered sewage (5µm)
	Water from a tropical reservoir in Singapore	Water from a tropical reservoir in Singapore	Water from a tropical reservoir in Singapore	Water from a tropical reservoir in Singapore
Major set up variables	Sediment supernatant, Filtered sewage (5 µm)	Sediments supernatant	Autoclaved sewage	Settled filtered sewage (5µm), Filtered sediments supernatant (5µm)
Constant variables	Autoclaved sediments	Autoclaved sediments	Autoclaved sediments	Autoclaved sediments

6.2.2 Enumeration of Fecal indicator bacteria

Fecal indicator bacteria (Total coliforms, *E. coli* and *Enterococcus*) were enumerated using Colilert and Colilert Quanti-Tray®2000 (IDEXX Laboratories, Westbrook, ME, USA) for total coliform and *E. coli*, and Enterolert™ (IDEXX Laboratories, Westbrook, ME, USA) for *Enterococcus*. Generated concentrations were expressed in Most Probable-Number (MPN/100ml). Enumeration of blank samples for *E. coli* and *Enterococcus* was done day during sampling period and showed negative results for both Colilert and Enterolert tests, thus confirming sterile work conditions and absence of cross contamination during sample processing.

6.2.3 PMA treatment

PMA was used following the protocol previously discussed by Bae and Wuertz (2009) (Bae and Wuertz, 2009a) with modification of the mechanisms used for light exposure. Briefly, PMA (Biotium Inc., Hayward, CA, USA) was dissolved in 20% of Dimethyl Sulfoxide (DMSO) and stocked at a concentration of 1mg/ml at -20°C into a light opaque bags. PMA was added onto 0.5 ml of the sample at a final concentration of 100 µM (Bae and Wuertz, 2009a). After adding PMA, samples were wrapped in aluminum foil, incubated for 5min and then exposed to light for 10 min using PMALite™ (Biotium Inc., Hayward, CA, USA). The PMA treated samples collected from all microcosms were stored in -20 degree freezer until DNA extraction.

6.2.4 DNA Extraction

Sediments supernatant and water column were DNA extracted using the PowerWater® DNA isolation kit (MO BIO Laboratories Inc., Carlsbad, CA, USA) following the manufacturer's protocol. Briefly, Sterivex filter cartridges membrane were aseptically removed, sliced in 15-20 strips using sterile blade and forceps. The membrane strips were then loaded into beads tubes and DNA was extracted using PowerWater® DNA isolation kit. For PMA-treated water, the entire sample volume (0.5ml) was loaded into 5ml beads tube and DNA was extracted using PowerWater® DNA isolation kit (MO BIO Laboratories Inc., Carlsbad, CA, USA) following manufactures protocol. Briefly, cells were lysed using PowerWater® beads loaded with lysing buffer pre-heated at 55°C for 10min. Non-PMA treated samples used a slightly modified protocol to increase the yield of protozoa lysing. A proteinase K digested step was added and 20µl Proteinase K (20mg/ml) was loaded to into the sample with bead beating and heated samples at 65°C for 30min. The following steps continued with the reduction of inhibitors by removing non-DNA organic and inorganic materials including humic acids using Inhibitor Removal Technology® (IRT). The supernatant was then treated by concentrated salt solutions to increase the binding of DNA to silica and increase the removal of the remaining inhibitors by spin filtration. All DNA samples were eluted in 100µl and their concentrations were measured using a nanoDrop-ND-1000 Spectrophotometer (NanoDrop Technologies, Wilmington, DE, USA) for quality and quantity verification, and stored in -20° C until quantification.

6.2.5 Droplet digital PCR (ddPCR) assays

The use of ddPCR followed the manufactures recommendations. Briefly, all B. theta and BacHum assays reaction mixes contained each 10µl of ddPCR™ Supermix for probes (NodUTP), primers and probes final concentration in the mix was 900nM and 250nM respectively, and 2µl genomic DNA. The 20µl of each sample mixture was loaded onto D8™ cartridge (Bio-Rad, Hercules, CA, USA) by 8 reactions at time and 70µl of ddPCR™ droplet generator oil (Bio-Rad, Hercules, CA, USA) for each mixture. The cartridges were then covered by gaskets rubber and placed into QX200™ Droplet generator machine (Bio-Rad, Hercules, CA, USA). After droplet generation, 40µl of each mixture were loaded to a 96 well-plate and sealed using PX1™ Plate Sealer (Bio-Rad, Hercules, CA, USA). Briefly, ddPCR assays were subjected to a thermocycle with enzyme activations at 95°C for 10min; denaturation 94°C for 30 seconds, annealing temperature 60°C for 1min; enzyme deactivation 98°C for 10 min followed by 40 cycles and a continuous hold at 4°C. Quantification results were read using ddPCR™ Droplet Reader Oil (Bio-Rad, Hercules, CA, USA); QX200™ Droplet Reader (Bio-Rad, Hercules, CA, USA) and QuantaSoft™ software (Bio-Rad, Inc., Hercules, CA, USA). Quality control for ddPCR was implemented as previous discussed by Cao and colleagues (Cao et al., 2015, Huggett et al., 2013, Pinheiro et al., 2012). Briefly, the number of accepted droplets generated per each sample run by ddPCR in this study was $\geq 10,000$. Positive concentrations of samples (positive controls and environmental samples) quantified in this study had a dynamic range from 0.12% to 99.0% positive droplets (equivalent to 1.5 to 10^4 of target copies/µl). Quantified replicates of each sample were closely similar with a coefficient of variation ≤ 0.05 (5%) for concentration below 100 copies/µl and ≤ 0.02 (2%) for concentrations above 100 copies/µl, but below

10⁴copies/μl across this study. The limit of detection were determined using LOD95 method used in chemical analysis and PCR application for and clinical diagnostics (Stewart et al., 2013, CODEX, 2010, Burd, 2010). For this study at least 95% of the tests at lower concentrations were above 1.5 or 5.0 gene copies/μl respectively for *B. theta* and BacHum. Concentrations of Human-associated *Bacteroidales* markers were expressed in gene copies/100ml. The reproducibility was confirmed by detecting and quantifying similar concentrations of a plasmid DNA standard (10³ copies/μl) containing the target for all run ddPCR plates (3±0.05 log copies/μl and 1.5 % CV).

6.2.6 Microscopic enumeration of protozoa

Protozoa such as flagellates, ciliates, and amoeba within sediments supernatant were enumerated using DNA staining dye combined with fluorescence microscopy. Samples from the water column were not examined as protozoa were in the lower number and not detectable by this method. To determine the number of protozoa within treated sediments supernatant samples, nucleic acid staining combined with fluorescent microscopy was used similar to as described by Hoffman and Atlas (1987). Sediments supernatant samples were preserved in 4 % glutaraldehyde and stored at 4 °C. Samples were diluted 10:1 with 1 × phosphate buffered saline (PBS), mixed and 100 μl added to a 96 well plate. Ten microliters of Lugol's iodine was added to enhance settlement of sediment supernatant. To stain eukaryotic DNA, 4', 6-diamidino-2-phenylindole (DAPI) was added at a concentration of 50 μM and the mixture was incubated at room temperature for 30 min.

Preparations were examined with epifluorescence microscopy, using an Axio Observer Z.1 wide field inverted microscope (Carl Zeiss, Oberkochen, Germany) fitted with a

DAPI filter observed at 50 × objective lens. The average number of protozoa > 10 µm was determined by counting the number of DAPI-stained nuclei in 15 × images. Images were processed by ZEN software (blue version, 2012) and analyzed by Image J software (v1.49). Concentrations of protozoa were reported first as protozoa per field of view (1.80 mm × 1.35 mm) +/- SD and the unit of volume was then included to report protozoa counts per 100 ml (Figure 2).

6.2.7 Decay rate calculations and statistical analysis

Collinearity in data processing was prevented by removing total coliforms (TC) counts in the final data processing as TC concentrations correlated with *E. coli* concentrations ($R > 0.5$). The decay rates of culturable fecal indicator bacteria (FIB) (*Enterococcus* and *E. coli*), viable *Bacteroidales* (B. theta and BacHum) were calculated using the first order exponential decay model ($N=N_0 \times e^{-kt}$) (Chick, 1908) where N is the concentration of FIB (MPN/100ml) or of human markers (B. theta and BacHum) Gene copies/100ml at time t (expressed in days), N_0 is the initial concentration, and k is the decay rate constant. The time for a 2-log reduction (T_{99}) was calculated based on the decay rate constant of the first order exponential decay used in this study. The lag-period was removed for FIB and human markers and all the decay rates were then calculated from the end of the lag period where necessary. All the data analysis related to the concentrations of human markers (B. theta and BacHum) were performed using only concentrations above the limit of detection (LOD). None-linear regression analysis and fitting was performed in R using Version 3.2.4 revised (2016-03-16 r70336). Constant decay rates (k) and T_{99} were compared across microcosms matrices (water column and sediments) using ANOVA and multiple comparison tests. Standard deviation and geometric means were calculated from three independent replicates of each sample using JMP Pro (SAS Institute Inc., Cary, NC USA).

6.3 Results

6.3.1 Physicochemical parameters

Microcosms achieved environmental set up factors (DO for anoxic conditions, and pH) over the course of the incubation period (20 d). Physicochemical parameters measured in a microcosm setup bottle using the miniDOT over the course of the incubation period (20 d) demonstrated that anoxic conditions were achieved ($\text{DO} = 0.12 \pm 0.005 \text{ mg/l}$) ($\leq 0.5 \text{ mg/l}$) (USGS, 2016). The pH changed gradually from 6 at the beginning for the experiment to 7.8 at the last sampling day. The overall pH average for the duration of microcosm's incubation was 6.85 ± 0.132 . The microcosm temperature inside the incubator remained at 30°C for 20 d (Table 2).

Table 2. Chemical parameters measured in the microcosm using the miniDO₂T logger (Precision Measurement Engineering Inc., California, USA)

Sampling Days	pH		DO (mg/l)	
	Geomean	SD	Geomean	SD
1	6.31	0.148	0.46	1.578
2	6.43	0.148	0.05	0.002
3	6.52	0.098	0.05	0.001
5	6.70	0.105	0.05	0.001
7	7.03	0.170	0.05	0.020
10	7.29	0.101	0.05	0.001
14	7.30	0.184	0.05	0.001
20	7.36	0.125	2.92	0.188

DO: Dissolved Oxygen; **SD:** Standard Deviation, **pH:** Potential of Hydrogen

6.3.2 Human-associated Bacteroidales markers (B. theta and BacHum) demonstrated a faster decay rate in sediments

The decay rates of human-associated *Bacteroidales* (BacHum and B. theta) in tropical sediments were examined by quantifying viable cells (PMA-treated DNA samples) and total DNA (from dead and live cells, No-PMA treatment) in microcosms prepared using sediments collected in a tropical freshwater reservoir. All decay profiles of human-associated *Bacteroidales* markers (BacHum and B. theta) for PMA and non-PMA treated samples exhibited a first order exponential decay ($N=N_0 \cdot e^{-kt}$) with $R^2 > 0.96$. Overall, the decay rates of DNA from viable cells and total DNA measured by PMA-ddPCR and No-PMA-ddPCR, respectively for B. theta and BacHum had decayed rates in the sediments different from decay rates in the water for either marker (B. theta or BacHum) quantified targeting viable cells (PMA-treated) or total DNA (Not-PMA treated) (comparison of T_{99} , PMA treated: $F=1.10$, $p = 0.033$ and not-PMA: $F = 0.07$, $p = 0.780$) (Table 4).

The decay of human-associated *Bacteroidales* in PMA-treated samples in sediments reached a 2-log removal by 9 d (Table 3), while it took up to 6 d to reach a 2-log removal in overlaying water (Table 3). BacHum and B. theta concentrations measured in PMA-treated samples from sediments demonstrated a closer range in 2-log removal timeframes varying from 4 to 9 d (Table 3). B. theta had a 2-log removal time of 5.5 and 9 d respectively in the limited grazing ($< 5 \mu\text{m}$ control) and the full grazing (treatment) microcosms, while BacHum 2-log removal under similar conditions was 6.3 d and 4 d (Table 3). BacHum and B. theta in overlaying water demonstrated similar decay profiles and 2-log removal ranging from 3.5 to 6 d (Figure 5 and Figure 6). B.

theta and BacHum demonstrated a similar 2-log removal time of 6 days in overlaying water of the grazing microcosm (treatment), while their decay rate in overlaying water of the limited grazing ($< 5 \mu\text{m}$ control) microcosms reached a 2-log removal in 3 days (Table 3).

Decay profiles of BacHum and B. theta measured in non-PMA treated samples had 2-log removal of total DNA with a range comparable to that observed in PMA treated samples (from 2.5 to 8.5 d) in sediments and overlaying water (Table 3). Both human-associated markers (B. theta and BacHum) demonstrated a faster decay in sediments ranging from 2.3 d up to 6.3 d (Table 3). B. theta decay rate in sediments reached 2-log removal in 4.5 d, while the same removal took 2.3 d for BacHum in sediment treatment microcosm.

The 2-log removal of BacHum and B. theta was respectively reached in 4.7 and 6.3 d in the limited grazing microcosm ($< 5 \mu\text{m}$ control). B. theta decay reached a 2-log removal up to 8 d, while BacHum only needed 7 d (Table 3). The decay in overlay water was faster for BacHum reaching a 2-log removal in 4 d in the treatment microcosm and in 2.4 d in the $< 5 \mu\text{m}$ control microcosm. B. theta took up 7 d in overlaying water in the treatment microcosm and 8.5 d in $< 5 \mu\text{m}$ control microcosm (Table 3, Figure 5 and 6). Human-associated markers (BacHum and B. theta) were not detected in PMA-treated and non-PMA treated samples collected in no-sewage and nutrient controls. The non-detection of human markers in controls confirmed that the nutrient in sewage did not add any effect on the microcosm and that sediments used for set up did not have a detectable number of autochthonous *Bacteroidales*.

Table 3. Kinetic parameters of fecal indicator bacteria (*Enterococcus* and *E. coli* measured by IDEXX Enterolert and Colilert respectively), intact cells (PMA-Treated) and total DNA of human-associated *Bacteroidales* genetic markers (BacHum and B. theta) in sediment microcosms

Matrices	Targets	Water column			Sediments		
		k ¹ (d ⁻¹)	T ₉₉ (d)	R ² (p<0.001)	k (d ⁻¹)	T ₉₉ (d)	R ² (p<0.001)
Treatment	<i>E. coli</i>	1.25 (±0.06)	3.68	0.99	1.10 (±0.22)	4.20	0.99
	<i>Enterococcus</i>	2.69 (±0.04)	1.72	0.99	0.44 (±0.04)	10.3	0.98
	PMA-ddPCR B. theta	0.79 (±0.06)	5.78	0.99	0.52 (±0.07)	8.77	0.96
	PMA-ddPCR BacHum	0.78 (±0.05)	5.86	0.93	1.20 (±0.15)	3.82	0.97
	Non-PMA-ddPCR B. theta	0.64 (±0.08)	7.18	0.99	1.02 (±0.05)	4.51	0.96
	Non-PMA-ddPCR BacHum	1.10 (±0.04)	4.18	0.93	1.99 (±0.11)	2.31	0.97
Control (< 5 µm)	<i>E. coli</i>	2.17 (±0.18)	2.12	0.99	0.32 (±0.05)	14.4	0.96
	<i>Enterococcus</i>	2.12 (±0.04)	2.17	0.99	1.56 (±0.03)	3.00	0.99
	PMA-ddPCR B. theta	1.32 (±0.13)	3.49	0.98	0.85 (±0.05)	5.37	0.99
	PMA-ddPCR BacHum	1.26 (±0.07)	3.50	0.99	0.72 (±0.03)	6.33	0.99
	Non-PMA-ddPCR B. theta	0.54 (±0.07)	8.47	0.98	0.72 (±0.02)	6.32	0.99
	Non-PMA-ddPCR BacHum	1.95 (±0.04)	2.42	0.99	0.97 (±0.05)	4.73	0.99

¹ Decay rate constant

²Standard error of the mean for non-linear first degree exponential decay model; T₉₉: 2-log removal calculated based on the first degree exponential decay model ($N=N_0 \cdot e^{-kt}$) ($T_{99} = \ln(0.01)/k$); and R²: Predictive squared correlation coefficients demonstrating the best fits at p<0.001 generated based on the iteration of convergence comparing the model to the measured data.

Figure 1. Intact cells (PMA-ddPCR) over the period of 20 days as quantified using human markers B. theta (A and C) and BacHum (B and D). Decay of B. theta (PMA-ddPCR) in treatment and < 5µm control of water column (A). Decay of BacHum (PMA-ddPCR) in treatment and < 5µm control of water column (B). Decay of B. theta (PMA-ddPCR) in treatment and < 5µm control of sediments microcosms (C). Decay of BacHum (PMA-ddPCR) in treatment and < 5µm control of sediments (D). Error bars represent the standard deviation of biological triplicates of each sampling day.

Figure 2: Total DNA over the period (intact cells and dead cells) of 20 days as quantified using human markers *B. theta* (A and C) and BacHum (B and D). Decay of *B. theta* in the treatment and < 5µm control of water column (A). Decay of BacHum in treatment and < 5µm control of water column (B). Decay of *B. theta* (in treatment and < 5µm control of sediments microcosms (C). Decay of BacHum in treatment and < 5µm control of sediments (D). Error bars represent the standard deviation of biological triplicates of each sampling day.

Table 4. ANOVA comparison for 2-log removal (T₉₉) in water column and sediments, and treatment and < 5µm control microcosms

Factors	Fecal Indicator Bacteria		PMA-Treated		Not PMA-Treated	
	<i>E. coli</i> and <i>Enterococcus</i>		<i>B. theta</i> and BacHum		<i>B. theta</i> and BacHum	
	F	P values	F	P values	F	P Values
Overlaying water vs Sediments	8.40	0.02***	1.10	0.03	0.07	0.78
Treatment vs < 5 µm control	0.07	0.80	1.10	0.33	0.11	0.74

***: Significant difference and vs: versus

6.3.3 Decay profiles of fecal indicator bacteria (FIB) in sediments

FIB (*E. coli* and *Enterococcus*) exhibited first order exponential decay in overlaying water and sediments (Figure 4). Two-log removal of FIB in overlaying water and sediments were significantly different (F=8.4, p=0.028) (Table 4). Overall, the decay rates of FIB demonstrated a faster 2-log removal from the overlaying water (≤ 3.5 days) than in sediments where the persistence was extended up 14 d (Table 3). The comparison of FIB decay rates based on the microcosm set up showed that the grazing treatment and limited grazing control (< 5 µm control) were not significantly different (p=0.796) (Table 3). *Enterococcus* demonstrated a faster removal rate in both matrices, overlaying water and sediments, than *E. coli* (Table 3). The 2-log removal (T₉₉) calculations revealed that *Enterococcus* could be detected up to 2 days in overlaying water and 10 days in sediments, while *E. coli* achieved 2-log removal in more than 2 days in water column and up to 14 days in sediments (Table 3).

Figure 3: Variation of culture concentrations of fecal indicator bacteria in treatment (live sewage and sediments supernatant, and autoclaved sediments) and control microcosms overtime. variation of concentrations of *E. coli* in treatment and < 5 μ m control of water column overtime (A); variation of concentrations of *E. coli* in treatment and < 5 μ m control in sediments overtime (B). variation of concentrations of *Enterococcus* in treatment and < 5 μ m control of water column (C) overtime; and variation of concentrations of *Enterococcus* in treatment and < 5 μ m control in sediments overtime (D). Error bars represent the standard deviation of biological triplicates of each *sampling* day.

Figure 4: First order exponential decay ($N=N_0 \cdot e^{-kt}$) profiles for fecal indicator bacteria in the water column and sediment layers of tropical freshwater sediments microcosms under dark and anoxic conditions. First order decay rates profiles of *E. coli* (A and B). First order decay rate profile for *Enterococcus* (C and D). Black colored points (triangles, diamonds, and circular) represent concentrations in the treatment, the blue colored points represent concentrations in limited grazing microcosms (< 5 μ m control). The curve line (black and blue) illustrates the model fitted to the measured data

The measurements of samples collected from no-sewage and nutrient controls microcosms demonstrated that natural occurring *Enterococcus* and *E. coli* in sediments sampled in a tropical reservoir were removed faster in the overlaying water than in sediments (Figure 3). *Enterococcus* was removed in less than a day from the overlaying water (No-sewage and nutrient controls), while *E. coli* took up to 4 d (Figure 3 and 4). Decay rate of FIB in no-sewage and nutrient sediments controls revealed a persistence of FIB where *E. coli* was detected up to 10 d and 14 d for *Enterococcus* (Figure 3).

Table 5. Protozoa Counts in Treatment and Three Controls (Log Protozoa counts/100ml)

Sampling day	Treatment		No-sewage control		Nutrient control (Autoclaved sewage)		< 5 µm control	
	Geomean	SD	Geomean	SD	Geomean	SD	Geomean	SD
1	5.3	0.22	5.4	0.20	5.6	0.14	5.6	0.33
2	5.5	0.26	5.6	0.25	5.7	0.19	5.7	0.15
3	5.7	0.21	5.6	0.23	5.5	0.16	5.6	0.17
5	5.6	0.29	5.7	0.18	5.9	0.15	5.8	0.24
7	5.8	0.17	5.8	0.08	5.5	0.07	5.5	0.09
10	5.6	0.13	5.6	0.16	5.5	0.21	5.4	0.30
14	5.8	0.13	5.8	0.11	5.7	0.26	5.6	0.10
20	5.4	0.26	5.6	0.29	5.3	0.18	5.3	0.27

SD: Standard Deviation, **Geomean:** Geometric mean; **< 5 µm control:** Filtered supernatant and sewage to decrease grazing

6.3.4 The effect of protozoa on the decay of FIB and human-associated *Bacteroidales* markers

Total protozoa (ciliates, flagellates and amoeba) were enumerated using fluorescence microscopy in sediment supernatants after preliminary microscopy tests revealed that protozoa were present mainly in the top layer of sediments in contact with the overlaying water.

Overall the levels of protozoa counts were closely similar through the entire experiment in all treatments and controls suggesting that they were not strongly impacted by the addition of sewage in the system (Table 5). The decay profiles of human markers (*B. theta* and BacHum) did not show any relationship with levels of protozoa counts (Supplemental Figure 1C and 1D). However, *B. theta* and BacHum showed slightly different T_{99} between grazing microcosm (Treatment) and limited grazing control (< 5 μm control) (e.g. BacHum and *B. theta* had a 2-log removal of 6d to 3d) (Table 3). The decay profiles of FIB did not show any relationship with protozoa though the entire experiment. However, different FIB 2-log removal were observed between grazing microcosms (Treatment) and limited grazing control (< 5 μm control) (e.g. 2-log removal of *E. coli* went from 4d to 14d respectively) (Table 3). The microscopy test could not distinguish among groups of protozoa to examine if any of them may have had more impact on the decay rates of FIB and human markers.

Figure 5: First order exponential decay ($N=N_0 \cdot e^{-kt}$) profiles for PMA-treated and no-PMA treated concentrations of *B. theta* in water column and sediments layers of tropical freshwater microcosms under dark and anoxic conditions. First order exponential decay rate of *B. theta* in PMA-treated samples (A, and B). First order exponential decay rate of *B. theta* profile for No-PMA *B. theta* (C and D). Black colored points (circles, triangles, squares and diamonds) represent concentrations in the treatment, the blue colored points represent concentrations in limited grazing microcosms ($< 5 \mu\text{m}$ control). The curve line (black and blue) illustrates the model fitted to the measured data.

Figure 6: First order exponential decay ($N=N_0 \cdot e^{-kt}$) profiles for PMA-treated and no-PMA treated concentrations of BacHum in water column and sediments layers of tropical freshwater microcosms under dark and anoxic conditions. First order exponential decay rate of BacHum in PMA-treated samples (A, and B). First order exponential decay rate of BacHum profile for No-PMA treated (C and D). Black colored points (triangles, squares and diamonds) represent concentrations in the treatment, the blue colored points represent concentrations in limited grazing microcosms ($< 5\mu\text{m}$ control). The curve line (black and blue) illustrates the model fitted to the measured data

6.4 Discussion

6.4.1 Survival and persistence of human-associated *Bacteroidales* markers in sediments

The role of environmental factors such as sunlight, salinity, and temperature on the decay of human-associated *Bacteroidales* markers and fecal indicator bacteria (FIB) in water column have been extensively studied (Schulz and Childers, 2011, Dick et al., 2010a, Bae and Wuertz, 2012, Bae and Wuertz, 2009b, Bae and Wuertz, 2014, Bae and Wuertz, 2015). The current use of human-associated *Bacteroidales* markers for microbial source tracking (MST) and monitoring of human fecal pollution sources has increased the need to understand their decay in sediments (Dick et al., 2010a, Eichmiller et al., 2013, Kim and Wuertz, 2015, Ahmed et al., 2008b, Bernhard and Field, 2000, Kildare et al., 2007). This interest is associated with living conditions of *Bacteroidales* as anaerobic bacteria (Fogarty and Voytek, 2005) and anoxic living environment for bacteria established in sediments (Nealson, 1997). The knowledge about the survival of *Bacteroidales* in sediments could provide vital information to the process of discerning recent from old human fecal pollution. A recent study examined the survival and persistence of host-associated *Bacteroidales* cells and DNA in comparison to FIB in the USA demonstrated that 56 d for *Bacteroidales* cells and 169 d for total DNA were needed to reach a 2-log removal at a temperature of 6° C (Kim and Wuertz, 2015). In addition, the study reported that the levels of *Bacteroidales* cells and DNA decay were six times slower in sediments compared to the overlying water (Kim and Wuertz, 2015). Despite these findings that provided insights of *Bacteroidales* cells and DNA decay under temperature climate, more studies are still needed under high

temperature of tropical environments. This study examined the decay rates of human-associated *Bacteroidales* cells (PMA Treated samples) and DNA (no-PMA treatment) in sediments under tropical conditions.

B. theta and BacHum in PMA-treated samples had similar decay rates in overlaying water and slightly different decay rates in sediments (Table 3). The faster decay rates of viable *Bacteroidales* cells in the water column observed in this study was previously reported in studies in the USA despite differences in studied temperatures (Bae and Wuertz, 2015, Kim and Wuertz, 2015). These studies compared decay rates of viable *Bacteroidales* cells in sediments, seawater and freshwater under sunlight and dark conditions and concluded that viable cells had a faster decay rate than total DNA (viable cells, dead cells and free DNA) (Bae and Wuertz, 2015, Kim and Wuertz, 2015). A recent study carried out in Singapore focused only on the aerobic water column and demonstrated similar decay rates for *B. theta* at 30° C (Liang et al., 2017). This study also did not observe any significant differences between decay rates of viable cells and total DNA ($p=0.576$) suggesting that DNA was also from sewage viable *Bacteroidales*. As the markers were not detected in controls (no-sewage and sewage nutrients), the only source of human *Bacteroidales* in microcosms was sewage.

The 2-log removal of *B. theta* and BacHum total DNA in the dark and at constant high temperature (30°C) was not significantly different between water column (< 6 d) and sediments (< 8 d) (PMA-treated samples: $p=0.334$ and total DNA: $p=0.780$, Table 4). A recent study by Kim and Wuertz (2015) demonstrated that decay rate of *Bacteroidales* cells and DNA (BacHum) in overlaying water in this study were similar to those measured at lower temperatures (6 and 20° C). Two-log removal in overlaying

water was 9 d for cells and 8 d for DNA, while the decay rates in sediments were longer and different from our findings (e.g. T₉₉ in top layer of sediments 0 – 2 cm was 37 d for cells and 50 d for DNA) (Kim and Wuertz, 2015). This difference could be attributed to high ambient temperatures (30° C) used in this study. The recent study by Liang et al. examined the decay rate of human-associated *Bacteroidales* markers (*B. theta* and *M. smithii*) in the water column of seawater and freshwater at constant temperature, in the sunlight and in dark conditions in Singapore suggested that marker was observed to decay at a faster rate (T₉₅ = 4 days) in our study versus Liang et al (T₉₅ = 5 days) (Liang et al., 2017). The 2-log removal of host-specific *Bacteroides* -*Prevotella* ssp. genetic marker (HF183/Bac242R) in freshwater (river water) microcosms incubated at 28°C in Japan was reported to be 6 days similar to the 2-log removal reported in the water column in this study (Kobayashi et al., 2013, Seurinck et al., 2005). The similar removal of human markers from the water column by both studies and removal from sediments presented in this study could be associated with the temperature range (e.g. 28° C and 30° C).

The human markers (*B. theta* and BacHum) did not show any variation in the presence of grazing in treatment microcosms versus reduced protozoa grazing in the filtered control (< 5µm control). Kobayashi et al. recently demonstrated the effect of predators on the persistence of host-specific *Bacteroidales* *Prevotella*-group (HF183/Bac242R) in water using a sequencing approach (Kobayashi et al., 2013). They examined 18S rRNA sequences and identified three clones of protozoa (flagellate: *Spumella*-like (99.5%), ciliates: *Colpidium camylum* (98.3%) and *Glaucoma scintillans* (98.9%)) that could affect the decay rates of *Bacteroidales* genetic markers in the environment (Kobayashi et al., 2013). In our study BacHum (PMA and non-PMA treated) exhibited

slightly lower 2-log removal times in the overlaying water in microcosms with reduced predation (< 5µm control) (Table 3). Dicks and colleagues (2010), although different sediments treatment conditions in our study were applied, demonstrated that the exposure to sunlight, sediments and reduced predation resulted in increased decay rates of *Bacteroidales* markers compared to *E. coli* (Dick et al., 2010b). These studies suggest protozoan grazing influences the fate of these markers, however our study suggests that this is not a significant factor in anoxic sediments, which is not surprising since many protozoa (especially larger cells) require oxygen for activity.

6.4.2 Survival and persistence of fecal indicator bacteria in sediments

Enterococcus decayed faster than *E. coli* in sediments overlaying water (Table 3). The lower time of 2-log removal of *Enterococcus* versus *E. coli* in the water column demonstrated in this study supported the use of *Enterococcus* as a fecal indicator of recent contamination. However, the potential growth of *Enterococcus* in soil or resuspension from sediments previously confirmed by other studies, was also observed in this study (Ferguson et al., 2005, Perkins et al., 2016, Fujioka et al., 1998). The no-sewage and < 5µm controls demonstrated three phases (growth, lag-time and decrease) of *Enterococcus* and *E. coli* found in freshwater sediments used in this study. These naturally occurring FIB were removed in more than 10 and 14 respectively for *E. coli* and *Enterococcus* and their concentrations varied below log 4.0 MPN/100ml for *E. coli* and log 2.5 MPN/100ml for *Enterococcus*. These findings suggested that a base-line should be established in order to evaluate significance of fecal pollution under tropical environment using FIB (Figure 3).

Fecal indicator bacteria (FIB) (*Enterococcus* and *E. coli*) demonstrated extended periods of survival in sediments. This prolonged survival of FIB in sediments was also confirmed by the Anderson et al. study of persistence and differential survival of FIB in subtropical waters and sediments (Anderson et al., 2005). In addition, the study showed that sewage contaminated sediments increased the survival of FIB compared to other matrices such as water column (Anderson et al., 2005). The potential of increase in FIB in water column by resuspension of sediments was recently confirmed by Perkins et al. whose study demonstrated a variation of FIB decay rates associated with suspended particulate matter in estuaries in the United Kingdom (Perkins et al., 2016).

Previous studies demonstrated that the source of fecal contaminations influenced the survival in sediments. The prolonged decay rate of FIB (*Enterococcus* and *E. coli*) in sediments in our study might have been associated with the use of sewage as an inoculum (Table 3, Figure 3B and D). Sewage-derived *Bacteroidales* are acclimated to survival outside of the gut, while feces-associated *Bacteroidales* may be more susceptible to environmental stresses. Sewage was previously compared with various sources such as soil and feces from different domesticated and ruminant animals. Findings of these studies suggested that an origin in sewage was among major contributing factors to the survival of FIB in sediments (Anderson et al., 2005, Korajkic et al., 2013a).

The presence or absence of maximum or reduced protozoa grazing demonstrated effect on the decay period of FIB. *E. coli* was reached a 2-log removal in 14 days in the microcosm with reduced grazing (< 5 µm control), while *Enterococcus* had 2-log removal of 10 days in the maxim protozoa grazing microcosm (treatment). Korajkic et

al. study recently demonstrated that the inactivation of protozoa and bacterial competitors by the use of a disinfectant increased the decline of *E. coli* in freshwater sediments and water column in the subtropical environment of Florida, USA (Korajkic et al., 2013b). Thus, confirming the role of protozoa in regulating the persistence of *E. coli* in sediments and water column. Despite the change in 2-log removal of FIB based on the presence/absence of maximum or reduced grazing, the analysis on protozoa in this study did not demonstrate a significant variation throughout the microcosm incubation period. Hence, the role of protozoan grazing in anoxic environments is not well known. More studies focusing on specific groups of protozoa or using other methods of quantification than microscopy or sequencing could provide more supporting evidence of the role of protozoa.

6.4.3 Decay models

Understanding the appropriate models for extrapolating FIB and human markers rate of decay in tropical freshwater would greatly contribute to watershed managements. In addition, the understanding of the fate of sewage and associated with health risks exposures in tropical freshwater.

The non-linear model fitting of the data generated in this study to a first order exponential decay explained removal trends to all FIB and human markers measured in this study and in all treatments (Table 3). The predictive squared correlation coefficient (R^2) of all first decay exponential models generated in this study supported that the predictive data highly matched our measured data ($R^2 > 0.93$ and $p < 0.001$) (Figure 4 to Figure 6). For FIB (*E. coli*), the model fitting to the decay data in sediments matched

the first degree exponential decay model after the removal of the first three days with constant concentration and only used the portion of the data where the first order decay was apparent. Previous studies with irregular fitting to the first exponential decay suggested that the bacterial inactivation was characterized by two or more phases such as a slow removal or early growth followed by exponential decay (Whitman et al., 2004). These findings led to a conclusion of using an improved decay models that would account for all the multiple environmental stressors (physicochemical or biological factors) (Whitman et al., 2004, Crane and Moore, 1986, Bae and Wuertz, 2009b). However, as the lag-time for our study was characterized by a trend of constant concentrations (Supplemental Figure 1B), only the section of the curve representing the decay was used for model fitting.

6.5. Conclusions

This study addressed whether sediments could contribute to the survival of human-associated *Bacteroidales* genetic markers (B. theta and BacHum), and fecal indicator bacteria (FIB) under tropical conditions. We also evaluated whether the presence grazing or reduced grazing could affect the survival of human-associated *Bacteroidales* markers and in freshwater sediments in tropical urban environment of Singapore. Our major findings are as follows:

- Human-associated *Bacteroidales* markers, B. theta and BacHum, demonstrated comparable decay rates in water column and sediments.
- Viable cells were identified to have a faster decay rate (< 6 d for T_{99}) in the water column and sediments supporting use of human-associated *Bacteroidales* markers

(*B. theta* and BacHum) for detecting and identifying recent human fecal contaminations.

- Fecal indicator bacteria (*Enterococcus* and *E. coli*) were identified to have a prolonged (*E. coli*) or short (*Enterococcus*) survival in sediments, while their survival in water column was faster for both *Enterococcus* and *E. coli* (< 2 d for T₉₉ removal). These findings confirmed the potential use of *Enterococcus* as indicators of sewage contamination, while sediments could be considered as a source of non-point of fecal contamination by resuspension of *E. coli* cells from sediments in the water column.
- Human-associated *Bacteroidales* markers and FIB were found to have a short survival period in sediments overlaying water, but survived longer in sediments. This study demonstrated BacHum and *B. theta* markers could be used as proxies for recent sewage contamination, supporting their adoption in the tropics.

References: Chapter 6

- AHMED, W., GOONETILLEKE, A., POWELL, D. & GARDNER, T. 2009. Evaluation of multiple sewage-associated *Bacteroides* PCR markers for sewage pollution tracking. *Water Research*, 43, 4872-7.
- AHMED, W., POWELL, D., GOONETILLEKE, A. & 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, 579-86.
- AHMED, W., STEWART, J., POWELL, D. & GARDNER, T. 2008b. Evaluation of *Bacteroides* markers for the detection of human faecal pollution. *Letters in Applied Microbiology*, 46, 237-42.
- AHMED, W., YUSUF, R., HASAN, I., GOONETILLEKE, A. & 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, 838-845.
- ANDERSON, M. L., WHITLOCK, J. E. & HARWOOD, V. J. 2005. Persistence and differential survival of fecal indicator bacteria in subtropical waters and sediments. *Applied and Environmental Microbiology*, 71, 8.
- BAE, S. & WUERTZ, S. 2009a. Discrimination of viable and dead fecal Bacteroidales bacteria by quantitative PCR with propidium monoazide. *Applied and Environmental Microbiology*, 75, 2940-4.
- BAE, S. & WUERTZ, S. 2009b. Rapid decay of host-specific fecal Bacteroidales cells in seawater as measured by quantitative PCR with propidium monoazide. *Water Research*, 43, 4850-9.
- BAE, S. & WUERTZ, S. 2012. Survival of host-associated bacteroidales cells and their relationship with *Enterococcus* spp., *Campylobacter jejuni*, *Salmonella enterica* serovar Typhimurium, and adenovirus in freshwater microcosms as measured by propidium monoazide-quantitative PCR. *Applied and Environmental Microbiology*, 78, 922-32.
- BAE, S. & WUERTZ, S. 2014. Differential decay kinetics of host-associated Bacteroidales markers in intact cells under continuous-flow freshwater and seawater conditions as measured by PMA-qPCR. In: ENGINEERING, D. O. C. A. E. (ed.). UC Davis.
- BAE, S. & WUERTZ, S. 2015. Decay of host-associated Bacteroidales cells and DNA in continuous-flow freshwater and seawater microcosms of identical

- experimental design and temperature as measured by PMA-qPCR and qPCR. *Water Research* 70, 205-13.
- BERNHARD, A. E. & FIELD, K. G. 2000. A PCR Assay To Discriminate Human and Ruminant Feces on the Basis of Host Differences in Bacteroides-Prevotella Genes Encoding 16S rRNA. *Applied and Environmental Microbiology*, 66, 4571-4574.
- BURD, E. M. 2010. Validation of laboratory-developed molecular assays for infectious diseases. *Clinical Microbiology Reviews*, 23, 6.
- CAO, Y., RAITH, M. R. & GRIFFITH, J. F. 2015. Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. *Water Research* 70, 337-49.
- CHICK, H. 1908. An Investigation of the Laws of Disinfection. *The Journal of Hygiene*, 8, 92-158.
- CODEX 2010. Guidelines on performance criteria and validation of methods for detection, identification and quantification of specific DNA sequences and specific proteins in foods. Rome.
- CRANE, S. R. & MOORE, J. A. 1986. Modeling enteric bacterial die-off: A review. *Water, Air, and Soil Pollution*, 27, 411-439.
- DAVIES, C. M., LONG, J. A., DONALD, M. & ASHBOLT, N. J. 1995. Survival of fecal microorganisms in marine and freshwater sediments. *Applied and Environmental Microbiology*, 61, 1888-96.
- DICK, L., STELZER, E., BERTKE, E., FONG, D. & STOECKEL, D. 2010a. Relative decay of Bacteroidales microbial source tracking markers and cultivated Escherichia coli in freshwater microcosms. *Applied and Environmental Microbiology*, 76, 3255-62.
- DICK, L. K., STELZER, E. A., BERTKE, E. E., FONG, D. L. & STOECKEL, D. M. 2010b. Relative decay of Bacteroidales microbial source tracking markers and cultivated Escherichia coli in freshwater microcosms. *Applied and Environmental Microbiology*, 76, 3255-62.
- EICHMILLER, J. J., HICKS, R. E. & SADOWSKY, M. J. 2013. Distribution of genetic markers of fecal pollution on a freshwater sandy shoreline in proximity to wastewater effluent. *Environmental Science & Technology*, 47, 3395-3402.
- ERVIN, J. S., VAN DE WERFHORST, L. C., MURRAY, J. L. & HOLDEN, P. A. 2014. Microbial source tracking in a coastal California watershed reveals canines as controllable sources of fecal contamination. *Environmental Science & Technology*, 48, 9043-52.

- FENCHEL, T. & FINLAY, B.J. (1990) Anaerobic free-living protozoa: growth efficiencies and the structure of anaerobic communities. *FEMS Microbial Ecology*, 74: 269–276.
- FENCHEL, T., KRISTENSEN, L.D. & RASMUSSEN, L. (1990) Water column anoxia: vertical zonation of planktonic protozoa. *Marine Ecology Progress Series*, 62: 1–10.
- FENCHEL T., & Finlay BJ (1995) *Ecology and Evolution in Anoxic Worlds*. Oxford University Press, Oxford, UK.
- FENG, F., GOTO, D. & YAN, T. 2010. Effects of autochthonous microbial community on the die-off of fecal indicators in tropical beach sand. *FEMS Microbiology Ecology*, 74, 214–25.
- FERGUSON, D. M., MOORE, D. F., GETRICH, M. A. & ZHOWANDAI, M. H. 2005. Enumeration and speciation of enterococci found in marine and intertidal sediments and coastal water in southern California. *Applied Microbiology*, 99, 598–608.
- FOGARTY, L. R. & VOYTEK, M. A. 2005. Comparison of bacteroides-prevotella 16S rRNA genetic markers for fecal samples from different animal species. *Applied and Environmental Microbiology*, 71, 5999–6007.
- FUJIOKA, R., SIAN-DENTON, C., BORJA, M., CASTRO, J. & MORPHEW, K. 1998. Soil: the environmental source of Escherichia coli and Enterococci in Guam's streams. *Applied Microbiology*.
- GAWLER, A. H., BEECHER, J. E., BRANDÃO, J., CARROLL, N. M., FALCÃO, L., GOURMELON, M., MASTERSON, B., NUNES, B., PORTER, J. & RINCÉ, A. 2007. Validation of host-specific Bacteroidales 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union. *Water Research*, 41, 3780.
- GREEN, H., SHANKS, O., SIVAGANESAN, M., HAUGLAND, R. & FIELD, K. 2011. Differential decay of human faecal Bacteroides in marine and freshwater. *Environmental Microbiology*, 13, 3235–49.
- GREEN, H. C., DICK, L. K., GILPIN, B., SAMADPOUR, M. & FIELD, K. G. 2012. Genetic Markers for Rapid PCR-Based Identification of Gull, Canada Goose, Duck, and Chicken Fecal Contamination in Water. *Applied and Environmental Microbiology*, 78, 503–510.

- HALLIDAY, E. & GAST, R. J. 2011. Bacteria in beach sands: an emerging challenge in protecting coastal water quality and bather health. *Environmental Science & Technology*, 45, 370-379.
- HARWOOD, V. J., BROWNELL, M., WANG, S., LEPO, J., ELLENDER, R. D., AJIDAHUN, A., HELLEIN, K. N., KENNEDY, E., YE, X. & FLOOD, C. 2009. Validation and field testing of library-independent microbial source tracking methods in the Gulf of Mexico. *Water Research*, 43, 4812-9.
- HOFFMAN, R. L. & ATLAS, R. M. 1987. Measurement of the effects of cadmium stress on protozoan grazing of bacteria (bacterivory) in activated sludge by fluorescence microscopy. *Applied and environmental microbiology*, 53, 2440-2444.
- HUGGETT, J. F., FOY, C. A., BENES, V., EMSLIE, K., GARSON, J. A., HAYNES, R., HELLEMANS, J., KUBISTA, M., MUELLER, R. D., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J., WITTEWER, C. T. & BUSTIN, S. A. 2013. The digital MIQE guidelines: Minimum Information for Publication of Quantitative Digital PCR Experiments. *Clinical Chemistry*, 59, 892-902.
- JENKINS, M. W., TIWARI, S., LORENTE, M., GICHABA, C. M. & WUERTZ, S. 2009. Identifying human and livestock sources of fecal contamination in Kenya with host-specific Bacteroidales assays. *Water Research*, 43, 4956-66.
- KILDARE, B., LEUTENEGGER, C., MCSWAIN, B., BAMBIC, D., RAJAL, V. & WUERTZ, S. 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal Bacteroidales: A Bayesian approach. *Water Research*, 41, 15.
- KIM, M., GUTIERREZ-CACCIABUE, D., SCHRIEWER, A., RAJAL, V. B. & WUERTZ, S. 2014. Evaluation of detachment methods for the enumeration of *Bacteroides fragilis* in sediments via propidium monoazide quantitative PCR, in comparison with *Enterococcus faecalis* and *Escherichia coli*. *Applied Microbiology* 117, 1513-22.
- KIM, M. & WUERTZ, S. 2015. Survival and persistence of host-associated Bacteroidales cells and DNA in comparison with *Escherichia coli* and *Enterococcus* in freshwater sediments as quantified by PMA-qPCR and qPCR. *Water Research*, 87, 182-192.
- KOBAYASHI, A., SANO, D. & OKABE, S. 2013. Effects of temperature and predator on the persistence of host-specific *Bacteroides-Prevotella* genetic markers in water. *Water Science and Technology*, 67, 838-45.
- KORAJKIC, A., MCMINN, B. R., HARWOOD, V. J., SHANKS, O. C., FOUT, G. S. & ASHBOLT, N. J. 2013a. Differential Decay of Enterococci and *Escherichia*

coli Originating from Two Fecal Pollution Sources. *Applied and Environmental Microbiology*, 79, 2488-2492.

KORAJKIC, A., WANJUGI, P. & HARWOOD, V. J. 2013b. Indigenous Microbiota and Habitat Influence *Escherichia coli* Survival More than Sunlight in Simulated Aquatic Environments. *Applied and Environmental Microbiology*, 79, 5329-5337.

LAYTON, B., CAO, Y., EBENTIER, D., HANLEY, K., BALLESTÉ, E., BRANDÃO, J., BYAPPANAHALLI, M., CONVERSE, R., FARNLEITNER, A., GENTRY SHIELDS, J., GIDLEY, M., GOURMELON, M., LEE, C., LEE, J., LOZACH, S., MADI, T., MEIJER, W., NOBLE, R., PEED, L., REISCHER, G., RODRIGUES, R., ROSE, J., SCHRIEWER, A., SINIGALLIANO, C., SRINIVASAN, S., STEWART, J., VAN DE WERFHORST, L. C., WANG, D., WHITMAN, R., WUERTZ, S., JAY, J., HOLDEN, P., BOEHM, A., SHANKS, O. & GRIFFITH, J. 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. *Water Research*, 47, 6897-6908.

LIANG, L., GOH, S. G. & GIN, K. Y. H. 2017. Decay kinetics of microbial source tracking (MST) markers and human adenovirus under the effects of sunlight and salinity. *Science of the Total Environment*, 574, 165-175.

NEALSON, K. H. 1997. Sediment bacteria: who's there, what are they doing, and what's new? *Annu Rev Earth Planet Sci*, 25, 403-34.

NEAVE, M., LUTER, H., PADOVAN, A., TOWNSEND, S., SCHOBEN, X. & GIBB, K. 2014. Multiple approaches to microbial source tracking in tropical northern Australia. *Microbiology Open*, 3, 860-874.

ODAGIRI, M., SCHRIEWER, A., HANLEY, K., WUERTZ, S., MISRA, P. R., PANIGRAHI, P. & JENKINS, M. W. 2015. Validation of Bacteroidales quantitative PCR assays targeting human and animal fecal contamination in the public and domestic domains in India. *Science of the Total Environment*, 502, 462-470.

OKABE, S., OKAYAMA, N., SAVICHTCHEVA, O. & ITO, T. 2007. Quantification of host-specific Bacteroides-Prevotella 16S rRNA genetic markers for assessment of fecal pollution in freshwater. *Applied Microbiology and Biotechnology*, 74, 890-901.

OKABE, S. & SHIMAZU, Y. 2007. Persistence of host-specific Bacteroides-Prevotella 16S rRNA genetic markers in environmental waters: effects of temperature and salinity. *Applied Microbiology and Biotechnology*, 76, 935-44.

PERKINS, T. L., PERROW, K., RAJKO-NENOW, P., JAGO, C. F., JONES, D. L., MALHAM, S. K. & MCDONALD, J. E. 2016. Decay rates of faecal indicator

bacteria from sewage and ovine faeces in brackish and freshwater microcosms with contrasting suspended particulate matter concentrations. *Science of the Total Environment*, 572, 1645-1652.

PINHEIRO, L. B., COLEMAN, V. A., HINDSON, C. M., HERRMANN, J., HINDSON, B. J., BHAT, S. & EMSLIE, K. R. 2012. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Analytical Chemistry*, 84, 1003-11.

PRIESTER, J. H., HORST, A. M., VAN DE WERFHORST, L. C., SALETA, J. L., MERTES, L. A. & HOLDEN, P. A. 2007. Enhanced visualization of microbial biofilms by staining and environmental scanning electron microscopy. *Micobiology Methods*, 68, 577-87.

REISCHER, G. H., EBDON, J. E., BAUER, J. M., SCHUSTER, N., AHMED, W., ÅSTRÖM, J., BLANCH, A. R., BLÖSCHL, G., BYAMUKAMA, D., COAKLEY, T., FERGUSON, C., GOSHU, G., KO, G., DE RODA HUSMAN, A. M., MUSHI, D., POMA, R., PRADHAN, B., RAJAL, V., SCHADE, M. A., SOMMER, R., TAYLOR, H., TOTH, E. M., VRAJMASU, V., WUERTZ, S., MACH, R. L. & FARNLEITNER, A. H. 2013. Performance Characteristics of qPCR Assays Targeting Human- and Ruminant-Associated Bacteroidetes for Microbial Source Tracking across Sixteen Countries on Six Continents. *Environmental Science and Technology*, 47, 8548-8556.

Sacca, A , Carles M. Borrego , C.M. , Renda, R. , Triado-Margarit, X. , Bruni, V. and Guglielmo, L. (2009). Predation impact of ciliated and flagellated protozoa during a summer bloom of brown sulfur bacteria in a meromictic coastal lake. *FEMS Microbial Ecology*, 70, pp. 42-53

SCHRIEWER, A., GOODWIN, K. D., SINIGALLIANO, C. D., COX, A. M., WANLESS, D., BARTKOWIAK, J., EBENTIER, D. L., HANLEY, K. T., ERVIN, J., DEERING, L. A., SHANKS, O. C., PEED, L. A., MEIJER, W. G., GRIFFITH, J. F., SANTODOMINGO, J., JAY, J. A., HOLDEN, P. A. & WUERTZ, S. 2013. Performance evaluation of canine-associated Bacteroidales assays in a multi-laboratory comparison study. *Water Research*, 47, 6909-20.

SCHULZ, C. J. & CHILDERS, G. W. 2011. Fecal bacteroidales diversity and decay in response to variations in temperature and salinity. *Applied and Environmental Microbiology*, 77, 2563-2572.

SEURINCK, S., DEFOIRDT, T., VERSTRAETE, W. & SICILIANO, S. 2005. Detection and quantification of the human-specific HF183 Bacteroides 16S rRNA genetic marker with real-time PCR for assessment of human faecal pollution in freshwater. *Environmental Microbiology*, 7, 249-59.

- SHUVAL, H. 2003. Estimating the global burden of thalassogenic diseases: human infectious diseases caused by wastewater pollution of the marine environment. *Journal of Water Health*, 1, 53-64.
- STEWART, J., BOEHM, A., DUBINSKY, E., FONG, T.-T., GOODWIN, K., GRIFFITH, J., NOBLE, R., SHANKS, O., VIJAYAVEL, K. & WEISBERG, S. 2013. Recommendations following a multi-laboratory comparison of microbial source tracking methods. *Water Research*, 47, 6829-38.
- TOLEDO-HERNANDEZ, C., RYU, H., GONZALEZ-NIEVES, J., HUERTAS, E., TORANZOS, G. A. & SANTO DOMINGO, J. W. 2013. Tracking the Primary Sources of Fecal Pollution in a Tropical Watershed in a One-Year Study. *Applied and Environmental Microbiology*, 79, 1689-1696.
- USGS. 2016. *Volatile Organic Compounds in the Nation's Ground Water and Drinking-Water Supply Wells: Supporting Information: Glossary* [Online]. US Geological Survey. Available: http://water.usgs.gov/nawqa/vocs/national_assessment/report/glossary.html [Accessed November 15, 2016 2016].
- WALTERS, S. & FIELD, K. 2009. Survival and persistence of human and ruminant-specific faecal Bacteroidales in freshwater microcosms. *Environmental Microbiology*, 11, 1410-1421.
- WANJUGI, P., FOX, G. A. & HARWOOD, V. J. 2016. The Interplay Between Predation, Competition, and Nutrient Levels Influences the Survival of Escherichia coli in Aquatic Environments. *Microbial Ecology*, 72, 526-37.
- WHITMAN, R. L., NEVERS, M. B., KORINEK, G. C. & BYAPPANAHALLI, M. N. 2004. Solar and temporal effects on Escherichia coli concentration at a Lake Michigan swimming beach. *Applied and Environmental Microbiology* 70, 4276-85.
- YAMPARA-IQUISE, H., ZHENG, G., JONES, J. E. & CARSON, C. A. 2008. Use of a Bacteroides thetaiotaomicron-specific alpha-1-6, mannanase quantitative PCR to detect human faecal pollution in water. *Applied Microbiology*, 105, 1686-93.

Chapter 7.

Summary and Future Research Considerations

Microbial water quality assessment has improved and reduced human health risks associated with waterborne diseases particularly in high-income countries. However, there is still a need to improve microbial water quality assessment as 3.4 million people are dying every year from water-related diseases (Pandey et al., 2014). Children are the most of the affected population with an estimate of 4,000 children dying each day as a result of contaminated water (Pandey et al., 2014, WHO, 2014). Globally, an estimated 1.8 billion people use fecal contaminated drinking water and this burden is shared in rural and urban areas, and most affected water sources are in low-income countries in Africa and Southeast Asia (Bain et al., 2014a, Bain et al., 2014b).

Major challenges for water quality assessment in tropical urban environments include natural occurrence and potential growth of fecal indicator bacteria (FIB), identification and discrimination of non-point fecal pollution sources, and lack of knowledge on survival and persistence of host-associated genetic markers in the environment. This research work identified potential technical and strategic solutions to these problems and provided an assessment of currently available tools to identify efficient applications. In this chapter I will discuss major findings of my research work and will suggest future research directions motivated by my work.

7.1 Implications and Future research considerations

7.1.1 Microbial community Dynamics and water quality

Environmental natural occurrence of *E. coli* in the tropics, and high baseline levels, have reduced the credibility for its use in water quality monitoring (Rochelle-Newall et al., 2015, Rochelle-Newall et al., 2016). To better understand the suitability of *E. coli* as a marker for water quality assessment in the tropics, we asked whether sites with high concentrations of measured *E. coli* in Singapore harbored microbial communities enriched in sewage-associated taxa (SAT) or pathogen-like bacteria with the expectation that significant correlations would be observed if *E. coli* levels were predictive for water quality impairment. Indeed, significant correlation between proportions of sequences from SAT and pathogen-like bacteria and the concentrations of *E. coli* suggested measured *E. coli* levels may be useful for predicting exposure to pathogens. Further work is needed to establish relevant environmental baselines and thresholds for categorizing sites by impairment risk. Sites with *E. coli* levels above the environmental background could be selected for additional tests to identify if they may be a source of human health risks.

We also examined HF183 sybr (HF183/Bac242/Bac282) approved by USEPA for human fecal source tracking. In contrast to our hypothesis we found little correlation with PLS, although some correlation to SAT. This suggested that HF183 Sybr assay might not be as good as *E. coli* at predicting impaired waters and indicated that further work is needed to validate *Bacteroidales* markers for application in Singapore.

In addition to determining proportions of sewage (sewage-associated taxa) and pathogen-like bacteria (pathogen-like sequences), next generation sequencing (NGS) information allows identification of specific taxa in samples that may pose public health risk. This survey capability can be used to identify and prioritize targets for more quantitative evaluation using culture-based or qPCR-based methods. This information is vital to the effective planning of water quality monitoring as it may orient the user about different pathogenic bacteria that could be prioritized for water quality monitoring. Currently, water quality monitoring of pathogenic strains focuses on the identification of a single strain of a known origin such as human feces (Harwood et al., 2014). However, there are evidence suggesting the increase in naturally occurring pathogens such as *legionella* (Whiley et al., 2014, Ashbolt, 2015). NGS is also being applied in the assessment of microbial water quality at different scales to provide vital information on unknown strains of bacteria or pathogens for assessment of relative abundance or diversity of less studied strains (Tan et al., 2015).

All methods for water quality assessment are subject to biases that may influence the reliability and usefulness of the results. For example cultivation-based methods do not recover viable-but not culturable cells, and identification of bacterial colonies that grow in response to specific culture-based assays reveal that certain populations of non-target organisms that may provide false-positives (Noble et al., 2010, Scott et al., 2002). Similarly, cultivation-independent methods based on a PCR step, such as NGS, may be subject to biases in amplification efficiency and detection of DNA from non-viable cells. Thus, a combination of approaches to study water quality may provide better information than single methods alone. Our study suggests the dual purpose usefulness of NGS to complement measurement by FIB to 1) provide information regarding the

proportion of pathogen-like sequences (PLS) or sewage associated taxa (SAT) which may be useful as an index for risk of water quality impairment while 2) providing information regarding the type of strains to measure in targeted waterways and water bodies.

Consideration for future research:

- Future research projects in water quality assessment of surface waters with interests in pathogens should include NGS as a preliminary step in order to ensure that naturally occurring pathogens are included in monitoring. In addition, NGS profiling could help decided what type of pathogens from a long list of known waterborne pathogens should be the focus of water quality monitoring in targeted waters. My research did not address the question of how pathogen-like bacteria sequences (PLS) could correspond to any health risk. Future epidemiological studies should include NGS where illness occurrences are correlated to FIB to assess if PLS could provide additional power as an indicator.
- Examining the variation of sewage-associated taxa (SAT) and pathogen-like sequences (PLS) in relation to measured water quality parameters could provide insights of what indicators (e.g. *Enterococcus* or *Bacteroidales* markers such as HF183) could be appropriate for sewage and pathogen monitoring. My findings demonstrated poor correlation between HF183 with sewage (Sewage associated Taxa) and PLS. Thus validating *Bacteroidales* markers for human fecal sources tracking and examining their survival under urban tropical environment of Singapore are recommended follow up studies.

- Current research is improving NGS to allow real-time sequencing such as the use of nanopore devices which would allow the real-time assessment of Pathogen-like sequencing (PLS) or sewage associated taxa (SAT). This novel progress will also lead to more efficiency in evaluating the relative abundance of pathogenic bacteria strains and indicator bacteria.
- For my research, I focused only on bacteria pathogens. However, the advancement of NGS allows also the quantification of viruses. Future research should consider assessing how the relative abundance of viruses compare to different indicators of viruses in water and other water quality indicators as it would provide valuable insights on virus monitoring.

7.1.2 *Bacteroidales* markers for microbial source tracking in Singapore and Southeast Asia

This study completed our first study described in Chapter 3. *Bacteroidales* HF183 Sybr (HF183/Bac242) was validated for MST in the USA. My study in Kranji catchment in Singapore demonstrated that HF183 (Bac242) was ubiquitous (Nshimiyimana et al., 2014). However, I observed some unexpected high concentrations of HF183 in reservoir samples and HF183 did not correlate with PLS suggesting non-specificity. Recent studies showed that different *Bacteroidales* assays work better for different environments (Kildare et al., 2007, Odagiri et al., 2015, Layton et al., 2013, Neave et al., 2014, Ahmed et al., 2010, Harwood et al., 2009). I evaluated the performance of host-associated *Bacteroidales* markers including HF183 sybr for microbial source tracking in Singapore. I asked the question whether *Bacteroidales* markers could be used to discriminate human from animal fecal sources in Singapore. I validated human-

associated *Bacteroidales* genetic markers, B. theta and BacHum, for Singapore with application in other Southeast Asia countries as they are located in the same geography and latitude (Suzuki and Worobey, 2014).

Validated human markers (B. theta and BacHum) could be applied in routine water quality monitoring in Singapore for the detection of sewage which may be discharged into the environment due to sewer leaks or poorly operated septic tanks (Nshimyimana et al., 2014). These validated markers will be used during the mitigation procedures as watershed managers will be able to focus on accurately identified sites of interests at watershed level. It should be noted that rabbits were identify as the major cross-reacting non-target in Singapore for the B. theta assay. Since rabbits are popular pets in Singapore, it may be advisable to combine use of B. theta with BacHum to remove false positives associated with rabbits in residential areas that may have sizable pet populations.

Human fecal pollution is prevalent in developing countries of Southeast Asia where available resources are not enough to address the existing sewage problems of booming cities and increasing population (Widmer et al., 2013, Jones, 2013). There is a need of adequate tools to identify and track fecal sources in order to mitigate impaired areas, thus preventing associated human health risks.

As my study validated markers for Southeast Asia, developing countries in the region would need to use a cost-effective approach in applying validated MST markers. This cost-effective approach would include a preliminary survey of urban waters using culture-based techniques such as *E. coli* or *Enterococcus* to identify sites with potential high load of fecal contaminations. Sites with high loads of FIB will then be sampled

and tested using MST validated human markers to identify sites with high load of human fecal pollution (e.g. sewage or septic tank leaks). Sewer leaks repair and other appropriate mitigation procedures will then follow the survey to prevent potential human health risk.

Consideration for future research:

- MST projects are still needed under tropical environments as most of the areas with prevalent human health risks associated with fecal pollutions are found in tropical areas. In addition, tropical areas are characterized by a variety of dietary preferences which may lead to gut microbial variations between regions in close proximity. These efforts could improve MST in the region and lead to efficient mitigation projects that will impact the prevention of human health risks.
- The development of new qPCR machines and master mixes has extended the variety of tools available for MST. However, changing amplification tools or reagents originally used for the design of MST assays could introduce errors and lead to the detection of false positives. There is a need to provide a standardized approach to calibrate or to adjust the equipment or reagents supplier without changing the overall performance of the genetic markers. This could be achieved in a multi-laboratory study in the region to ensure that markers are being used with accuracy, laboratory sample preparation and processing are standardized, and that all results are reproducible and comparable to be a baseline of sharing best practices.

7.1. 3 Real-time PCR (qPCR) and Digital droplet PCR (ddPCR): Which tool is better for microbial source tracking

The findings of my microbial source tracking study discussed in chapter 4 demonstrated that cross reactions concentrations were lower, except in the rabbits fecal samples where human-associated *Bacteroidales* markers concentrations were comparable to those found in human stool and sewage.

Digital droplet PCR (ddPCR), which quantifies molecular material at the scale of droplet, has recently gained popularity in the medical field and it is reported to have high sensitivity, specificity, inhibition control and quantification without the use of a standard curve (Hindson et al., 2011, Pinheiro et al., 2012, McDermott, 2013, Huggett et al., 2013, Hindson et al., 2013, Pillet et al., 2016). These current findings increased my interest to examine whether ddPCR could perform better in quantifying targets in environmental samples.

I demonstrated that overall the performance of ddPCR and qPCR leads to comparable results (e.g. quantification of composite and environmental samples) with improvement of sensitivity for BacHum. However, the analysis of the standard curve (BacHum) was found to be associated with a systematic error likely related to pipetting leading to significantly higher observed concentrations than ddPCR despite passing all quality control for qPCR. E.g. BacHum qPCR concentrations were then 14-fold higher than those generated by ddPCR while *B. theta* concentrations by qPCR and by ddPCR were not that different as their difference was only 3-fold. The similarity of *B. theta* results between both platforms (ddPCR and qPCR) demonstrated how some differences in markers could be missed as qPCR quantification are based on a standard curve and

quantification in 20µl. DDPCR also demonstrated a constant reproducibility even among lower concentrations, confirming its usefulness for quantifying environmental samples. However, poor capacity in quantifying high concentrations by ddPCR may be a limitation. QPCR could be used if the target of interest is in high concentration in the environment (e.g. ongoing sewage pollution leading to high concentrations in targeted water body) or a sample could be diluted and brought to a concentration quantifiable by ddPCR.

Overall the decision to replace qPCR by ddPCR becomes a question of cost effectiveness and intended use as the overall performance of ddPCR did not differ greatly from that of qPCR. Currently, the cost of operating ddPCR is still at least 3X higher than that of qPCR and the time involved is different as well. For example the cost of running one sample using a probe would cost up to 3 times on ddPCR the cost the same run would be priced for on a qPCR platform. In addition, using a BioRad ddPCR system as an example, it may take up to 1.5 hour to prepare and run a plate by qPCR, while the time doubles when using the ddPCR as the processes involves lengthy stages of droplet generation before quantification and reading after qPCR. As the cost and time for sample preparation and runs differ for both platforms, (ddPCR versus qPCR), the selection of what system to buy or use could be influenced by what type of targets to quantify and what are their concentrations in samples to be run. Here, if one would be interested in quantify pathogens, which are generally known to be of a lower concentration in environmental sample, then ddPCR could be the best choice. Studies have shown that qPCR can be as good as ddPCR if the standards are quantified accurately. For qPCR standard curve improvement, one can use ddPCR to measure

standard concentrations, and then run qPCR as a less expensive way to quantify environmental samples off a calibration curve that's been validated by ddPCR.

Considerations for future research:

- The use of ddPCR for quantifying targets in environmental samples is still limited. However, the absence of a standard curve suggests it may be more reproducible across laboratories. Future research work should consider applying ddPCR to large scale cross-laboratory water quality monitoring projects with interest in quantifying pathogens, viruses, and samples with high concentrations of inhibitors. This will avoid inter-laboratory variability e.g. associated with dilution errors in qPCR standard curves.
- The use of ddPCR should also be considered for quantifying the target in fecal samples, which are considered to be highly inhibited.
- Standard curves for qPCR could also be calibrated using ddPCR to reduce the effect of systematic errors that could be associated with the standard curve. This would increase the efficiency and accuracy of the qPCR quantification while enabling researchers to do the bulk of the analysis on the lower-cost qPCR platform.

71.4 Survival and persistence of human-associated *Bacteroidales* cells and DNA in sediments

Finding from the three first studies discussed in Chapter 3, Chapter 4 and Chapter 5 led to additional question about whether the human-associated *Bacteroidales* markers (B. theta and BacHum) validated for MST could be applied for the detection of recent or

aged human fecal contamination. In addition, I wanted to address whether free DNA from dead cells of *Bacteroidales* persists longer in the environment as such long persistence could affect the credibility of the markers in providing accurate information about ongoing fecal contamination such as a leak from a sewer pipeline.

Bacteroidales anaerobic living conditions also contributed to my curiosity and intent to rule out all potential poor information that could be associated with use of the markers in Southeast Asia. Looking at a water body and also based on recent scientific literature, I decided to focus on anoxic and dark conditions found at the level of sediments as they are favorable living conditions for anaerobic bacteria (e.g. *Bacteroidales*). Through microscopic examination of freshwater samples (water column and sediments overlaying water) from a tropical reservoir in Singapore, I learned that the sediments level had a high number of protozoa (ciliates, flagellates, and amoeba). I added then a question of examining whether protozoa could be a factor influencing decay of validated human markers. As the first chapter suggested FIB (*E. coli*) as a better predictor of Pathogen-like species (PLS) and sewage associated sequencing (SAT), I wanted to examine if sediments could be a storage site for such bacteria that would lead to a resuspension of FIB in water column. Finally, the knowledge of fate and transport of these markers in a tropical freshwater system could be included in models used for the prediction and potential impact assessment.

My work on this project demonstrated that human-associated *Bacteroidales* markers took up to 6 d for a 2-log removal from sediments and water column. *Bacteroidales* cells and DNA were not found to survive longer in sediments confirming the use for human markers for monitoring recent human fecal contamination. In contrast, *E. coli*

persisted the most in sediments reaching a 2-log removal after 14 d. In addition, the control microcosms showed evidence of growth during the time of experiment. However, the concentration of *E. coli* in controls did not exceed 10^4 MPN/g. on the contrary, *Enterococci*, another FIB studied, reached a 2-log removal after 3 d. However, it took up to 10 d to reach a 2-log removal in sediment overlaying water.

The combination of the results of Chapter 3 that reported high correlation of *E. coli* and proportion of pathogen-like bacteria are complemented by my decay results of chapter 6 that showed different 2-log removal times for *E. coli* (up to 14 d) and *Bacteroidales* viable cells (up to 6 d). The faster decay rate of human *Bacteroidales* viable cells in sediments relative to *E. coli* could indicate that *Bacteroidales* is a better tracer for recent contamination of anoxic sediments and could potentially contribute to the correlation between *E. coli* and PLS, but not between the *Bacteroidales* marker HF183 and PLS as demonstrated in Chapter 3. In addition, analysis of controls indicated that the background level of *E. coli* was below 10^4 MPN/ 1 g of sediments or 100ml of water confirming that high concentrations of *E. coli* ($>10^4$) could be linked to sewage addition and the likely presence of PLS and SAT.

Presence of protozoa was found to not have any effect to *Bacteroidales* cells or DNA as quantified by human markers. However, limited grazing (reduction of protozoa by size $< 5\mu\text{m}$) could influence the survival period of FIB in sediments. This information obtained by comparing microcosm with grazing versus those with limited grazing (protozoa size $< 5\mu\text{m}$) confirmed observations made by related studies (different set up) in temperate climate or subtropical climate. The decay rate of FIB in sediments

could be prolonged by their interaction with autochthonous bacteria, protozoa, and not only environmental physicochemical factors.

Considerations for future research:

- In this study decay of human-associated *Bacteroidales* genetic markers were studied under dark anoxic and high temperature (30°C) conditions associated with sediments in tropical freshwaters. I observed faster decay of *Bacteroidales* in sediments under tropical conditions than Kim and Wuertz (2015) observed under conditions associated with temperate waters (6 and 20°C) suggesting a potential role for temperature in the survival of the marker in different regions (Kim and Wuertz, 2015). Therefore, as different human-associated markers are validated in different regions worldwide, more decay studies are needed.
- There is a growing evidence of the role of protozoa in influencing decay and survival of *Bacteroidales* (Kobayashi et al., 2013) and fecal indicator bacteria (Wanjugi et al., 2016). My study showed that under anoxic sediment conditions protozoan grazing could reduce survival of FIB in sediments. In contrast my study did not show that protozoa influenced the decay of human-associated *Bacteroidales* markers. To my knowledge, this is the first study to address the role of grazing on the fate of water quality markers under anoxic conditions under tropical environment. Whether *Bacteroidales* survival under aerobic conditions are affected by grazing remains to be determined and future studies could focus on studying the role of protozoa under aerobic conditions.

7.1.5 Outlook

My doctoral research work applied molecular-based tools to contribute to the improvement of microbial water quality assessment under tropical regions and provided scientific insights for the application of molecular tools for water quality monitoring worldwide. My main goal of improving microbial source detection and tracking, and evaluating the relationship of fecal indicator and microbial community dynamic in tropical urban environment was achieved.

The combination of the results generated using molecular-based tool such as Next generation sequencing (NGS), qPCR and culture counts (FIB measurements) revealed that *E. coli* and *Enterococcus* could be applied for freshwater microbial water quality evaluation with a high threshold cut-off to remove environmental backgrounds. Next generation sequencing (NGS) could be applied to simultaneously detect multiple human pathogens-like bacteria (PLS) and sewage associated taxa (SAT), thus providing a useful packaged information on the status of freshwater water quality assessed. As EPA recommended HF183 for human fecal source monitoring in the USA and other studies confirming its use in other temperature climates, my study demonstrated that other markers such as *B. theta* are better under the tropical environmental conditions of Southeast Asia. Validated human-*Bacteroidales* markers were also studied for decay rates together with FIB, which provide an environmental manager with necessary information for microbial water quality assessment and modeling their fate under tropical environment of Southeast Asia.

The increasing population in coast tropical urban cities in Southeast Asia and elsewhere coupled with the increasing of the burden of waterborne disease in developing countries support the need of advanced and accurate microbial water quality tools to improve watershed management at local scale rural or urban. The efforts to improve microbial water quality monitoring under tropical environments and worldwide is still needed as global warming brings more microbial water quality related challenges.

References: Chapter 7

- AHMED, W., YUSUF, R., HASAN, I., GOONETILLEKE, A. & 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, 838-845.
- ASHBOLT, N. J. 2015. Microbial Contamination of Drinking Water and Human Health from Community Water Systems. *Current Environmental Health Reports*, 2, 95-106.
- BAIN, R., CRONK, R., HOSSAIN, R., BONJOUR, S., ONDA, K., WRIGHT, J., YANG, H., SLAYMAKER, T., HUNTER, P., PRÜSS-USTÜN, A. & BARTRAM, J. 2014a. Global assessment of exposure to faecal contamination through drinking water based on a systematic review. *Tropical Medicine & International Health*, 19, 917-927.
- BAIN, R., CRONK, R., WRIGHT, J., YANG, H., SLAYMAKER, T. & BARTRAM, J. 2014b. Fecal Contamination of Drinking-Water in Low- and Middle-Income Countries: A Systematic Review and Meta-Analysis. *PLOS Medicine*, 11, e1001644.
- FOGARTY, L. R. & VOYTEK, M. A. 2005. Comparison of *Bacteroides*-prevotella 16S rRNA genetic markers for fecal samples from different animal species. *Applied and Environmental Microbiology*, 71, 5999-6007.
- HARWOOD, V., STALEY, C., BADGLEY, B., BORGES, K. & KORAJKIC, A. 2014. Microbial source tracking markers for detection of fecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews*, 38, 1-40.
- HARWOOD, V. J., BROWNELL, M., WANG, S., LEPO, J., ELLENDER, R. D., AJIDAHUN, A., HELLEIN, K. N., KENNEDY, E., YE, X. & FLOOD, C. 2009. Validation and field testing of library-independent microbial source tracking methods in the Gulf of Mexico. *Water Research*, 43, 4812-9.
- HINDSON, B. J., NESS, K. D., MASQUELIER, D. A., BELGRADER, P., HEREDIA, N. J., MAKAREWICZ, A. J., BRIGHT, I. J., LUCERO, M. Y., HIDDESEN, A. L., LEGLER, T. C., KITANO, T. K., HODEL, M. R., PETERSEN, J. F., WYATT, P. W., STEENBLOCK, E. R., SHAH, P. H., BOUSSE, L. J., TROUP, C. B., MELLEN, J. C., WITTMANN, D. K., ERNDT, N. G., CAULEY, T. H., KOEHLER, R. T., SO, A. P., DUBE, S., ROSE, K. A., MONTESCLAROS, L., WANG, S., STUMBO, D. P., HODGES, S. P., ROMINE, S., MILANOVICH, F. P., WHITE, H. E., REGAN, J. F., KARLIN-NEUMANN, G. A., HINDSON, C. M., SAXONOV, S. & COLSTON, B. W. 2011. High-Throughput Droplet Digital PCR System for Absolute Quantitation of DNA Copy Number. *Analytical Chemistry*, 83, 8604-8610.

- HINDSON, C. M., CHEVILLET, J. R., BRIGGS, H. A., GALLICHOTTE, E. N., RUF, I. K., HINDSON, B. J., VESSELLA, R. L. & TEWARI, M. 2013. Absolute quantification by droplet digital PCR versus analog real-time PCR. *Nature Methods*, 10, 1003-5.
- HUGGETT, J. F., FOY, C. A., BENES, V., EMSLIE, K., GARSON, J. A., HAYNES, R., HELLEMANS, J., KUBISTA, M., MUELLER, R. D., NOLAN, T., PFAFFL, M. W., SHIPLEY, G. L., VANDESOMPELE, J., WITTEWER, C. T. & BUSTIN, S. A. 2013. The digital MIQE guidelines: Minimum Information for Publication of Quantitative Digital PCR Experiments. *Clinical Chemistry*, 59, 892-902.
- JONES, G. W. 2013. The population of Southeast Asia. *Working paper*. Singapore: Asia Research Institute.
- KILDARE, B., LEUTENEGGER, C., MCSWAIN, B., BAMBIC, D., RAJAL, V. & WUERTZ, S. 2007. 16S rRNA-based assays for quantitative detection of universal, human-, cow-, and dog-specific fecal Bacteroidales: A Bayesian approach. *Water Research*, 41, 15.
- KIM, M. & WUERTZ, S. 2015. Survival and persistence of host-associated Bacteroidales cells and DNA in comparison with *Escherichia coli* and *Enterococcus* in freshwater sediments as quantified by PMA-qPCR and qPCR. *Water Research*, 87, 182-192.
- KOBAYASHI, A., SANO, D. & OKABE, S. 2013. Effects of temperature and predator on the persistence of host-specific *Bacteroides-Prevotella* genetic markers in water. *Water Science and Technology*, 67, 838-45.
- LAYTON, B., CAO, Y., EBENTIER, D., HANLEY, K., BALLESTÉ, E., BRANDÃO, J., BYAPPANAHALLI, M., CONVERSE, R., FARNLEITNER, A., GENTRY SHIELDS, J., GIDLEY, M., GOURMELON, M., LEE, C., LEE, J., LOZACH, S., MADI, T., MEIJER, W., NOBLE, R., PEED, L., REISCHER, G., RODRIGUES, R., ROSE, J., SCHRIEWER, A., SINIGALLIANO, C., SRINIVASAN, S., STEWART, J., VAN DE WERFHORST, L. C., WANG, D., WHITMAN, R., WUERTZ, S., JAY, J., HOLDEN, P., BOEHM, A., SHANKS, O. & GRIFFITH, J. 2013. Performance of human fecal anaerobe-associated PCR-based assays in a multi-laboratory method evaluation study. *Water Research*, 47, 6897-6908.
- MCDERMOTT, G. 2013. Multiplexed Target Detection Using DNA-Binding Dye Chemistry in Droplet Digital PCR. *Analytical chemistry*, 85, 11619.
- NEAVE, M., LUTER, H., PADOVAN, A., TOWNSEND, S., SCHOBEN, X. & GIBB, K. 2014. Multiple approaches to microbial source tracking in tropical northern Australia. *Microbiology Open*, 3, 860-874.

- NOBLE, R. T., BLACKWOOD, A. D., GRIFFITH, J. F., MCGEE, C. D. & WEISBERG, S. B. 2010. Comparison of Rapid Quantitative PCR-Based and Conventional Culture-Based Methods for Enumeration of *Enterococcus* spp. and *Escherichia coli* in Recreational Waters. *Applied and Environmental Microbiology*, 76, 7437-7443.
- NSHIMYIMANA, J. P., EKKLESIA, E., SHANAHAN, P., CHUA, L. H. C. & THOMPSON, J. R. 2014. Distribution and abundance of human-specific *Bacteroides* and relation to traditional indicators in an urban tropical catchment. *Applied Microbiology*, 116, 1369-1383.
- ODAGIRI, M., SCHRIEWER, A., HANLEY, K., WUERTZ, S., MISRA, P. R., PANIGRAHI, P. & JENKINS, M. W. 2015. Validation of *Bacteroidales* quantitative PCR assays targeting human and animal fecal contamination in the public and domestic domains in India. *Science of the Total Environment*, 502, 462-470.
- PANDEY, P. K., KASS, P. H., SOUPIR, M. L., BISWAS, S. & SINGH, V. P. 2014. Contamination of water resources by pathogenic bacteria. *AMB Express*, 4, 51-51.
- PILLET, S., BERTHELOT, P., GAGNEUX-BRUNON, A., MORY, O., GAY, C., VIALON, A., LUCHT, F., POZZETTO, B. & BOTELHO-NEVERS, E. 2016. Contamination of healthcare workers' mobile phones by epidemic viruses. *Clin Microbiol Infect*, 22, 456.e1-6.
- PINHEIRO, L. B., COLEMAN, V. A., HINDSON, C. M., HERRMANN, J., HINDSON, B. J., BHAT, S. & EMSLIE, K. R. 2012. Evaluation of a droplet digital polymerase chain reaction format for DNA copy number quantification. *Analytical Chemistry*, 84, 1003-11.
- ROCHELLE-NEWALL, E., NGUYEN, T. M., LE, T. P., SENGTAHEUANGHOUNG, O. & RIBOLZI, O. 2015. A short review of fecal indicator bacteria in tropical aquatic ecosystems: knowledge gaps and future directions. *Frontiers in Microbiology*, 6, 308.
- ROCHELLE-NEWALL, E. J., RIBOLZI, O., VIGUIER, M., THAMMAHACKSA, C., SILVERA, N., LATSACHACK, K., DINH, R. P., NAPORN, P., SY, H. T., SOULILEUTH, B., HMAIMUM, N., SISOUVANH, P., ROBAIN, H., JANEAU, J.-L., VALENTIN, C., BOITHIAS, L. & PIERRET, A. 2016. Effect of land use and hydrological processes on *Escherichia coli* concentrations in streams of tropical, humid headwater catchments. *Scientific Reports*, 6, 32974.
- SCOTT, T., ROSE, J., JENKINS, T., FARRAH, S. & LUKASIK, J. 2002. Microbial source tracking: current methodology and future directions. *Applied and Environmental Microbiology*, 68, 5796-803.

- SUZUKI, T. A. & WOROBEY, M. 2014. Geographical variation of human gut microbial composition. *Biology Letters*, 10.
- TAN, B., NG, C., NSHIMYIMANA, J. P., LOH, L. L., GIN, K. Y. & THOMPSON, J. R. 2015. Next-generation sequencing (NGS) for assessment of microbial water quality: current progress, challenges, and future opportunities. *Frontiers in Microbiology*, 6, 1027.
- WANJUGI, P., FOX, G. A. & HARWOOD, V. J. 2016. The Interplay Between Predation, Competition, and Nutrient Levels Influences the Survival of *Escherichia coli* in Aquatic Environments. *Microbial Ecology*, 72, 526-37.
- WHILEY, H., KEEGAN, A., FALLOWFIELD, H. & ROSS, K. 2014. Uncertainties associated with assessing the public health risk from *Legionella*. *Frontiers in Microbiology*, 5, 501.
- WHO 2014. Water Quality and Health. Drinking water chlorination – A review of disinfection practices and issues. *WORLD Health Organization*.
- WIDMER, K., VAN HA, N. T., VINITNANTHARAT, S., STHIANNOPKAO, S., WANGSAATMAJA, S., PRASETIATI, M. A., THANH, N. C., THEPNOO, K., SUTADIAN, A. D., THAO, H. T., FAPYANE, D., SAN, V., VITAL, P. & HUR, H. G. 2013. Prevalence of *Escherichia coli* in surface waters of Southeast Asian cities. *World Journal of Microbiology and Biotechnology*, 29, 2115-24.

Appendix: Chapter 3

Supplemental Table 1: Average Bray-Curtis resemblance of bacterial communities within and between sites sampled in January and July 2009.

Taxonomy	Criteria	Reservoir sites included		Reservoir sites excluded	
		Inter-site	Intra-site	Inter-site	Intra-site
OTUs level	Total number	612	18	420	15.0
	Average B-C similarity	15.3	33.7	18.3	29.5
	Standard deviation	12.6	21.1	9.4	20.6
	Standard error	0.6	8	0.9	7.6
	Minimum	0.6	2.2	2.2	2.6
	Maximum	86.6	73	46.3	73
	Median	12.2	30.4	17.3	26
	T test (two tails, type 3)	F=0.3; P=0.002		F=0.2; P=0.05	

Supplemental Table 2: Summary of PERMANOVA results for evaluation of land-use and sample month as factors that structure the bacterial community in the catchment.

Source	df	SS	MS	Pseudo-F	P (Value)	Number of Permutatio ns
Bacterial community structure (OTU)						
Land use	2	8,788	4,394	2.0	0.009	998
Sampling Month	1	2,725	2,725	1.2	0.16	999
Interaction of land use and month	2	2,224	1,112	0.5	1	996
Sewage-associated taxa (SAT) at genus level						
Land use	2	9,242	4,621	3	0.004	999
Sampling Month	1	1,323	1,323	1	0.5	999
Interaction of land use and month	2	1,514	757	0.4	1	999
Pathogen-like sequences at genus level (PLSg)						
Land use	2	3,614	1,807	2	0.031	998
Sampling Month	1	2,505	2,505	2	0.008	999
Interaction of land use and month	2	1,543	771	0.7	1	999
Pathogen-like sequences at species level (PLSs)						
Land use	2	2,226	1,113	2	0.031	999
Sampling Month	1	1,161	1,161	2	0.057	998
Interaction of land use and month	2	632	316	0.5	1	997

Supplemental Table 3. Sewage associated taxa (Genus level) from this study

Family	Genus	Taxa associated with human gut or sewage*	Taxa identified as sewage-associated (McLellan et al. 2010)
<i>Bifidobacteriaceae</i>	<i>Bifidobacterium</i>	√	√
<i>Bacteroidaceae</i>	<i>Bacteroides</i>	√	√
<i>Porphyromonadaceae</i>	<i>Parabacteroides</i>	√	√
<i>Ruminococcaceae</i>	<i>Faecalibacterium</i>	√	√
<i>Lachnospiraceae</i>	<i>Roseburia</i>	√	√
<i>Ruminococcaceae</i>	<i>Ruminococcus</i>	√	√
<i>Alcaligenaceae</i>	<i>Sutterella</i>	√	√
<i>Akkermansiaceae</i>	<i>Akkermansia</i>	√	√
<i>Bacteroidaceae</i>	<i>Papillibacter</i>	√	√
<i>Bacteroidaceae</i>	<i>Subdoligranulum</i>	√	√
<i>Prevotellaceae</i>	<i>Prevotella</i>	√	-
<i>Rikenellaceae</i>	<i>RC9_gut_group</i>	√	-
<i>Clostridiaceae</i>	<i>Clostridium</i>	√	-
<i>Veillonellaceae</i>	<i>Veillonella</i>	√	-
<i>Streptococcaceae</i>	<i>Streptococcus</i>	√	-

* based on sequence annotation or literature search

“√” Taxa classified by indicated criteria

“-” Taxa not classified by indicated criteria

Supplemental Table 4. Pathogen-like bacterial 16S rRNA gene sequences recovered in this study.

OTUs from this study	Pathogen Genus*	Human species**	pathogen	Accession Numbers	Diseases	Clinical References
OTU0005	<i>Acinetobacter</i>	<i>Acinetobacter junii</i>		Z93438.1	Septicemia	(Linde et al. 2002)
OTU0511	<i>Acinetobacter</i>	<i>Acinetobacter ursingii</i>		NR_025392.1	Bacteremia	(Loubinoux et al. 2003)
OTU1889	<i>Acinetobacter</i>	<i>Acinetobacter soli</i>		NR_044454.1	Bloodstream infection	(Pellegrino et al. 2011)
OTU0070 OTU0706 OTU0382	& <i>Clostridium</i>	<i>Clostridium paraputrificum</i>		NR_113021.1	Bloodborne pathogen	(Woo et al. 2005)
OTU8629 OTU0303	& <i>Clostridium</i>	<i>Clostridium disporicum</i>		NR_026491.1	Opportunistic infection for immuno-compromised	(Shandera et al. 1988)
	& <i>Clostridium</i>	<i>Clostridium perfringens</i>		NR_121697.1	Diarrhea	(Bennett et al. 2013)
OTU2955	<i>Clostridium</i>	<i>Clostridium intestinale</i>		AY781385.1	Bacteremia	(Elsayed and Zhang 2005)
OTU0965	<i>Rickettsia</i>	<i>Rickettsia honei</i>		NR_114449.1	Fever	(Murphy et al. 2011)
OTU0131	<i>Rickettsia</i>	<i>Rickettsia montanensis</i>		CP003340.1	Fever	(Ammerman et al. 2004)
OTU0025	<i>Pseudomonas</i>	<i>Pseudomonas mendocina</i>		KF862906.1	Endocarditis in a patient with situs inversus	(Mert et al. 2007)
OTU0232	<i>Pseudomonas</i>	<i>Pseudomonas oryzihabitans</i>		NR_114041.1	Peritonitis, endophthalmitis, and bacteremia	(Esteban et al. 2002)
OTU3872 OTU7452	& <i>Pseudomonas</i>	<i>Pseudomonas stutzeri</i>		AF094748	Opportunistic pathogen	(Lalucat et al. 2006)
OTU9011 OTU3265	& <i>Legionella</i>	<i>Legionella pneumophila</i>		NR_074231.1	Pneumonia / Opportunistic	(Muder et al. 2002)
OTU0963 OTU1273 OTU0916	& <i>Legionella</i>	<i>Legionella gresilensis</i>		NR_028744.1	Legionnaires' Disease	(Lo Presti et al. 2001)
	<i>Legionella</i>	<i>Legionella rowbothamii</i>		NR_036804.1	Legionnaires' Disease	(Birtles et al. 1996)
OTU5429	<i>Staphylococcus</i>	<i>Staphylococcus sciuri</i>		AB212276.1	Endocarditis	(Stepanovic et al. 2003)
OTU1038	<i>Staphylococcus</i>	<i>Staphylococcus epidermidis</i>		AB680360.1	Bacteremia	(Blum and Rodvold 1987)
OTU0440	<i>Yersinia</i>	<i>Yersinia enterocolitica</i>		HE803717.1	Gastroenteritis	(Sihvonen et al. 2012)
OTU0163	<i>Aeromonas</i>	<i>Aeromonas tecta</i>		NR_118043.1	Bacteremia	(Demarta et al. 2008)
OTU0024	<i>Shigella</i>	<i>Shigella sonnei</i>		NR_074894.1	Shigellosis	(Yang et al. 2005)
OTU0622	<i>Streptococcus</i>	<i>Streptococcus oralis</i>		AY485602.1	Septicemia	(Arbique et al. 2004)
OTU0758	<i>Enterococcus</i>	<i>Enterococcus casseliflavus</i>		NR_041704.1	Enterococcal meningitis	(Iaria et al. 2005)
OTU1693	<i>Haemophilus</i>	<i>Haemophilus parainfluenzae</i>		NR_116168.1	Bronchiectasis	(Mitchell and Hill 2000)
OTU1204	<i>Serratia</i>	<i>Serratia marcescens</i>		FJ584421	Sepsis (neonatal unit)	(Mahlen 2011)
OTU0887	<i>Burkholderia</i>	<i>Burkholderia dolosa</i>		NR_104973.1	Lung disease	(Moreira et al. 2014)
OTU0321	<i>Campylobacter</i>	<i>Campylobacter helveticus</i>		U03022.1	Gastroenteritis	(Linton et al. 1994)
OTU1837	<i>Mycobacterium</i>	<i>Mycobacterium setense</i>		EU371507.1	Lower Respiratory Tract Infection	(Toro et al. 2008)
OTU_1166	<i>Corynebacterium</i>	<i>Corynebacterium tuberculostearicum</i>		NR_028975.1	Isolated in patients with mastitis	(Paviour et al. 2002)
OTU_16& OTU_6394	<i>Arcobacter</i>	<i>Arcobacter cryaerophilus</i>		EU669899.1	Diarrhea	(Figueras et al. 2014)

* Taxonomic affiliation with genus listed on US NIH etiological agents list

** Top BlastN match with >99% nucleotide identity

Supplemental Table 5. OTUs, SAT and PLS bacteria contributing to the dissimilarity between the catchment land uses and reservoir and the catchment (Attached excel spread sheet)

Criteria	Taxa	Catchment: Average Abundance	Reservoir: Average Abundance	Average Dissimilarity	Dissimilarity /Standard deviation	% Contribution	% Average Dissimilarit y
Dissimilarity between Catchment and Reservoir Microbial Communities	OTUs level contributing to dissimilarity between Catchment and Reservoir						
	OTU_36 (<i>Flexibacter</i>)	0.7	5.3	0.1	1.7	0.2	74.7
	OTU_31 (<i>Chitinophagaceae Spp.</i>)	1.1	5.5	0.1	1.7	0.2	
	OTU_26 (<i>Flexibacter</i>)	1.4	5.2	0.1	1.5	0.1	
	OTU_47 (<i>hgcI_clade</i>)	1.5	4.9	0.1	1.5	0.1	
	OTU_30 (<i>Microcystis</i>)	1.7	5.0	0.1	1.6	0.1	
	OTU_18 (<i>Alcaligenaceae Spp.</i>)	1.6	4.7	0.1	1.6	0.1	
	OTU_87 (<i>NS9_marine_group Spp.</i>)	0.5	3.6	0.1	1.5	0.1	
	OUT_7 (<i>Arcobacter</i>)	3.2	0.6	0.1	1.7	0.1	
	OUT_6 (<i>12up</i>)	3.1	2.9	0.1	1.3	0.1	
	OUT_32 (<i>Hydrogenophaga</i>)	3.5	2.0	0.1	1.5	0.1	
	SAT genus level contributing to dissimilarity between Catchment and Reservoir						
	<i>Prevotella</i>	2.8	0.2	8.0	2.3	14.7	54
	<i>Bacteroides</i>	2.9	0.9	7.0	1.9	12.9	
	<i>Parabacteroides</i>	2.6	0.8	6.1	1.6	11.2	
	<i>Clostridium</i>	3.3	1.9	4.6	2.1	8.5	
	<i>Bifidobacterium</i>	1.8	0.6	4.4	1.5	8.2	
	<i>Faecalibacterium</i>	1.4	0.4	3.8	1.2	7.0	
	<i>RC9_gut_group</i>	1.1	0.0	3.4	1.0	6.2	
	<i>Ruminococcus</i>	1.1	0.3	3.1	1.2	5.7	
	<i>Subdoligranulum</i>	1.1	0.4	3.0	1.2	5.4	
	<i>Sutterella</i>	1.0	0.2	2.9	1.2	5.4	

Supplemental Table 5.

Criteria	Taxa	Catchment: Average Abundance	Reservoir: Average Abundance	Average Dissimilarity	Dissimilarity /Standard deviation	% Contribution	% Average Dissimilarity
Dissimilarity between Catchment and Reservoir Microbial Communities	PLS genus level contributing to dissimilarity between Catchment and Reservoir						
	<i>Arcobacter</i>	4.3	1.1	4.4	2.3	10.0	44.1
	<i>Aeromonas</i>	3.0	0.4	3.6	1.9	8.1	
	<i>Acinetobacter</i>	5.0	2.1	3.6	1.7	8.1	
	<i>Enterobacter</i>	2.5	0.6	2.6	1.7	6.0	
	<i>Erysipelothrix</i>	2.0	0.0	2.6	1.8	6.0	
	<i>Pseudomonas</i>	3.3	2.0	2.1	1.6	4.7	
	<i>Yersinia</i>	1.5	0.0	2.0	1.5	4.5	
	<i>Clostridium</i>	3.3	2.0	1.9	1.9	4.3	
	PLS species level contributing to dissimilarity between Catchment and Reservoir						
	<i>Arcobacter cryaerophilus</i>	3.0	1.0	4.1	1.4	8.6	47.1
	<i>Aeromonas tecta</i>	2.0	0.0	3.6	1.8	7.6	
	<i>Acinetobacter junii</i>	3.6	2.0	2.9	1.5	6.2	
	<i>Yersinia enterocolitica</i>	1.5	0.0	2.6	1.5	5.6	
	<i>Legionella gresilensis</i>	0.4	1.5	2.6	1.3	5.5	
	<i>Shigella sonnei</i>	1.6	0.3	2.5	1.4	5.4	
	<i>Pseudomonas mendocina</i>	3.0	2.0	2.4	1.6	5.1	
	<i>Rickettsia montanensis</i>	2.0	2.7	2.1	1.2	4.4	
	<i>Enterococcus casseliflavus</i>	1.3	0.3	2.1	1.6	4.4	
	<i>Mycobacterium setense</i>	1.1	0.0	2.0	1.4	4.3	

Supplemental Table 5.

Criteria	Taxa	Horticultural: Average abundance	Reservoir: Average Abundance	Average Dissimilarity	Dissimilarity /Standard deviation	% Contribution	% Average Dissimilarity
Dissimilarity between Horticulture and Residential area microbial communities	OTUs level contributing to dissimilarity between horticulture and Residential Land Uses						
	OTU_6 (<i>12up</i>)	2.4	4.4	0.1	1.1	0.1	69
	OTU_2 (<i>Cynobacteria Spp.</i>)	2.0	2.3	0.1	0.7	0.1	
	OTU_11 (<i>Novispirillum</i>)	1.6	3.7	0.1	1.0	0.1	
	OTU_7827 (<i>12up</i>)	1.5	3.0	0.1	0.8	0.1	
	OTU_14 (<i>Tolumonas</i>)	3.4	2.4	0.1	0.8	0.1	
	OTU_35 (<i>Megasphaera</i>)	2.0	1.5	0.1	0.7	0.1	
	OTU_2638 (<i>Prevotella</i>)	2.7	0.7	0.1	1.3	0.1	
	OTU_112 (<i>Comamonadaceae Ssp.</i>)	1.6	2.9	0.1	0.9	0.1	
	OTU_32 (<i>Hydrogenophaga</i>)	3.6	4.0	0.1	1.1	0.1	
	OTU_5002 (<i>Aeromonadaceae Ssp.</i>)	2.1	1.9	0.1	0.8	0.1	
	SAT genus level contributing to dissimilarity between horticulture and Residential Land Uses						
	<i>Prevotella</i>	3.6	2.2	4.4	1.6	11.8	38
	<i>Faecalibacterium</i>	2.2	0.7	3.8	1.5	10.1	
	<i>Bifidobacterium</i>	2.4	1.3	3.7	1.5	9.7	
	<i>Ruminococcus</i>	1.8	0.5	3.1	1.7	8.1	
	<i>Subdoligranulum</i>	1.8	0.5	3.0	1.7	7.9	
	<i>Parabacteroides</i>	3.1	2.2	2.8	1.3	7.3	
	<i>RC9_gut_group</i>	1.4	0.9	2.7	1.4	7.2	
	<i>Bacteroides</i>	3.2	2.6	2.6	1.4	7.0	
	<i>Sutterella</i>	1.6	0.6	2.5	1.5	6.6	
	<i>Veillonella</i>	1.4	0.7	2.5	1.0	6.5	

Supplemental Table 5.

Criteria	Taxa	Horticultural: Average abundance	Reservoir: Average Abundance	Average Dissimilarity	Dissimilarity /Standard deviation	% Contribution	% Average Dissimilarity
Dissimilarity between Horticulture and Residential area microbial communities	PLS genus level contributing to dissimilarity between horticulture and Residential Land Uses						
	<i>Acinetobacter</i>	5.0	5.0	2.1	1.3	7.4	29.0
	<i>Aeromonas</i>	3.0	3.0	1.6	1.3	5.6	
	<i>Arcobacter</i>	4.5	4.3	1.6	1.3	5.5	
	<i>Rickettsia</i>	1.5	2.5	1.4	1.4	5.0	
	<i>Pseudomonas</i>	3.6	3.0	1.4	1.2	5.0	
	<i>Enterobacter</i>	2.7	2.4	1.3	1.1	4.6	
	<i>Yersinia</i>	1.3	2.0	1.3	1.4	4.5	
	<i>Erysipelothrix</i>	1.8	2.0	1.3	1.2	4.5	
	<i>Legionella</i>	1.3	2.0	1.2	1.3	4.2	
	<i>Escherichia-Shigella</i>	2.0	1.4	1.2	1.2	4.2	
	PLS species contributing to dissimilarity between horticulture and Residential land uses						
	<i>Arcobacter cryaerophilus</i>	3.5	2.5	3.0	1.2	8.1	35.3
	<i>Acinetobacter junii</i>	3.5	4.0	2.1	1.3	6.0	
	<i>Aeromonas tecta</i>	2.0	2.1	2.0	1.2	5.4	
	<i>Yersinia enterocolitica</i>	1.3	2.0	2.0	1.3	5.2	
	<i>Rickettsia montanensis</i>	1.4	2.3	2.0	1.4	5.0	
	<i>Pseudomonas oryzihabitans</i>	1.1	1.0	1.7	1.2	5.0	
	<i>Clostridium disporicum</i>	2.3	1.3	1.7	1.2	5.0	
	<i>Shigella sonnei</i>	2.0	1.4	1.7	1.1	5.0	
	<i>Pseudomonas mendocina</i>	3.1	3.0	1.6	1.3	4.6	
	<i>Clostridium perfringens</i>	1.4	0.7	1.6	1.4	4.4	
	<i>Rickettsia honei</i>	0.5	1.0	1.4	1.2	4.0	

Supplemental Figure 1: Hierarchical clustering of the distribution of selected pathogen-like sequences identified to species level (PLSs) including the ten most highly-represented and the five most closely related to gastro-intestinal pathogens. Comparison includes the proportion of % SAT, *E. coli*, and HF183. Heat map indicate sequence abundance as a fraction of the total sequences per sample and are individually scaled. *E. coli* and HF183 concentrations are log10 transformed.

Appendix: Chapter 4

Supplemental Table 1: Number of samples collected, pooled and quantified in this study

Sample type	n	Pooled samples*	Samples analyzed by QPCR
Human			
Stool	35	Not pooled	35
Sewage	30	(10) = 1	21
Domesticated animals			
Dogs	50	(40) = 4	14
Cats	50	(40) = 4	14
Rabbits	40	(30) = 3	13
Chickens	16	Not pooled	10
Wild animals			
Wild Boars	14	Not pooled	10
Monkeys	10	Not pooled	10
Birds	50	(40) = 4	14
Total	295	-	141

* Total number of samples pooled in batches of 10 given in brackets

Supplemental Table 2: Comparison of human-associated gene markers concentrations in host target samples (human stool and sewage samples)

<i>Bacteroidales</i>- associated Assay	Human stool¹	Sewage¹	T-test P value
HF183_SYBR	5.8 ± 1.80 (B)	4.2 ± 0.52 (b)	p<0.0001
HF183	6.0 ± 1.80 (B)	4.0 ± 0.64 (b)	p<0.0001
BacHum	6.0 ± 1.55 (B)	4.1 ± 0.60 (b)	p<0.0001
BacH	5.70 ± 1.45 (B)	4.1 ± 0.68 (b)	p<0.0001
B. theta	6.0 ± 0.97 (B)	3.5 ± 0.42 (c)	p<0.0001
BacCan	3.90 ± 0.60 (C)	0.30 ± 0.06 (d)	p<0.0001
BacUni	7.7 ± 0.80 (A)	5.2 ± 0.27 (a)	p<0.0001
ANOVA	F=20.73; p<0.0001	F=112.30; p<0.0001	

¹ Values reported are the mean and standard deviation of log-transformed marker concentrations reported in gene copies/gr wet feces for human stool and gene copies/ml for sewage samples. Letters in brackets denote results from Tukey's honestly significant post-hoc test for comparison of marker concentrations indicated by different assays in the same type of sample. Different letters indicate statistically significant differences between groups (p<0.05).

Supplemental Table 3. Reproducibility measured by two standard (10^2 and 10^3 copies/reaction of the positive control) run with every qPCR plate per every *Bacteroidales*

<i>Bacteroidales</i> markers	Cp values of 10^2 Copies/ μ l			Cp values of 10^3 Copies/ μ l			
	Average Cp values	SD	% CV	Average values	Cp	SD	%CV
HF183	32.2	0.7	2.0	28.1		0.6	2.1
HF183-SYBRG	29.2	0.6	2.1	26.3		0.3	1.3
BacHum	32.6	0.5	1.5	28.1		0.1	0.3
BacH	30.2	0.5	1.6	26.9		0.5	2.0
B. theta	32.3	0.2	0.7	29.0		0.3	1.0
BacCan	33.0	0.8	2.4	30.9		0.7	2.2
BacUni	30.7	0.5	1.5	27.6		0.6	2.0

markers

Notes: SD: Standard deviation;

%CV: Percentage of Coefficient of variation;

Cp: Crossing point and six plates were run per each marker

Supplemental Table 4. Number of template and qPCR control samples run with all samples quantified using *Bacteroidales* assays

<i>Bacteroidales</i> Markers	Number 96-wells plates	Template control Number NTD	qPCR reagent mix control of Number NTD
HF183	6	12	12
HF183-SYBR	6	12	12
BacHum	6	12	12
B. theta	6	12	12
BacH	6	12	12
BacCan	6	12	12
BacUni	6	12	12

Notes: Duplicates were run for template and qPCR controls and **NTD:** No Target Detected

Supplemental Figure 1: Comparison of measured and expected concentrations of a serial dilution of HF183 plasmid DNA quantified using SYBR Green and TaqMan fast advanced master mix on LightCycler480. The blue colored equation and dots represent HF183-SYBR Green (reverse primer designed Seurinck et al. 2005); Orange colored equation and dots represent the TaqMan-based HF183 assay (reverse primer designed by Haugland et al. 2010); and black colored equation and dots represent TaqMan-based HF183 assay (reverse primer designed by Green et al. 2014). Trend line represent the linear regression representing the best fit between expected and measured with a slope of 1 and R^2 ranging from 0.99 to 1.00.

Appendix: Chapter 5

Supplemental Figure 1 (S1). Concentrations of standard in positive control run on separate 96-well plates by ddPCR (S2A: 10^4 copies/reaction and S2B: 10^3 copies/reaction)

Supplemental Table 1 (S2). Summary of ANOVA analysis for ddPCR and qPCR quantification of sewage and human stool samples

Samples	Factors	Indicators	F values	P values
Sewage	qPCR versus ddPCR	B. theta	24.5	p < 0.001
		BacHum	12.5	p = 0.0011
		Both markers	12.0	p = 0.0009
	B. theta versus BacHum	ddPCR	133.6	p < 0.001
		qPCR	306.6	p < 0.0001
		Both platforms	195.4	p < 0.001
Human stool	qPCR versus ddPCR	B. theta	8.0	p = 0.0062
		BacHum	13.0	p = 0.0006
		Both markers	13.5	p = 0.0003
	B. theta versus BacHum	ddPCR	29.8	p < 0.001
		qPCR	47.0	p < 0.001
		Both platforms	63.4	p < 0.001

Appendix: Chapter 6

Supplemental Figure 1. Comparison of protozoa counts, enumeration of fecal indicator bacteria (*E. coli* and *Enterococcus*) and quantification of intact cells (ddPCR-PMA) of BacHum and *B. theta* in treatment and < 5 μ m control microcosms. Protozoa counts, *Enterococcus* and *E. coli* in treatment microcosm (A). Protozoa counts, *Enterococcus* and *E. coli* in < 5 μ m control microcosm (B). Protozoa counts, qPCR-PM