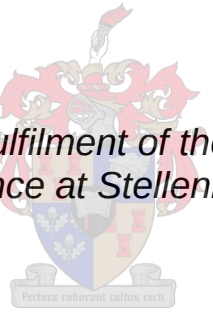
Identifying the Primary Microbial and Chemical Source Tracking Markers in Harvested Rainwater for the Detection of Faecal Contamination

by

Monique Waso



Thesis presented in partial fulfilment of the requirements for the degree Master of Science at Stellenbosch University

Supervisor: Prof. Wesaal Khan

Co-supervisor: Dr Sehaam Khan

Department of Microbiology

Faculty of Science

March 2017

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

March 2017

Signature:…… Date: ……

SUMMARY

Rainwater harvesting has been earmarked as an additional source of fresh water. However, research has indicated that the microbiological quality is substandard as pathogens have been detected in this water source. As it is impractical to monitor for the presence of all pathogens in a water source, indicator organisms are routinely utilised to monitor water quality and predict the presence of pathogens in contaminated environmental waters. Various research groups have however indicated that the analysis of indicator organisms in a water source may not be sufficient to accurately identify the source of contamination. Supplementary indicators are therefore required to accurately identify contamination sources, with chemical and microbial source tracking markers currently being investigated and applied to various water sources. The primary focus of the current study was thus to identify a toolbox of microbial source tracking (MST) and chemical source tracking (CST) markers that could be utilised to supplement indicator organism analysis of domestic rainwater harvesting (DRWH) systems.

To achieve this aim, harvested rainwater (n = 60) and rooftop debris (n = 60) samples were screened for a range of MST (conventional PCR) and CST (high-performance liquid chromatography tandem mass spectrometry) markers previously utilised in literature to analyse various water sources (**Chapter two**). All the tank water samples collected at the Kleinmond Housing Scheme site (Kleinmond, Western Cape), were also screened for traditional indicator organisms using culture based techniques. Additionally, *Escherichia coli* (*E. coli*) and enterococci were screened for in all tank water and rooftop debris samples using quantitative PCR (qPCR) analysis. Based on the conventional PCR results, *Bacteroides* HF183, adenovirus, Lachnospiraceae and human mitochondrial DNA (mtDNA) were the most prevalent MST markers. These markers were subsequently quantified in the tank water and rooftop debris samples by qPCR. The HF183 marker was then detected at a mean concentration of 5.1×10^3 and 4.7×10^3 gene copies/µL in the tank water and rooftop debris, respectively. Adenovirus was detected at 3.2×10^2 and 6.4×10^3 gene copies/µL; human mtDNA was detected at 1.1×10^6 and 3.0×10^5 gene copies/µL and Lachnospiraceae was detected at 3.0×10^4 and 6.9×10^3 gene copies/µL in the tank water and rooftop debris samples, respectively. Additionally, *E. coli* and enterococci were quantifiable in all tank water and rooftop debris samples by qPCR analysis. The CST markers caffeine, salicylic acid, acetaminophen, triclosan, triclocarban and methylparaben were then detected at µg/L levels in all the tank water [except salicylic acid (98%)] and rooftop debris samples. A secondary aim was to establish correlations between the MST and CST markers as well as indicator organisms to ascertain which markers may be employed to supplement indicator organism analysis of DRWH systems. In the tank water samples, significant positive correlations were observed for adenovirus versus *E. coli* (enumerated with the culturing techniques) (p = 0.000), the HF183 marker versus *E. coli*

(quantified by qPCR) (p = 0.023), Lachnospiraceae versus heterotrophic bacteria (p = 0.000) and human mtDNA versus enterococci (enumerated with the culturing techniques) (p = 0.026). In addition, significant positive correlations were observed for caffeine versus enterococci (quantified by qPCR) (p = 0.000); faecal coliforms (p = 0.001); total coliforms (p = 0.000) and enterococci (enumerated with culturing techniques) (p = 0.002). Salicylic acid also positively correlated with total coliforms (p = 0.024) in the tank water samples. For the rooftop debris samples, significant positive correlations were observed for *E. coli* (quantified by qPCR) versus methylparaben (p = 0.000) and salicylic acid (p = 0.042), respectively. Based on the results obtained, it is thus evident that faecal contamination and anthropogenic activities may be the primary sources of contamination in the DRWH systems. Moreover, the markers *Bacteroides* HF183, Lachnospiraceae, human mtDNA, adenovirus, caffeine, salicylic acid and methylparaben may be utilised to supplement traditional indicator organism analysis for the monitoring of harvested rainwater. It is however recommended that future studies focus on correlation analysis of the source tracking markers with pathogens frequently detected in harvested rainwater, in order to determine which source tracking markers may be utilised as surrogates for these pathogens and subsequently as supplementary indicators.

Avian species are vectors of microorganisms in the environment and have been identified as major sources of faecal contamination of DRWH systems. The focus of **Chapter three** was thus to design and validate (on a small-scale) novel MST markers for the detection of avian faecal contamination in the DRWH systems. Three primer sets [AVF1 and AVR (designated AV1); AVF2 and AVR (designated AV2); and ND5F and ND5R (designated ND5)] were subsequently designed to target regions of the NADH dehydrogenase subunit 5 mitochondrial DNA gene of avian species. Mitochondrial DNA is abundant in animal faecal matter and may thus be readily detected. Conventional PCR assays were optimised for each of the three primer sets. Avian and non-avian faecal samples were then screened to validate the host-specificity and host-sensitivity of the mtDNA markers. The mtDNA markers AV1, AV2 and ND5 displayed a host-sensitivity of 1.00, 0.892 and 0.622, respectively. While the host-specificity of each assay was equal to 0.316, 0.0526 and 0.237 for AV1, AV2 and ND5, respectively. Tank water samples (n = 60) and rooftop debris (n = 60) were then screened for the prevalence of the three markers. Overall, AV1 was the dominant marker detected in the tank water (85%) and rooftop debris (90%) samples. Bayes' theorem then indicated that there was an 89.2% and 92.9% probability that the AV1 marker detected true avian faecal contamination in the tank water and rooftop debris samples, respectively. The AV1 marker thus exhibited the greatest potential as an avian mtDNA marker for the detection of avian faecal contamination in DRWH systems. However, based on the low host-specificity obtained for all three primer sets (AV1, AV2 and ND5), further optimisation should include the use of a Taqman™ probe to increase the specificity of this marker.

OPSOMMING

Geoeste reënwater is geïndentifiseer as ‘n addisionele vars waterbron, maar navorsing het bewys dat die mikrobiese kwaliteit substandaard is aangesien ‘n verskeidenheid patogene al in geoeste reënwater gevind is. Aangesien dit onprakties is om vir alle patogene in ‘n waterbron te toets, word indikator organismes algemeen gebruik om die kwaliteit van waterbronne te monitor en om die teenwoordigheid van patogene in die water te voorspel. Verskeie navorsingsgroepe het egter gewys dat om vir indikator organismes te toets, nie voldoende is om die bron van kontaminasie te identifiseer nie. Daar is dus ‘n behoefte aan aanvullende indikators om die bronne van kontaminasie te identifiseer. Daarom word chemiese en mikrobiese bron spoor merkers deesdae nagevors en toegepas op verskeie waterbronne. Die primêre doel van die huidige studie was dus om ‘n versameling mikrobiese bron spoor (MBS) en chemiese bron spoor (CBS) merkers te identifiseer wat gebruik mag word om die analise van indikator organismes in huishoudelike reënwater oesting (HRWO) sisteme, aan te vul.

Hierdie doel is behaal deur geoeste reënwater monsters (n = 60) en detritus monsters vanaf die dakoppervlak (n = 60) te toets vir ‘n paneel MBS (konvensionele PKR) en CBS (hoë-verrigting vloeistof chromatografie tandem massaspektrometrie) merkers, wat voorheen in die literatuur aangewend is om water te analiseer (**Hoofstuk twee**). Die tenk water monsters wat by die Kleinmond Behuisings-skema (Kleinmond, Wes-Kaap) geneem is, is ook getoets vir tradisionele indikator organismes deur gebruik te maak van groei-gebaseerde tegnieke. Daarby is daar ook vir *Escherichia coli* (*E. coli*) en enterokokkie met kwantitatiewe PKR (kPKR) in die tenk water en detritus monsters getoets. Die konvensionele PKR resultate het getoon dat *Bacteroides* HF183, adenovirus, Lachnospiraceae en menslike mitokondriale DNS (mtDNS) die mees algemene MBS merkers in die monsters was. Hierdie merkers is dus gekwantifiseer in die tenk water en detritus monsters met behulp van kPKR. Die HF183 merker is toe teen ‘n gemiddelde konsentrasie van 5.1×10^3 en 4.7×10^3 geen kopieë/µL in die tenk water en detritus monsters gekry. Adenovirus is teen 3.2×10^2 en 6.4×10^3 geen kopieë/µL; menslike mtDNS is teen 1.1×10^6 en 3.0×10^5 geen kopieë/µL en Lachnospiraceae is teen 3.0×10^4 en 6.9×10^3 geen kopieë/µL in onderskeidelik die tenk water en detritus monsters gekry. Daarbenewens was die *E. coli* en enterokokkie ook kwantifiseerbaar in al die tenk water en detritus monsters, onderskeidelik. Die CBS merkers kafeïen, salisielsuur, asetaminofen, metielparabeen, triklosaan en triklokarbaan is teen µg/L vlakke in al die tenk water [behalwe salisielsuur (98%)] en detritus monsters gekry. ‘n Tweede doel van hierdie studie was om korrelasies tussen die MBS en CBS merkers en indikator organismes te ondersoek, om vas te stel watter merkers gebruik mag word om indikator organisme analises aan te vul. In die tenk water monsters is daar beduidende positiewe korrelasies waargeneem vir adenovirus teenoor *E. coli* (groei-gebaseerd) (p = 0.000), die HF183 merker teenoor *E. coli* (kPKR) (p = 0.023), Lachnospiraceae teenoor heterotrofiese bakterieë (p = 0.000) en menslike mtDNS teenoor

enterokokkie (groei-gebaseerd) (p = 0.026). Daaropvolglik, is beduidende positiewe korrelasies opgemerk vir kafeïen teenoor enterokokkie (kPKR) (p = 0.000); fekale koliforme (p = 0.001); totale koliforme (p = 0.000) en enterokokkie (groei-gebaseerd) (p = 0.002). Salisielsuur het ook positief gekorreleer met totale koliforme (p = 0.024) in die tenk water monsters. Vir die detritus monsters is beduidende positiewe korrelasies opgemerk vir *E. coli* (kPKR) teenoor metielparabeen (p = 0.000) en salisielsuur (p = 0.042), onderskeidelik. Hierdie resultate dui dan aan dat fekale kontaminasie en antropogeniese aktiwiteite die primêre bronne van kontaminasie van die HRWO sisteme is. Verder kan *Bacteroides* HF183, Lachnospiraceae, menslike mtDNS, adenovirus, kafeïen, salisielsuur en metielparabeen gebruik word om tradisionele indikator organisme analises aan te vul om die kwaliteit van geoeste reënwater te monitor. Daar word egter aanbeveel dat toekomstige studies op korrelasies tussen bron spoor merkers en patogene, wat gereeld in geoeste reënwater gevind word, ondersoek word om vas te stel watter bron spoor merkers as surrogate vir hierdie patogene en verder as aanvullende indikators gebruik kan word.

Voël spesies is vektore van mikroorganismes in die omgewing en is geïdentifiseer as bronne van fekale kontaminasie in HRWO sisteme. Die fokus van **Hoofstuk drie** was dus om nuwe MBS merkers, om fekale kontaminasie van voëls in HRWO sisteme op te spoor, te ontwerp en op ‘n klein skaal te verifieer. Drie inleier stelle [AVF1 en AVR (benoem AV1); AVF2 en AVR (benoem AV2); en ND5F en ND5R (benoem ND5)] is dus ontwerp om dele van die NADH dehidrogenase subeenheid 5 mtDNS geen van voëls te teiken. Mitokondriale DNS is vollop in die fekale materiaal van diere en kan dus maklik geamplifiseer word. Konvensionele PKR toetse is vir elke inleier paar geoptimiseer. Fekale monsters van voël spesies en nie-voël spesies is gevolglik geanaliseer om die gasheer-sensitiwiteit en -spesifisiteit van die mtDNS merkers te verifieer. Die gasheer-sensitiwiteit was dus gelyk aan 1.00, 0.892 en 0.622 vir die AV1, AV2 en ND5 merkers, onderskeidelik, terwyl die gasheer-spesifisiteit gelyk was aan 0.316, 0.0526 en 0.237 vir die AV1, AV2 en ND5 merkers, onderskeidelik. Tenk water (n = 60) en detritus (n = 60) monsters is toe getoets vir die teenwoordigheid van die drie merkers. Die AV1 merker is as die dominante merker in die tenk water (85%) en detritus (90%) monsters geïdentifiseer. Bayes se stelling het aangedui dat daar ‘n 89.2% en 92.9% waarskynlikheid is dat die AV1 merker opgespoor is weens ware voël verwante kontaminasie in die tenk water en detritus monsters. Die AV1 merker het dus die grootste potensiaal om as ‘n mtDNS merker, vir die opsporing van voël verwante kontaminasie in HRWO sisteme, gebruik te kan word. As gevolg van die lae gasheer-spesifisiteit wat opgemerk is vir die drie inleier stelle (AV1, AV2 en ND5), word daar egter voorgestel dat hierdie merkers verder geoptimiseer moet word deur gebruik te maak van Taqman™ ondersoekers spesifiek vir voëls, om dan die spesifisiteit van die merkers te verbeter.

ACKNOWLEDGEMENTS

First and foremost, to **God Almighty**, I am grateful for the talents and abilities that I have been granted and for the strength and courage to complete my MSc. For it is only by grace that I have achieved this milestone.

Dr Wesaal Khan, my supervisor and mentor. I could not have picked a better supervisor for my MSc study. I extend my sincerest thanks for all the support, compassion, patience, guidance, encouragement and for believing in my potential. Without her guidance and support I would not have been able to successfully complete my MSc study. Her immense knowledge, leadership, commitment to excellence and passion for microbiology has been a great inspiration during my MSc.

Dr Sehaam Khan for her co-supervision, knowledge, academic assistance and continued interest and enthusiasm. Without whose assistance and knowledge Chapter three would not have been completed successfully.

Khan Lab for their support and friendship during my MSc study and their assistance during sample collection in Kleinmond. In particular, I would also like to offer my sincerest thanks to **Penny** and **Thando** for their help and guidance in the lab and for always being there.

Harry Crossley Foundation for financial support during the second year of my MSc study.

Department of Microbiology and Stellenbosch University for all the support, friendships and the opportunity to complete a MSc degree.

My family:

Brandon Reyneke for his endless love and support, continuous motivation and understanding throughout my studies. Without whose companionship, I would not have been able to achieve this milestone.

Piet and **Veronica Reyneke** for all their love, support and continuous motivation. For being my second home and always believing in my potential.

My brother, **Frikkie Waso**, for always being there for me and for his continuous love and support.

A special thanks to my parents, **Dominique** and **Jenny Waso**, for their constant love, support, encouragement and motivation. For always being there for me and for believing in my dreams. Without whom I would not have been able to complete my MSc study. The sacrifices and unrelenting support in the last two years has not gone unnoticed.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND ACRONYMS

ADWG	Australian Drinking Water Guidelines
BDL	Below Detection Limit
BLAST	Basic Local Alignment Search Tool
CAF	Central Analytical Facility
CDC	Centres for Disease Control and Prevention
cDNA	Complementary Deoxyribonucleic Acids
CFU	Colony Forming Units
CST	Chemical Source Tracking
CSIR	Council for Scientific and Industrial Research
DRWH	Domestic Rainwater Harvesting
DWAF	Department of Water Affairs and Forestry
EDTA	Ethylenediaminetetraacetic Acid
EHEC	Enterohemorrhagic *Escherichia coli*
EMA	Ethidium Monoazide Bromide
FAO	Food and Agricultural Organisation
FIB	Faecal Indicator Bacteria
FN	False Negative
FP	False Positive
HLB	Hydrophilic-Lipophilic Balanced
HPLC/MS/MS	High Performance Liquid Chromatography Tandem Mass Spectrometry
LLOD	Lower Limit of Detection
MAMA	Mismatch Amplification Mutation Assay
MDG	Millennium Development Goals
MST	Microbial Source Tracking
mtDNA	Mitochondrial Deoxyribonucleic Acids
NCBI	National Centre for Biotechnology Information
NHMRC	National Health and Medical Research Council
NRMMC	Natural Resource Management Ministerial Council
NPV	Negative Predictive Value
PCR	Polymerase Chain Reaction
PMA	Propidium Monoazide
PPV	Positive Predictive Value
qPCR	Quantitative or Real-Time Polymerase Chain Reaction
r^2	Correlation Coefficient
R2A	Reasoner's 2 Agar
RWH	Rainwater Harvesting
SABS	South African Bureau of Standards
SANS	South African National Standards
SDG	Sustainable Development Goals
SLT	Shiga-like Toxin
SPE	Solid Phase Extraction
TBE	Tris Borate Ethylene-diaminetetraacetic Acid
TE	Tris Ethylene-diaminetetraacetic Acid
TN	True Negative
TP	True Positive
UK	United Kingdom
UN	United Nations
UNICEF	United Nations International Children's Emergency Fund
USA	United States of America
US EPA	United States Environmental Protection Agency
UV	Ultraviolet
VBNC	Viable But Non-Culturable
WHO	World Health Organization
WRC	Water Research Commission

Chapter 1: Literature Review

(UK spelling is employed)

1.1 Introduction

The Millennium Development Goals (MDG) were approved by 193 United Nations (UN) member states and 23 international organisations. The overall goals were to: eradicate extreme poverty and hunger; achieve universal primary education; promote gender equality and empower women; reduce child mortality; improve maternal health; combat disease; ensure environmental sustainability and develop a global partnership for development and progress particularly for underdeveloped countries (UN, 2015a). One of the main aims of the MDG was to notably decrease the proportion of people without access to potable water and adequate sanitation by 2015. The global goal for drinking water was achieved by 2010, five years ahead of schedule however, the goal for sanitation was not met [World Health Organisation/United Nations International Children's Emergency Fund (WHO/UNICEF), 2015]. It is estimated that globally, 663 million people still lack access to a safe water source and 2.4 billion people lack access to adequate sanitation facilities (WHO/UNICEF, 2015). Moreover, sub-Saharan Africa did not meet the MDG for potable water by 2015 and it is estimated that 391 million people in this region are still without access to a safe drinking water source (WHO/UNICEF, 2015).

In December 2015 the UN member states, including South Africa, adopted the Sustainable Development Goals (SDG) which came into effect in January 2016. The SDG aim to continue the efforts and plans set in motion by the MDG and the targets for water and sanitation are consequently to: achieve universal and equitable access to safe and affordable drinking water for all by 2030; achieve access to adequate and equitable sanitation and hygiene for all by 2030; improve water quality by reducing pollution, eliminating dumping and minimising the release of hazardous chemicals and materials; halve the proportion of untreated wastewater and substantially increase recycling and safe water reuse globally by 2030; protect and restore water-related ecosystems by 2020; expand international co-operation and support to developing countries in water- and sanitation-related programmes including water harvesting, desalination, water efficiency, wastewater treatment, recycling and reuse technologies by 2030; support and strengthen the participation of local communities in improving water and sanitation management strategies (UN, 2015b).

In line with these goals and to subsequently alleviate the pressure on existing freshwater sources and potable water supply systems, strategies which include the use of rainwater as an alternative freshwater source are being investigated and implemented. Worldwide, rainwater is harvested to augment freshwater supplies and in some countries such as Australia, harvested rainwater is frequently utilised as the primary freshwater source, particularly in households located in regions where water is scarce (Ahmed et al. 2010a; 2011a). Countries such as Ireland (Li et al. 2010), Bermuda (Levesque et al. 2008), United States of America (Jones & Hunt, 2010; Steffen et al. 2013) and South Africa (Mwenge Kahinda et al. 2010) amongst

others, are thus all investigating and implementing rainwater harvesting systems to augment freshwater supplies.

Domestic rainwater harvesting (DRWH) refers to the collection of rainwater from a catchment surface into a storage tank (Mwenge Kahinda & Taigbenu, 2011). This harvested rainwater is frequently used for potable and non-potable purposes particularly in regions where people lack access to safe drinking water and basic sanitation (Mwenge Kahinda et al. 2007). The quality of harvested rainwater is however a major concern particularly where this water source is utilised to augment drinking water supplies. Previous studies have detected the presence of various pathogens in harvested rainwater, which include virulent *Escherichia coli* (*E. coli*) (Dobrowsky et al. 2014a), *Aeromonas* spp. (Simmons et al. 2008; Dobrowsky et al. 2014b), *Salmonella* spp. (Simmons et al. 2008; Uba & Aghogho, 2000; Ahmed et al. 2008a; 2010a; 2012; Dobrowsky et al. 2014b), *Legionella* spp. (Albrechtsen, 2002; Ahmed et al. 2008a; 2010a; Dobrowsky et al. 2014b), *Campylobacter* spp. (Ahmed et al. 2008a; 2010a; 2012) and *Cryptosporidium* spp. (Crabtree et al. 1996; Albrechtsen, 2002).

Indicator organisms are commonly used to monitor water quality and have also been utilised to monitor the quality of harvested rainwater (Dobrowsky et al. 2014c). This may be attributed to the fact that indicator organisms occur abundantly in faecal matter and wastewater, are generally associated with low pathogenicity and therefore are safe and easy to work with and indicator organisms may display relationships with pathogens in contaminated water sources [Department of Water Affairs and Forestry (DWAF) 1996; Harwood et al. 2014]. Therefore, indicator organisms have served as surrogates for the presence of pathogens in contaminated water sources [including harvested rainwater (Dobrowsky et al. 2014c)]. Indicator organisms may include total coliforms, *E. coli*, enterococci, faecal coliforms, *Clostridium perfringens* and heterotrophic bacteria (DWAF, 1996; Harwood et al. 2014). A subset of the indicator organisms, the faecal indicator bacteria (FIB), are then utilised to specifically assess the presence of faecal contamination in a water source and generally includes analysing for *E. coli*, enterococci and faecal coliforms (Harwood et al. 2014). Despite the benefits of monitoring water sources for the presence of indicator organisms, some disadvantages have been noted in literature and is now commonly referred to as the indicator paradigm (Field & Samadpour, 2007). To elucidate, in previous studies the presence of indicator organisms could not always be correlated with the presence of pathogens (Harwood et al. 2005; Harwood et al. 2014). In addition, studies have reported the persistence and proliferation of indicator organisms and in particular FIB strains, that have adapted to the natural environment in various habitats (Anderson et al. 2005; Harwood et al. 2014). Moreover, indicator organisms cannot be utilised to identify the source of contamination as they are present in a wide range of hosts and are therefore not host-specific. It is therefore apparent that the health risk associated with the use of contaminated water cannot

always be accurately assessed by using established indicator detection methods. The remediation of a particular water source thus becomes complex (Harwood et al. 2014).

Source tracking which may be defined as an investigation plan utilising host-specific markers to identify sources of contamination threatening water quality, has the potential to resolve some of the pitfalls associated with the use of FIB to indicate faecal pollution of water sources (Harwood et al. 2014). Microbial source tracking (MST) refers to the utilisation of microbial host-specific markers which may include organisms or genes of organisms generally associated with a specific animal or human host to screen for faecal contamination originating from these hosts. The chemical host-specific markers employed in chemical source tracking (CST) strategies are chemical compounds associated with waste from specific animal or human sources. The use of genetic markers or chemical compounds associated with faecal matter or waste from a known host to screen for host-specific contamination is now considered a more accurate method of determining the primary source of contamination and could thus be utilised to monitor water quality. In addition, employing a set of ST markers could increase confidence when identifying contamination sources, improve discrimination between recent and prior contamination events and aid in accurately assessing the health risk associated with the use of a particular water source (Sidhu et al. 2013). Source tracking has thus been employed in various studies to determine the origin of pollution in seawater (Muscillo et al. 2008; Ahmed et al. 2010b), rivers (Seurinck et al. 2005; Ahmed et al. 2010c; Kobayashi et al. 2013), lakes (Jones-Lepp, 2006), stormwater run-off (Sidhu et al. 2013) and harvested rainwater (Ahmed et al. 2016; Waso et al. 2016). Some of the common microbial and chemical markers employed in source tracking studies include human-specific *Bacteroides* HF183 and *Methanobrevibacter smithii nifH* (*M. smithii nif*H) markers (Seurinck et al. 2005; Ufnar et al. 2006; Sercu et al. 2011; Sidhu et al. 2013; Waso et al. 2016), human adenovirus and human polyomavirus (Muscillo et al. 2008; Sauer et al. 2011; Sidhu et al. 2013; Waso et al. 2016), pharmaceuticals such as paracetamol and aspirin (Hagedorn & Weisberg, 2009; Sidhu et al. 2013; Waso et al. 2016), sterols/stanols (metabolic by-products of cholesterol) such as coprostanol, optical brighteners found in detergents and caffeine (Hagedorn & Weisberg, 2009; Waso et al. 2016).

The aim of the current study was thus to identify a toolbox of MST and CST markers present in DRWH systems, which may be utilised to augment or supplement indicator organism analysis in future screenings of rainwater harvesting systems. In the current study, this was achieved by: i) screening tank water and gutter debris samples for a range of MST and CST markers shown elsewhere to be promising candidates for ST, ii) monitoring indicator numbers in tank water samples, iii) optimising and applying quantitative Polymerase Chain Reaction (qPCR) assays for the assessment of the predominant MST markers, *E. coli* and enterococci in tank water and rooftop debris samples, iv) performing correlation analysis for the MST markers, CST markers and indicator numbers detected in tank water and gutter debris samples, v) designing and

optimising a novel PCR assay to detect avian faecal contamination in gutter debris and tank rainwater samples and vi) validating the accuracy of a novel avian ST marker developed during the course of this study by screening host and non-host faecal samples.

1.2 Rainwater Harvesting

Domestic rainwater harvesting (DRWH) refers to the collection of rainwater from rooftops, courtyards or compacted surfaces into holding tanks above or below the ground for domestic and agricultural use (Gould & Nissen-Peterson, 1999; Mwenge Kahinda & Taigbenu, 2011). As one millimetre of rainwater collected per square metre of collection surface is equivalent to one litre of harvested water [Food and Agriculture Organisation (FAO), 1985], this water source has been earmarked as an effective means to increase the volume of freshwater available for potable and non-potable use in rural communities and urban informal settlements in South Africa (Mwenge Kahinda & Taigbenu, 2011). By 2010, South Africa had approximately 34 000 DRWH tanks dispersed across the country and 96% of these were located in rural areas, particularly the Eastern Cape and KwaZulu-Natal (Mwenge Kahinda et al. 2010). This number has now increased to approximately 69 746 DRWH tanks located across South Africa, providing a primary supply of freshwater to households (**Fig. 1.1.**) (Malema et al. 2016).

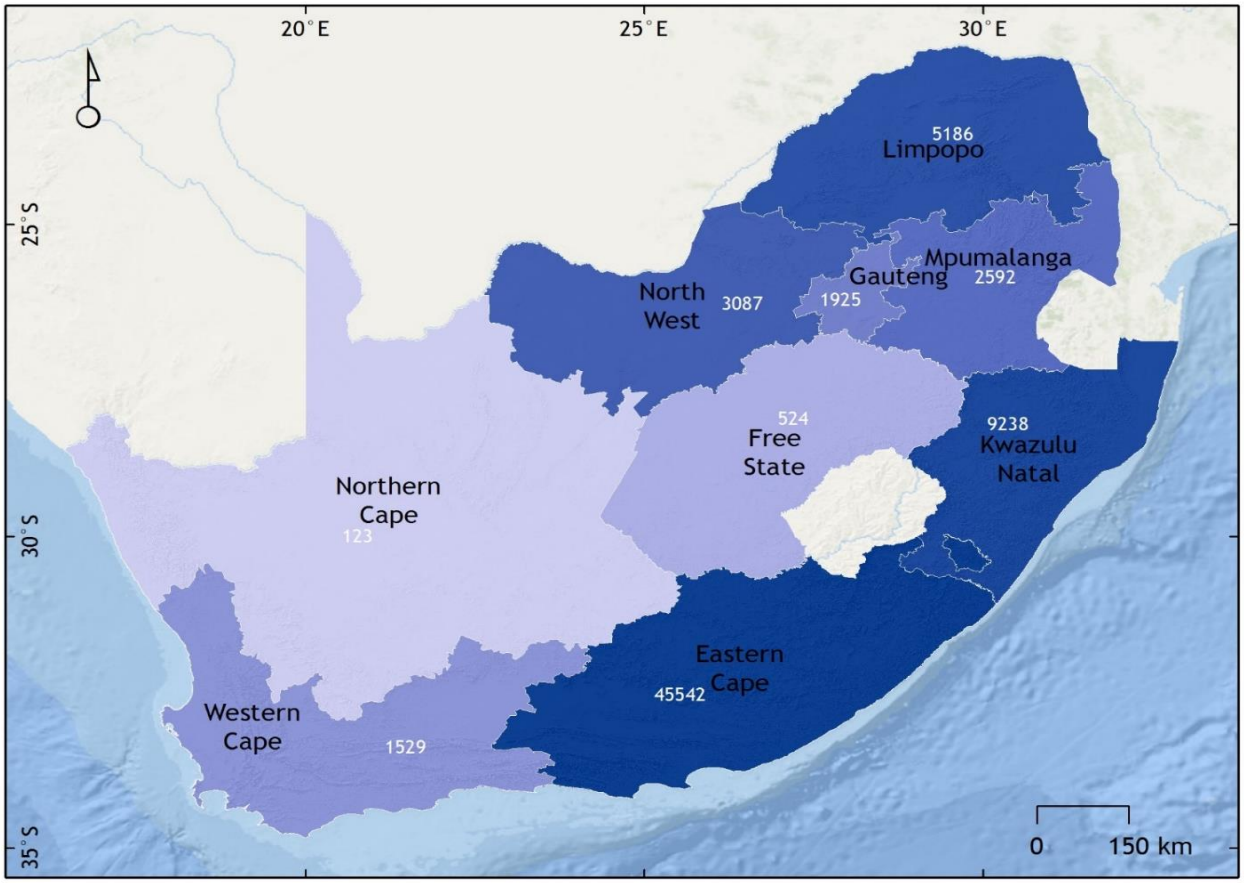


Fig. 1.1. Number of households using DRWH tanks as the primary water source in the nine provinces of South Africa (Adopted from Malema et al. 2016).

In order to successfully implement rainwater harvesting technologies, it is important to qualitatively and quantitatively assess the quality of the rainwater and if required, also to implement treatment strategies to ensure that the rainwater is safe to drink. While legislation regarding the use of harvested rainwater is not available internationally there is an ongoing initiative implemented by the Department of Water and Sanitation and the Water Research Commission of South Africa to stipulate rainwater quality guidelines [Water Research Commission (WRC) Reference Group Meeting, 2015, personal communication]. The quality of the rainwater then ultimately relies on several factors including human activity in close proximity to the tanks, maintenance and topography of the tanks and the type of catchment area (Mwenge Kahinda et al. 2007).

1.2.1 Catchment Systems

In order to collect rainwater a variety of catchment systems such as roof, rock and ground-catchment systems are used and these have all been investigated (Gould & Nissen-Peterson, 1999). Roof and rock-catchments are generally utilised for domestic rainwater harvesting to augment domestic water supplies. The most common method is the roof-catchment system (Gould & Nissen-Peterson, 1999), which typically consists of three basic components, namely the catchment surface, the conveyance or gutter system and the storage tank (Gould & Nissen-Peterson, 1999; Mwenge Kahinda & Taigbenu, 2011).

In developing regions such as Africa and parts of Asia, clay tiles, aluminium, galvanized metal sheets, concrete, plastic, grass thatch and asbestos are the most frequently utilised roofing materials (Gould & Nissen-Peterson, 1999; Farreny et al. 2011; Mwenge Kahinda & Taigbenu, 2011). Smooth materials including galvanized iron sheets, plastics and tiles are ideal materials for the construction of the catchment surfaces as limited accumulation of debris is associated with these surfaces (Gould & Nissen-Peterson, 1999). Materials with irregular surfaces may also be used to construct catchment surfaces provided they are cleaned regularly in order to minimise debris from accumulating on these surfaces and subsequently being washed into the storage tank during a rainfall event. Generally, non-painted materials are preferred for the construction of catchment surfaces as paint flakes or chemical compounds in paint contribute to the contamination of rainwater harvested from these catchments (Gould & Nissen-Peterson, 1999). Downpipes constructed from metal or plastic then convey the rainwater from the catchment surface to the storage tanks. The latter may be constructed from cement, concrete, brick or polymeric materials as these are usually watertight, durable and cost-effective materials (Li et al. 2010).

The most important requirement of the catchment area is that it should not contaminate the rainwater (Gould & Nissen-Peterson, 1999). To further preserve the quality of rainwater it has

been suggested that catchment areas and the interior of the rainwater harvesting tanks should be cleaned regularly (Mwenge Kahinda et al. 2007). In addition, first-flush diverters could be installed to divert the first few millilitres of rainwater after a dry spell, as this may eliminate some of the debris and thus contaminants that have accumulated on the roof surface during the dry period. Alternatively, leaf screens and/or fine filters may be installed to prevent debris from the rooftops washing into the tanks (Martinson & Thomas, 2005; Mwenge Kahinda et al. 2007; De Kwaadsteniet et al. 2013). However, research has shown that these efforts rarely improve the microbial quality of the harvested rainwater and may only improve the physico-chemical quality of the water source (Gikas & Tsihrintzis, 2012).

1.3 Primary Chemical and Microbial Contaminants Associated with Roof-Catchment Systems

Rainwater quality is commonly compromised as raindrops traverse polluted air, by contaminated catchment areas and by contaminated storage tanks (De Kwaadsteniet et al. 2013). In addition, factors that may also influence the quality of roof-harvested rainwater include the roof geometry, the roof material, the proximity of the roof to pollution sources, maintenance of the roof, rainfall events, seasons, wind direction and speed, dry periods and the presence of contaminants in the atmosphere (Abbasi & Abbasi, 2011). These factors influence both the chemical and the microbial quality of harvested rainwater.

Chemical contaminants of harvested rainwater are less studied than are microbial contaminants, as chemical pollutants do not pose an immediate health risk to the consumer (De Kwaadsteniet et al. 2013). Various cations and anions, including iron, copper, calcium, potassium, magnesium, sodium, ammonium, zinc, fluoride, phosphate, nitrate, chlorine, phosphorous and sulphate have been detected in harvested rainwater. However, most of the concentrations were within limits set by national and international drinking water standards (De Kwaadsteniet et al. 2013). In contrast, previous studies have detected lead in harvested rainwater at concentrations exceeding the drinking water guidelines of various countries (Simmons et al. 2001; Huston et al. 2012; De Kwaadsteniet et al. 2013). This water had been collected from rooftops painted with lead-based paints (Abbasi & Abbasi, 2011). Thus, it is generally recommended that such paints are not applied to the catchment area. In contrast, Uyger et al. (2010) detected high levels of aluminium, as well as the trace elements chromium, cobalt, nickel, vanadium and lead in harvested rainwater. It was concluded that these contaminants washed into the tanks by means of raindrops as they traversed polluted air. Therefore, air quality in the vicinity of a rainwater harvesting system could also influence the quality of the harvested rainwater. Hence in areas experiencing high levels of air pollution, possible treatment and preventative strategies should be investigated (Uyger et al. 2010).

Numerous studies have identified faecal matter as the major source of microbial contamination in harvested rainwater (Field & Samadpour, 2007; Ahmed et al. 2008a; Simmons et al. 2008). Possible sources of faecal contamination in stored rainwater include birds and small mammals such as rats as well as insects and reptiles, which have access to the rooftops utilised as catchment surfaces (Mwenge Kahinda et al. 2007; De Kwaadsteniet et al. 2013). Thus, undesirable pathogens present in the faecal matter of these animals, insects and reptiles may be washed into the tanks (De Kwaadsteniet et al. 2013). Pathogens detected in harvested rainwater include virulent *E. coli* (Dobrowsky et al. 2014a), *Aeromonas* spp. (Simmons et al. 2008; Dobrowsky et al. 2014b), *Salmonella* spp. (Simmons et al. 2008; Uba & Aghogho, 2000; Ahmed et al. 2008a; 2010a; 2012; Dobrowsky et al. 2014b), *Pseudomonas* spp. (Uba & Aghogho, 2000; Albrechtsen, 2002; Dobrowsky et al. 2014b), *Shigella* spp. (Uba & Aghogho, 2000), *Legionella* spp. (Albrechtsen, 2002; Ahmed et al. 2008a; 2010a; Dobrowsky et al. 2014b), *Campylobacter* spp. (Ahmed et al. 2008a; 2010a; 2012) and *Cryptosporidium* spp. (Crabtree et al. 1996; Albrechtsen, 2002). Furthermore, indicator organisms which include total coliforms, faecal coliforms, *E. coli* and enterococci have been detected in harvested rainwater in several countries including Australia (Verrinder & Keleher, 2001; Ahmed et al. 2008a; 2010a; 2012), Canada (Despins et al. 2009), US Virgin Islands (Crabtree et al. 1996) and South Africa (Dobrowsky et al. 2014c).

It is thus clear that harvested rainwater is not a pure water source and the presence of contaminants (microbial and chemical) pose an immediate and possible long-term health risk to the consumer. Furthermore, the influence of polluted air on the quality of harvested rainwater has not been studied extensively and should be an important consideration particularly when rainwater harvesting systems operate in urban areas where air pollution may be considerable. The current study will however focus on sources of faecal pollution of rainwater harvesting systems and identify markers specific to these sources that could be utilised to supplement FIB analyses during future monitoring procedures implemented for this water source.

1.4 Monitoring Water Quality

Monitoring water sources such as harvested rainwater for pathogens provides valuable information regarding the risks associated with the use and consumption of the specific water source. However, monitoring water sources for all known pathogens is costly and time-consuming. This is largely attributed to the occurrence of rare pathogens in water where their presence in very low numbers makes them difficult to culture and identify. Furthermore, if such pathogens are not evenly distributed within the water source, the scale of the problem increases (DWAF, 1996; Field & Samadpour, 2007; Harwood et al. 2014). Pathogens present in the water source may also be diverse and monitoring for only a few pathogens would provide a false impression regarding the overall quality of the water and the risk associated with the use of the

water (Field & Samadpour, 2007; Harwood et al. 2014). It is thus standard practice to monitor water quality by using indicator organisms which include *E. coli*, enterococci, coliforms, faecal coliforms, *Clostridium perfringens*, various other heterotrophic bacteria and bacteriophages (DWAF, 1996; Field & Samadpour, 2007; Harwood et al. 2014). Faecal indicator bacteria are a subset of the indicator organisms which are used to screen specifically for faecal contamination in a water source and includes analysing for *E. coli*, enterococci and faecal coliforms. The FIB are abundant in faecal matter and wastewater and their presence in environmental waters may thus indicate faecal pollution and the possible presence of waterborne pathogens (DWAF, 1996). Furthermore, because of their abundance in faecal matter, the FIB are easily detected and cultured from contaminated water and are therefore easily monitored in water sources. The FIB may also display relationships with pathogens in a water source (Harwood et al. 2014).

There are however certain limitations to assessing only FIB to monitor water sources for faecal contamination. Numerous studies have indicated that the presence of indicator organisms, and in particular the FIB, do not necessarily correlate positively with the pathogen content of a water source (Lund, 1996; Bonadonna et al. 2002; Lemarchand & Lebaron, 2003; Anderson et al. 2005; Harwood et al. 2005; 2014). This could in part be attributed to the differences in physiology and phylogeny between FIB and possible pathogens, as pathogens include bacteria, viruses, fungi, yeasts and protozoa, whereas FIB are comprised solely of bacteria (Harwood et al. 2014). Moreover, some environmentally adapted FIB strains have been shown to proliferate and persist after excretion from a host in many different habitats ranging from terrestrial soils and aquatic sediments to aquatic vegetation (Harwood et al. 2014). These factors negatively influence the reliability of FIB as indicators of faecal pollution in an environment. In addition, FIB do not display host-specific distributions and may be found ubiquitously associated with a wide variety of warm- and cold-blooded animals (Field & Samadpour, 2007; Harwood et al. 2014). Finally, well-characterised culture techniques are commonly used for the detection of indicator organisms and FIB. These methods introduce a bias towards the detection of viable and culturable organisms and exclude viable but non-culturable organisms. As a result, monitoring the quality and assessing the health risk associated with the use of a particular water source is complex. It is clearly apparent that supplementary indicators of faecal pollution in water sources are required to improve the accuracy and reliability of water quality monitoring strategies (Harwood et al. 2014).

1.5 Source Tracking

Source tracking may be described as both a collection of methods and an investigative strategy to identify possible sources of pollution in environmental waters (Harwood et al. 2014). Source tracking has the potential to resolve some of the pitfalls associated with the use of FIB to monitor for faecal pollution in water sources. The method relies on the premise that certain

genes of microorganisms and/or chemical compounds may be associated with the faecal matter of a specific host (animal or human). These markers may be screened for in a water body to link pollution of the water to a specific host source (Field & Samadpour, 2007; Harwood et al. 2014). Numerous methods for source tracking have been developed over the past few years to identify host-specific faecal contamination in environmental waters (Simpson et al. 2002; Meays et al. 2004; Field & Samadpour, 2007; Stoeckel & Harwood, 2007; Harwood et al. 2014; Villemur et al. 2015). These methods may be divided into two categories. One is MST where organisms or genes of organisms are screened for by using molecular methods such as the PCR technique. The other is CST where compounds associated with faecal matter or other waste (for example household waste generated by humans) originating from specific hosts, are screened for in environmental waters (Field & Samadpour, 2007; Harwood et al. 2014).

1.5.1 Microbial Source Tracking

Microbial source tracking is usually the primary focus of source tracking strategies and relies on the premise that certain microorganisms are specific to certain hosts. The molecular markers which include specific DNA sequences or genes, are traced in the environment (Harwood et al. 2014) by utilising molecular techniques such as the PCR and qPCR. Characteristics of an ideal MST marker include: the marker should be specific to the target host-group; must be present in all members of the target host group; must be temporally and geographically stable in the host group and the decay rates of the markers should correlate with the decay rates of the pathogens present in a water source (Ahmed et al. 2015). Microbial source tracking may in turn be divided into library-dependent and library-independent strategies.

Library-dependent methods consist of phenotypic and genotypic tests including antibiotic resistance assays, carbon-source utilisation profiling, ribotyping/DNA fingerprinting and screening for the *uid* genes associated with *E. coli* (Field & Samadpour, 2007). The basis of this approach is to construct a library or host origin database from known hosts/sources. Using the database, any new isolates identified from the chosen sampling site are then compared with known isolates. In turn this should identify the most likely source of contamination (Field & Samadpour, 2007). However, this approach has proven to be time-consuming and expensive and requires advanced statistical analysis of results obtained to confirm findings (Field & Samadpour, 2007; Ahmed et al. 2015). As a result, library-dependent methods have largely been replaced by library-independent methods.

Library-independent methods emerged with the development of molecular techniques and technologies and rely on the detection of genes of organisms by the use of PCR assays. This allows for the rapid detection of molecular markers, rare microorganisms and viable but non-culturable organisms (Field & Samadpour, 2007; Harwood et al. 2014). Various markers have

been proposed as promising MST markers and these include human-specific *Bacteroides* HF183, *M. smithii nif*H, *Bifidobacterium* spp. and enteric viruses (Field & Samadpour, 2007; Harwood et al. 2014), amongst others. In addition, microorganisms commonly investigated as potential MST markers include faecal anaerobic bacteria which were previously not screened for when assessing water quality as they are difficult to culture. With the development of advanced molecular techniques, the presence of these microbes is now used as a potential indicator of faecal contamination in environmental waters as these markers are thought to have co-evolved with their specific hosts and therefore they could display notable host-specific distributions (Johnston et al. 2013).

1.5.1.1 *Bacteroides* spp.

Bacteroides spp. are Gram-negative anaerobic organisms present in the digestive tract of warm-blooded animals and humans and often occur in a higher abundance than traditional faecal coliforms (Kildare et al. 2007). These organisms have been proposed as promising markers for source tracking and were among the first MST markers to be developed, as they are present in high concentrations in the faecal matter of hosts and display highly host-specific distributions (Harwood et al. 2014). In addition, because of their anaerobic physiology it is believed that these bacteria do not persist for extended periods of time in a natural aerobic environment. This characteristic is beneficial when recent contamination sources need to be identified (Ballestè & Blanch, 2011).

Primers specific for the identification of *Bacteroides* spp. have been developed to detect faecal contamination originating from and specific to ruminants, pigs and humans, amongst others (Bernhard & Field, 2000a; Layton et al. 2006; Okabe et al. 2007; Field & Samadpour, 2007). The most well-known and extensively studied *Bacteroides* marker, which is specific for faecal pollution of human origin, is the HF183 marker (**Table 1.1**). The HF183 primer set is complementary to a specific segment of the 16S rRNA gene of *Bacteroides* spp. This segment is conserved among *Bacteroides* strains of human origin and has been shown to be highly specific for the detection of sewage and human faecal material in environmental waters (Harwood et al. 2014). In a study conducted by Ahmed et al. (2010c) the sensitivity (proportion of target host samples identified as positive) and specificity (proportion of non-target hosts that produce negative results) of the HF183 marker was assessed by screening for its presence in human, cattle, dog, cat and chicken faecal samples and the marker was subsequently utilised to determine the quality of the water from an urban lake in Dhaka, Bangladesh. The HF183 marker was detected in 13 of 15 human samples and in none of the animal faecal samples with the exception of one cat and one dog sample. The sensitivity of the marker was found to be 87% and the specificity was calculated to be 93%. The relatively high specificity of this marker is valuable for the determination of the source of faecal pollution (Ahmed et al. 2010c).

Furthermore, it is hypothesised that the detection of the human-associated HF183 marker in companion animals such as cats and dogs may be ascribed to gut microorganisms being transferred between hosts that live in close proximity to one another (Field & Samadpour, 2007). Cross-reactivity among host species living in close proximity should thus be considered when screening a specific site for host-associated markers. The specificity and sensitivity of a ruminant-specific *Bacteroides* marker, CF128 (**Table 1.1**) was also investigated by Ahmed et al. (2010c). The CF128 marker was found to be 100% specific and the sensitivity was calculated to be 75%. Although the sensitivity of the marker was less than that observed for the human-specific *Bacteroides* HF183 marker, the sensitivity is still regarded as high. Hence, the use of *Bacteroides* spp. as a faecal source tracking marker has been widely accepted as these host-specific markers are able to distinguish between host and non-host sources of faecal pollution with relatively high sensitivity and specificity percentages (Ahmed et al. 2010c).

The human-associated HF183 marker has also been developed as a qPCR assay by Seurinck et al. (2005). The limit of detection of the marker was found to be 4.7×10^5 human-specific HF183 *Bacteroides* genetic markers per litre of freshwater and the qPCR assay was more sensitive than conventional PCR assays. It was shown that five out of six human faecal samples tested positive for the HF183 marker by qPCR whereas four out of six human faecal samples tested positive by conventional PCR assays (Seurinck et al. 2005). The qPCR assay thus allows for accurate, rapid detection and quantification of the HF183 marker in water samples (Seurinck et al. 2005). *Bacteroides* HF183 analysis has subsequently been applied to various water sources utilising both conventional PCR and qPCR assays to indicate faecal contamination as well as to distinguish between human and animal faecal contamination sources (**Table 1.1**). For example, Sidhu et al. (2013) utilised the *Bacteroides* HF183 marker to indicate sewage contamination in stormwater run-off. In addition, Waso et al. (2016) detected the HF183 marker in harvested rainwater samples and gutter debris collected from DRWH systems.

Table 1.1 Microbial source tracking markers applied to various water sources utilising conventional PCR and qPCR assays.

Organism	Marker (Specific Host)	Gene Target	Water Source	Reference
Bacteroides	HF183 (Human)	16S rRNA	Wastewater, stormwater run-off, freshwater, seawater, river water, surface water, harvested rainwater	Seurinck et al. 2005; Ahmed et al. 2009; Jenkins et al. 2009; Ahmed et al. 2010c; Gourmelon et al. 2010; Shanks et al. 2010; Sauer et al. 2011; McQuaig et al. 2012; Sidhu et al. 2013; Waso et al. 2016
	HuBac (Human)		Surface water, wastewater	Layton et al. 2006; Ahmed et al. 2009; Shanks et al. 2010
	BacHum-UCD (Human)		Wastewater	Kildare et al. 2007; Ahmed et al. 2009; Jenkins et al. 2009
	BacH (Human)		Wastewater	Reischer et al. 2007; Ahmed et al. 2009
	Human-Bac1 (Human)		River water	Okabe et al. 2007
	HumM2 (Human)	Hypothetical protein B3236	Wastewater	Shanks et al. 2010
	HumM3 (Human)	Putative RNA polymerase sigma factor	Wastewater	Shanks et al. 2010
	B. theta α (Human)	*B. thetaiotomicron* α-mannanase	Wastewater	Yampara-Iquise et al. 2008
	CF128 (Bovine)	16S rRNA	Surface water	Ahmed et al. 2010c
	AllBac (All *Bacteroides* spp.)		Surface water, wastewater, river water	Layton et al. 2006; Gourmelon et al. 2010

Table 1.1 (Continued) Microbial source tracking markers applied to various water sources utilising conventional PCR and qPCR assays.

Organism	Marker (Specific Host)	Gene Target	Water Source	Reference
Bacteroides	GenBac (All *Bacteroides* spp.)	16S rRNA	Surface water, freshwater	Bernhard & Field, 2000b; Sauer et al. 2011
	Rum-2-Bac (Bovine spp.) Pig-1-Bac (Porcine spp.) Pig-2-Bac (Porcine spp.)		Wastewater, river water	Gourmelon et al. 2010
	BacPre1 (General *Bacteroides-Prevotella*) Cow-Bac2 (Bovine spp.) Pig-Bac2 (Porcine spp.)		River water	Okabe et al. 2007
	BoBac (Bovine spp.)		Surface water	Layton et al. 2006
Bifidobacterium spp.	*Bifidobacterium* (Human)	16S rRNA	Wastewater, surface water	Bernhard & Field, 2000b; Gourmelon et al. 2010
Lachnospiraceae	Lachno2 (Human)	16S rRNA V6 region	Wastewater, harbour water, freshwater	Newton et al. 2011
Enterococcus	*Esp*-1 (Human)	Enterococcal surface protein (*esp*-1)	Wastewater, septic tank waste	Scott et al. 2005
	E. faecium esp (Human)			Ahmed et al. 2008b
Methanobrevibacter smithii	*nif*H (Human)	*nif*H	Wastewater, seawater, surface water, stormwater run-off	Johnston et al. 2010; McQuaig et al. 2012; Sidhu et al. 2013
F^+ RNA Coliphages	FRNAPH I / IV (Animal) FRNAPH II / III (Human)	Viral genome	Wastewater, river water	Wolf et al. 2010; Gourmelon et al. 2010
Pepper Mild Mottle Virus	PMMoV (Human)	Viral genome	Wastewater	Rosario et al. 2009
Teschovirus	PTV (Porcine)	Polyprotein	Wastewater	Jimenez-Clavero et al. 2003

Table 1.1 (Continued) Microbial source tracking markers applied to various water sources utilising conventional PCR and qPCR assays.

Organism	Marker (Specific Host)	Gene Target	Water Source	Reference
Polyomavirus	HPyV (Human)	T antigen	Wastewater, seawater, river water, stormwater run-off	Albinana-Gimenez et al. 2009; McQuaig et al. 2009; Ahmed et al. 2010b; McQuaig et al. 2012; Sidhu et al. 2013
	BPyV (Bovine spp.)	VP1	Wastewater, river water	Hundesa et al. 2006; 2010
Adenovirus	AdV (General)	Hexon gene	Harvested Rainwater	Waso et al. 2016
	HAdV (Human)		Wastewater, River Water, Seawater, Stormwater run-off, Harvested Rainwater	Noble et al. 2003; Hundesa et al. 2006; Ahmed et al. 2010b; McQuaig et al. 2012; Sidhu et al. 2013; Waso et al. 2016
	HAdV-C (Human)		Wastewater	Wolf et al. 2010
	HAdV-F (Human)			
	BAdV (Bovine spp.)		Wastewater	Ahmed et al. 2010b
	BAdV (Bovine spp.)		Wastewater	Hundesa et al. 2006
	PAdV (Porcine spp.)		Wastewater and River Water	
	PAdV (Porcine spp.)		Wastewater and River Water	Wolf et al. 2010
	OAdV (Ovine spp.)			
Enterovirus	HEV (Human)	NTR	Wastewater	Noble et al. 2003
Norovirus	NoVGI (Human)	Capsid protein	Wastewater	Wolf et al. 2010
	NoVGII (Human, Porcine)			
	NoVGIII (Bovine, Ovine)	RNA polymerase		
Atadenovirus	AtAdV (Sheep, Cattle, Deer, Goat)	Hexon gene	Wastewater	Wolf et al. 2010
Mitochondrial DNA	Human, Bovine, Porcine, Dog, Cat, Canada Goose, Deer	NADH dehydrogenase subunit 5	Wastewater	Caldwell et al. 2007; Caldwell & Levine, 2009
	mtCytb	Cytochrome b	Wastewater	Schill & Mathes, 2008

Another human-specific *Bacteroides* marker, BacHum-UCD, was developed by Kildare et al. (2007). This marker together with bovine-specific (BacCow-UCD), dog-specific (BacCan-UCD) and universal (BacUni-UCD) *Bacteroides* markers were designed and the specificity and sensitivity of the markers were compared (**Table 1.1**). The universal primer set (BacUni-UCD) detected *Bacteroides* in all faecal samples tested and displayed 100% sensitivity and 100% specificity for faecal matter from all hosts tested. The human-specific marker (BacHum-UCD) was detected in all mixed human faecal matter samples such as sewage and exhibited 100% sensitivity towards mixed sources of human faecal matter. However, the marker was not detected in all individual human faecal samples (12/18) and was also detected in faecal matter originating from dogs (non-host faecal sources). Therefore, the marker was calculated to be only 87% specific to individual human faecal samples. Once again, the detection of the *Bacteroides* human marker in dog faecal material was attributed to cross-reactivity between host groups living in close proximity to one another. Additionally, the sensitivity of the bovine-specific marker was 100% and the specificity was 62% as cross-reactivity with horse faecal matter was observed. The lowest sensitivity (63%) and specificity (57%) was obtained when the canine-specific marker was utilised. Kildare et al. (2007) also calculated the conditional probability of the markers to identify the correct host and calculated this to be between 0.84 and 1.00 (with the maximum probability score being equal to 1.00) for all the *Bacteroides* markers tested. Moreover, Kildare et al. (2007) tested these markers in the field and found these results to be highly accurate, reproducible and the markers were markedly host-specific. The Kildare et al. (2007) study however highlights some disadvantages in utilising *Bacteroides* spp. as host-specific markers. These were attributed to the fact that some markers were detected in the faecal matter of non-target hosts, as was observed for the HF183 marker (Ahmed et al. 2010b). An example is the human marker (BacHum-UCD) which was detected in the faeces of companion animals (cats and dogs). Thus, cross-reactivity of *Bacteroides* spp. from different hosts may lead to false-positive results and it is suggested that new markers be developed in order to minimise cross-reactivity among hosts living in close proximity (Kildare et al. 2007). In addition, the study emphasised a lack of markers specific to seagull guano, companion animals and wildlife animals and therefore markers specific to these host groups should be investigated and developed (Kildare et al. 2007).

1.5.1.2 *Bifidobacterium* spp.

Bifidobacterium spp. are anaerobic Gram-positive bacteria which form part of the phylum Actinobacteria. They are known to colonise the gut of healthy infants and provide many benefits to the hosts. These include enhancement of the immune system and the production of vitamins and antimicrobial compounds (Ballestè & Blanch, 2011). *Bifidobacterium* spp. have been described as being highly host-specific (which may be attributed to their anaerobic physiology) and have accordingly been considered as promising source tracking markers for the detection

of human and animal faecal contamination in the environment (Ballestè & Blanch, 2011). Species that are associated with human hosts include *Bifidobacterium adolescentis* (*B. adolescentis*), *Bifidobacterium angulatum*, *Bifidobacterium catelatum* and *Bifidobacterium longum* (Ballestè & Blanch, 2011). Some species such as *Bifidobacterium pseudolongum* and *Bifidobacterium thermophilum* have also been described as bovine-specific as these organisms are widespread in cattle (Ballestè & Blanch, 2011). In addition, *Bifidobacterium pullorum* and *Bifidobacterium gallinarum* have been shown to display host-specificity in chickens (Ballestè & Blanch, 2011).

In a study conducted by Gourmelon et al. (2010), *B. adolescentis* was utilised as a MST marker for the detection of human waste and was detected in all wastewater treatment plant (WWTP) effluent samples and in 90% of individual human faecal samples (**Table 1.1**). The sensitivity of the marker was calculated to be 92%. The marker also occurred in two bovine and two avian faecal samples (non-target hosts) and therefore displayed a specificity of 94.5%. It was concluded by the authors that *B. adolescentis* could serve as a promising marker for discriminating between human and non-human sources of faecal pollution when combined with other markers as the marker did not display 100% specificity for human faecal matter (Gourmelon et al. 2010).

1.5.1.3 Enterococcal Surface Protein

Enterococci are facultative anaerobic Gram-positive cocci and are recognised as important human pathogens as they can cause nosocomial bacteraemia, endocarditis, neonatal and urinary tract infections (Eaton & Gasson, 2002). Enterococci are also commensal organisms commonly found in the oral cavity, gastrointestinal tract and female genital tract of both humans and animals (Mohamed & Huang, 2007). *Enterococcus faecalis* (*E. faecalis*) is responsible for 80 to 90% of human enterococcal infections and is the most common enterococcal species isolated from clinical infections. *Enterococcus faecium* (*E. faecium*) is less common than *E. faecalis* but accounts for the remaining infections caused by enterococci in humans (Jett et al. 1994, Jones et al. 2004). For many years, enterococci were considered medically insignificant and harmless to humans. However, due to advances in medical research, enterococci are presently known to be one of the leading pathogens causing nosocomial infections and are associated with mortality rates as high as 61% (Lopes et al. 2005; Fisher & Phillips, 2009). Over the last decade enterococci have been used in the food industry as a probiotic or starter culture because of their ability to produce bacteriocins (Fisher & Phillips, 2009). Enterococci have also been used as FIB for many years as they are associated with the faecal matter of warm-blooded animals (DWAF, 1996; Field & Samadpour, 2007; Harwood et al. 2014). Some studies have indicated that the presence and distribution of *Enterococcus* spp.

may be dependent on host type and consequently enterococci may be promising targets for source tracking (Scott et al. 2005).

A study by Scott et al. (2005) investigated the possibility of utilising enterococci as a target to develop MST markers, specifically by using the enterococcal surface protein (*esp*) gene of these bacteria to indicate human sewage contamination in a water source (**Table 1.1**). The *esp* gene is considered a putative virulence factor of *Enterococcus* spp. and has been reported to be associated with *E. faecalis* and *E. faecium* strains isolated predominantly from a clinical environment (Shankar et al. 1999; Di Rosa et al. 2006). The ESP protein is encoded for by a chromosomal gene and is found localised on the cell surface of enterococci. It is hypothesised that the *esp* gene product confers virulence by altering the structure of the bacterial cell surface, which possibly contributes to enterococci evading the host immune system. Alternatively, it may enhance the binding of the bacterium to host cells thus improving persistence of enterococci at infection sites (Shankar et al. 1999). In the study conducted by Scott et al. (2005), the marker was found in all wastewater and septic tank samples, but was not detected in any animal faecal samples. Thus it is possible that this marker could be used to differentiate between human and animal sources of faecal contamination. Ahmed et al. (2008b) used the enterococcal *esp* marker to develop a qPCR assay for the quantification of the marker in sewage and environmental waters. The marker was found in all sewage samples and in eight of the 12 septic tank samples monitored (**Table 1.1**). The marker was not detected in any of the animal faecal samples tested in the study (Ahmed et al. 2008b) and the specificity of the *esp* marker was calculated to be 100%. The authors recommended the *esp* marker as a promising source tracking marker to identify human sewage as the contamination source of environmental waters. However, it was further recommended that the marker should be used in combination with enterococcal plate counts (Ahmed et al. 2008b) for optimal water quality monitoring.

1.5.1.4 Lachnospiraceae

The Lachnospiraceae is a robust family of bacteria belonging to the order Clostridiales, and they commonly occur in the gut of humans and other mammals. This family is morphologically diverse with rods, vibrio and cocci being evident (Vos et al. 2009). All known members of this family are obligate anaerobes and are thus unlikely to proliferate in any aerobic environment which occurs outside the anaerobic gastrointestinal tract of mammals (Meehan & Beiko, 2014). The family contains 24 named genera including *Lachnospira* (type genus), *Butyrivibrio*, *Lachnobacterium*, *Pseudobutyrivibrio* and *Roseburia*, amongst others (Vos et al. 2009; Meehan & Beiko, 2014). Recently, research has focussed on butyric acid production by members of the Lachnospiraceae, with a particular interest as to the manner in which the production of butyric acid influences other microorganisms as well as the host epithelial cells in the gastrointestinal tract of mammals (Meehan & Beiko, 2014).

The Lachnospiraceae have been proposed as an indicator of sewage or faecal pollution in environmental waters as members of this family occur abundantly in both wastewater and faecal matter (Newton et al. 2011; McLellan et al. 2013). In a study conducted by Newton et al. (2011) a harbour in Milwaukee, United States of America (USA) was screened for sewage contamination by screening for the presence of Lachnospiraceae. Detection of these microorganisms was compared with the detection of *Bacteroides* source tracking markers. Both Lachnospiraceae and *Bacteroides* spp. were present in all harbour water samples (Newton et al. 2011) (**Table 1.1**). The markers displayed a notable positive correlation and were detected at approximately the same concentrations throughout the sampling period in all of the harbour water samples monitored (Newton et al. 2011). As a result of the correlation noted between *Bacteroides* and Lachnospiraceae, it was suggested by the authors that the markers be utilised together to provide a consistent measure of faecal contamination in environmental waters rather than using either marker independently (Newton et al. 2011).

1.5.1.5 *Methanobrevibacter smithii nif*H

Methanobrevibacter smithii is the dominant archaeon found in the gastrointestinal tract of humans and can occur at concentrations as high as 10^{10} cells per gram of dry faeces (Johnston et al. 2013). In addition, *M. smithii* has been found in the vaginal tract of females (Belay et al. 1990). *Methanobrevibacter smithii* is the only species of *Methanobrevibacter* known to colonise the large intestine of humans. In contrast, other *Methanobrevibacter* spp. colonise the gut of animals, are found in decaying plant materials and are also present in anaerobic sludge and wastewater treatment plants (Bond et al. 1971; Miller & Wolin, 1983; Lin & Miller, 1998). In a study conducted in the United Kingdom (UK) and the USA, it was found that 33% of humans harbour methanogens. For this reason, *M. smithii* may be well-suited to trace composite faecal contamination sources or sewage in the environment, rather than detecting contamination originating from single individuals or hosts (Bond et al. 1971; Johnston et al. 2013).

A study conducted in Southeast Queensland (Ahmed et al. 2011b) evaluated the sensitivity and specificity of the *M. smithii nif*H marker, which targets the DNA sequence for the nitrogenase reductase gene grouped in the Pseudo-*nif* cluster which occurs only in methanogens (Ufnar et al. 2006). This sequence is highly conserved in *M. smithii* strains. In total, 188 faecal samples from 11 host groups were screened for the *nif*H marker and of the 64 human faecal samples tested, 52 (81%) tested positive. All primary influent sewage samples tested positive for the marker and 15 of the 22 secondary effluent samples also tested positive (Ahmed et al. 2011b) (**Table 1.1**). In the case of animal faecal samples, one of the 30 avian faecal samples and six of the 20 porcine faecal samples tested positive for the *nif*H marker (Ahmed et al. 2011b). The specificity of the *nif*H marker was found to be 96% and the sensitivity of the marker was determined to be 81% for the detection of human faecal matter. The relatively high sensitivity

and specificity of this marker highlights the potential for using the *nif*H marker to screen for human faecal contamination in the environment (Ahmed et al. 2011b).

In a follow-up study conducted by Johnston et al. (2013) the *M. smithii nif*H marker was screened for in 485 individual faecal samples from 20 different hosts as well as from sewage samples and samples collected from pit toilets and septic tanks. The *nif*H marker was detected in 100% of the sewage samples (n = 24), 45% of the pit toilet samples (n = 15) and 47% of the septic tank samples (n = 20) (**Table 1.1**). In addition, the marker was found in 78% of the individual human faecal samples. The marker was also detected in all goat faecal samples (n = 2), 46% of cow faecal samples (n = 13) and in one of the sheep faecal samples (n = 2). The marker was not detected in any of the other individual animal faecal samples taken from birds, cats, dogs, geese, seagulls and mice. Neither was the marker detected in any composite animal faecal samples. The *nif*H marker once again displayed high but not absolute sensitivity to the specific target (humans). However, this marker, together with the *Bacteroides* HF183 marker, was detected consistently in human-source faecal samples (Johnston et al. 2013). For this reason, it was concluded that the *nif*H marker combined with the human-specific HF183 marker could be used as toolbox markers for the routine screening of environmental waters for human faecal contamination (Johnston et al. 2013).

1.5.1.6 Enteric Viruses

Enteric viruses are a heterogeneous group of viral agents associated with subclinical infections and diseases in humans and animals (Staggemeier et al. 2015). These viruses have been studied for their potential use as source tracking markers as they may be identified by using molecular methods without culturing and more importantly, some enteric viruses have also demonstrated host-specificity (Field & Samadpour, 2007). The viruses identified with source tracking potential include polyomavirus, enterovirus, teschoviruses and adenovirus (refer to Section 1.5.1.7).

Polyomavirus is the sole genus belonging to the *Polyomaviridae.* The viruses are icosahedral and encapsidated and possess a 5 kb double-stranded DNA genome. Polyomavirus has gained particular interest as the causative agent of kidney nephritis (inflammation) and progressive multifocal leukoencephalopathy (inflammation of the brain) in immunocompromised individuals, Merkel cell carcinoma (a neuroendocrine tumour of the skin) and trichodysplasia spinulosa (McQuaig et al. 2006; La Rosa et al. 2015). Studies have shown that more than 70% of adults harbour antibodies towards polyomavirus and this virus is responsible for lifelong asymptomatic viruria (viruses present in urine) in infected immunocompetent individuals. Polyomavirus is shed primarily in urine, but may also be found in faecal matter and is known to be prevalent in human populations (McQuaig et al. 2006; Harwood et al. 2014). In addition, polyomavirus has been

detected in high numbers in municipal sewage (McQuaig et al. 2006). This makes polyomavirus a promising indicator of human faecal pollution in the environment (McQuaig et al. 2006) and the virus has thus been used as a source tracking marker in environmental waters (McQuaig et al. 2006; McQuaig et al. 2009; Albinana-Gimenez et al. 2009; McQuaig et al. 2012; Sidhu et al. 2013) (**Table 1.1**).

Enteroviruses belong to the order *Picornavirales* and family *Picornaviridae*, and have a positive-sense single-stranded RNA genome. Enteroviruses are associated with a variety of diseases in humans and other mammals. Infection by these viruses results in a wide variety of symptoms ranging from the common cold, hand-, foot- and mouth-disease, myocarditis, aseptic meningitis, acute haemorrhagic conjunctivitis, severe neonatal sepsis-like disease to acute flaccid paralysis [Centres for Disease Control and Prevention (CDC), 2014]. However, enteroviruses may also cause asymptomatic infections. Enterovirus currently has 71 identified serotypes and infected individuals may excrete enterovirus in their faeces. Viral particles can also occur in body fluids such as saliva and nasal mucus. Hence as a result of the prevalence of enterovirus in the human population and the association of enteroviruses with faeces, body fluids and wastewater, these viruses have been suggested as potential source tracking markers (**Table 1.1**) (Noble et al. 2003; Harwood et al. 2014).

Teschoviruses also belong to the order *Picornavirales* and family *Picornaviridae*. The natural hosts of these viruses are pigs. Teschovirus has a single-stranded, linear, non-segmented RNA genome and is the causative agent of porcine enteroviral encephalomyelitis, which is transmitted by the faecal-oral route. Porcine teschovirus has been found exclusively in pig manure, rendering these viruses highly host-specific (Jimenez-Clavero et al. 2003). Therefore, teschovirus could be targeted as an indicator of porcine faecal contamination in environmental waters.

Some examples of the application of enteric viruses include the application of teschovirus and adenovirus to trace pig manure in the environment (Jimenez-Clavero et al. 2003; Maluquer de Motes et al. 2004; Hundesa et al. 2006). In addition, adenovirus, polyomavirus and enterovirus have been employed for the detection of bovine faecal matter in environmental waters (Jimenez-Clavero et al. 2003; Maluquer de Motes et al. 2004; Hundesa et al. 2006).

1.5.1.7 Adenovirus

Adenoviruses have been described as host-specific viruses and human adenovirus has been investigated as an indicator of sewage in source tracking studies (Sidhu et al. 2013; Rusinõl et al. 2014). Adenovirus is a non-enveloped virus with a linear double-stranded DNA genome and is covered by an icosahedral protein shell (Van Heerden et al. 2005). The size of the DNA genome of these viruses ranges from 26 to 45 kb (Jones et al. 2007). There are 51 serotypes

classified into six groups, A to F (Van Heerden et al. 2003). The use of adenoviruses to assess the quality of water is important as they are pathogenic and may cause eye infections, gastroenteritis, respiratory infections, pneumonia, meningitis and hepatitis (Van Heerden et al. 2003; Jones et al. 2007). Therefore, directly monitoring for the presence of human adenovirus which is associated with disease, could provide additional information on the health risk associated with the use of a water source. In addition, these viruses could act as indicators of the presence of other pathogens such as the protozoans *Giardia* and *Cryptosporidium* spp. and other less prevalent infectious viruses (Field & Samadpour, 2007).

Adenoviruses are more resistant than are other enteric viruses to various water disinfection processes such as ultraviolet (UV) radiation. This may be attributed to the double-stranded DNA genome and the presence of DNA repair mechanisms in host cells which enable the repair of viral DNA upon infection. Thus adenovirus persistence in the environment is enhanced and it remains infectious for longer periods of time than do other enteric viruses (Van Heerden et al. 2003; Botes et al. 2013; Sidhu et al. 2013). In addition, adenoviruses are unable to proliferate in the natural environment as they require a host cell to replicate (Rusinõl et al. 2014). This makes adenovirus a promising source tracking marker to trace continuous long-term faecal contamination of a water source (Sidhu et al. 2013). Other adenoviruses that show host-specific distributions are bovine, ovine and porcine adenoviruses (Field & Samadpour, 2007).

In a study conducted by Sidhu et al. (2013), human adenovirus was utilised as a MST marker to assess the extent of human faecal contamination of stormwater run-off. Human adenovirus was identified in 91% of the samples analysed and was detected at a higher frequency than was human polyomavirus (HPv) (56%) (**Table 1.1**). The wide-spread presence of human adenovirus in the stormwater indicated a potential health risk associated with the use of this water for domestic purposes. It was noted however, that the viral particles had to be quantified in order to accurately assess the health risk. Hence, any similar future research should include quantitative assays (Sidhu et al. 2013). It is likely that human adenovirus and human polyomavirus are both present in sewage-contaminated waters as it has been reported that these viruses are present in sewage in high numbers, ranging from 10^5 to 10^6 virus particles per litre. Their presence in such high numbers increases confidence for the use of adenoviruses (and other enteric viruses such as polyomavirus) as source tracking markers. Furthermore, in a study by Waso et al. (2016) adenoviruses were detected in 42.5% of harvested rainwater samples and 52.5% of gutter debris samples analysed. This included the detection of human-specific adenovirus in some of the harvested rainwater and gutter debris samples.

1.5.1.8 *Bacteroides* spp. Bacteriophages

Bacteriophages are viruses that infect bacteria (Jofre et al. 2014). It is known that phages are more abundant than are bacteria in the environment and these viruses only replicate inside an active bacterial host cell (Jofre et al. 2014). Bacteriophages which infect *Bacteroides fragillis* (*B. fragilis*), *Bacteroides thetaiaotaomicron* (*B. thetaiaotaomicron*), *Bacteroides ruminicola* and *Bacteroides ovatus* have been regularly detected in faecal matter and wastewater and have been proposed as suitable MST markers. The replication of *Bacteroides* infective phages in the external environment is highly unlikely as their hosts (specific *Bacteroides* strains) require an anaerobic environment and specific nutrients that are unlikely to occur simultaneously outside of an animal or human host (Jofre et al. 2014). In addition, most phages that infect *Bacteroides* spp. have a narrow host range and as a result usually infects only specific *Bacteroides* strains. The latter are generally associated with a specific animal or human host, thus displaying strict host-specificity. All of these factors indicate that the use of *Bacteroides* phages as MST markers could provide reliable assays for the detection of faecal contamination originating from specific hosts in environmental waters.

To detect phages, specific *Bacteroides* host strains are used in a double layer agar assay. A selected *Bacteroides* strain from a known host (human or animal) is mixed with appropriate culture media and dispensed as the first layer of the agar plate. To prepare the upper layer, the sample to be tested for the presence of phages is mixed with appropriate agar culture media and is poured over the lower layer. This solidifies to form the upper layer. The formation of plaques after incubation of the plates indicates the presence of a phage that infects the *Bacteroides* strain selected for the assay. *B. fragilis* RYC2056 and VPI3625, for example, may be used to detect phages specific to human and non-human faecal matter (thus indicating general faecal contamination of a water source). However other *Bacteroides* strains associated with specific phages may be limited to only human or animal hosts. For example, Bacteriophages B40-8 and φB124.14 infect *B. fragilis* HSP40 and *Bacteroides* sp. GB-124, respectively, and these *Bacteroides* strains are specific to human faecal matter (Jofre et al. 2014). In contrast, Gómez-Doñate et al. (2011) reported isolating *B. fragilis* and *B. thetaiotaomicron* strains which can be infected by bacteriophages specifically associated with faecal matter from pigs, cattle and poultry.

It is possible that obligate faecal anaerobes and associated infecting phages have co-evolved within specific animal or human hosts and accordingly the importance of investigating phages specific to animal and human hosts is emphasised. Future research should focus on further phage discovery (Jofre et al. 2014) and the subsequent development of source tracking markers targeting these phages.

1.5.1.9 F^+ RNA Coliphages

The F^+ RNA phages are a group of icosahedral phages that attach specifically to the F-pili of bacteria, while coliphages refer to viruses that only infect coliforms (Cimenti et al. 2007). There are four subtypes of F^+ RNA coliphages namely I, II, III and IV (Harwood et al. 2014). Subtypes II and III are generally associated with human wastewater whereas subtypes I and IV are associated with animal faecal matter (Hsu et al. 1995; Beekwilder et al. 1996; Stewart-Pullaro et al. 2006; Harwood et al. 2014).

In a study by Wolf et al. (2010) F^+ RNA coliphages were utilised as part of a viral toolbox approach using multiplex qPCR. The coliphages were used as MST markers together with norovirus, adenovirus and atadenovirus (Wolf et al. 2010). The human subtypes (II and III) were more prevalent in porcine, deer and cattle faecal matter than in human faeces. Human subtype II was found in 80% of pig manure samples and 50% of cattle faecal samples. The human-specific subtypes II and III were also detected in high concentrations in sewage influent and these were also the only subtypes detected in sewage biosolids (Wolf et al. 2010; Harwood et al. 2014). Group IV coliphages were detected in sewage and group I was the only subtype detected in abattoir effluent (Wolf et al. 2010; Harwood et al. 2014). These results indicated that F^+ RNA coliphages could be utilised as host-associated markers, but not as host-specific markers as frequent cross-reactivity between human and animal faecal sources was observed across all the coliphage groups. The F^+ RNA coliphages should therefore be combined with source tracking markers that display higher specificity to increase the accuracy of these markers when trying to identify a source of faecal contamination (Harwood et al. 2014).

1.5.1.10 Pepper Mild Mottle Virus

Pepper Mild Mottle Virus (PMMoV) consists of an encapsidated rod-shaped viral particle which has a positive sense linear single-stranded RNA genome (Hamza et al. 2011). This virus is a plant pathogen that infects *Capsicum* spp. including bell, hot and ornamental peppers (Harwood et al. 2014; Kuroda et al. 2015) and is responsible for extensive economic losses wherever it occurs (Hamza et al. 2011). Metagenomic analysis found PMMoV to be the most abundant RNA virus detected in human faeces in numbers as high as 10^6-10^{10} virions per gram of faeces (Hamza et al. 2011). It is hypothesised that these viruses could be consumed in food which contains peppers. Thus high numbers of the virus would occur in human faecal material (Zhang et al. 2005; Colson et al. 2010; Kuroda et al. 2015). The virus has also been isolated from animal faecal matter originating from chickens, cows, geese and seagulls but at much lower concentrations than those detected in human faecal matter (Rosario et al. 2009; Hamza et al. 2011; Kuroda et al. 2015). In addition, studies have indicated that PMMoV is detected in wastewater, surface water and seawater more frequently and in higher concentrations than

enteric viruses such as adenovirus and polyomavirus (Rosario et al. 2009; Hamza et al. 2011; Haramoto et al. 2013; Kitajima et al. 2014; Kuroda et al. 2015). Pepper Mild Mottle Virus has been found to persist in seawater for up to 1.5 days and displays persistence throughout wastewater treatment processes (which may be attributed to the capsid structure of the viral particle). Accordingly, the virus has been proposed as a possible MST marker to trace sewage pollution in environmental waters (Hamza et al. 2011; Harwood et al. 2014).

1.5.1.11 Mitochondrial DNA

Mitochondrial DNA has been suggested as a target for source tracking markers, as the mitochondrial DNA sequences for many animals are readily available in DNA databases. Moreover, mitochondrial DNA is highly specific to animals and humans, is abundant in faecal matter and can be detected in wastewater, surface water and agricultural run-off (Villemur et al. 2015). This is because cells in the digestive tract of animals and humans are sloughed off as consumed food moves through the digestive tract during the process of digestion (Caldwell et al. 2007). The cells then all have multiple mitochondria and therefore results in high concentrations of mitochondrial DNA in the faecal matter from the specific animal or human host (Caldwell et al. 2007). Targeting mitochondrial DNA then ultimately aims at improving on the idea that microorganisms are associated with hosts, and targets the hosts directly by screening for their mitochondrial DNA in contaminated water sources.

In a study conducted by Villemur et al. (2015) water samples were collected from five watersheds and human mitochondrial DNA was detected in 46% of the water samples analysed. Bovine and porcine mitochondrial DNA were identified in 23% and in 6% of the samples, respectively (Villemur et al. 2015) (**Table 1.1**). Poultry mitochondrial DNA was detected in only 3% of the samples collected. A complete absence of ovine mitochondrial DNA was reported (Villemur et al. 2015). The same study also screened for the *Bacteroides* HF183 marker and was detected in 50% of the samples analysed. The authors also observed a strong positive correlation between the occurrence of human mitochondrial DNA and the human-specific *Bacteroides* HF183 marker (Villemur et al. 2015). In addition, co-linearity (two variables are highly correlated, so one variable can be linearly predicted from the other variable with great accuracy) was observed between the levels of coliforms and the human-specific *Bacteroides* HF183 marker. However, no co-linearity was observed between coliforms and human mitochondrial DNA (Villemur et al. 2015).

The use of mitochondrial DNA as a marker may thus provide a direct assay to target specific faecal sources and could ascribe faecal contamination to a specific host more accurately than screening for organisms associated with hosts. In addition, combining mitochondrial DNA

markers with other MST markers could improve discrimination of faecal contamination sources and provide valuable MST targets to include in source tracking studies.

1.5.2 Chemical Source Tracking

Chemical source tracking is often used in conjunction with MST as a supplementary screening tool to increase confidence in results obtained from MST markers (Field & Samadpour, 2007). Chemical source tracking refers to the utilisation of chemical compounds such as fragrance compounds and chemicals found in personal care products, faecal sterols and stanols, optical brighteners and compounds associated with food and pharmaceuticals to trace faecal or sewage pollution associated with human activity, but could also include chemical compounds associated with animal faecal matter or agricultural run-off (Field & Samadpour, 2007; Harwood et al. 2014). Chemical markers that persist in the environment such as the artificial sweeteners; acesulfame and aspartame, are useful in tracing routes of faecal or sewage contamination in environmental waters. Biodegradable compounds such as caffeine can then be utilised as indicators of untreated wastewater contamination in fresh water (Buerge et al. 2003; 2009; Sidhu et al. 2013). Research efforts have emphasised the identification of chemical markers specific to anthropogenic activity and therefore chemical compounds associated with animal faecal matter have not been studied as extensively (Glassmeyer et al. 2005).

In a study conducted by Glassmeyer et al. (2005), wastewater effluent compounds originating from human waste were identified and it was suggested that these could serve as downstream indicators of faecal pollution in rivers and streams, which had possibly been contaminated by closely located wastewater treatment plants. The study identified 35 potential chemical faecal indicators for use in tracing human sewage contamination in environmental waters. Of these, caffeine, acetaminophen (paracetamol), cholesterol, carbamazepine and triclosan showed the most promise (**Table 1.2**). Similarly, Nakada et al. (2008) examined 13 pharmaceuticals and personal care products as prospective indicators of sewage contamination in riverine, coastal and groundwater. The study identified triclosan, ibuprofen, crotamiton and carbamazepine as promising sewage indicators and emphasised that the presence of these markers in river and groundwater samples could indicate contamination by sewage effluent. Some of these potential CST markers will be discussed in more detail in the following sections.

Table 1.2 Chemical Source Tracking Markers Applied to Various Water Sources as Indicators of Anthropogenic Contamination

Compound (Concentration)	Water Analysed	Detection Frequency in Samples Analysed	Reference
Acesulfame (µg/L)	Stormwater run-off	96%	Sidhu et al. 2013
	Wastewater	100%	Buerge et al. 2009; Scheurer et al. 2009
	Groundwater	65% - 100%	
	Surface water	100%	Scheurer et al. 2009
Acetaminophen (µg/L)	Surface water	73%	Glassmeyer et al. 2005
	Wastewater	-	Al-Rifai et al. 2007
	Stormwater run-off	87%	Sidhu et al. 2013
	Harvested rainwater	5%	Waso et al. 2016
Salicylic acid (µg/L)	Wastewater	-	Al-Rifai et al. 2007
	Stormwater run-off	78%	Sidhu et al. 2013
	Harvested rainwater	37.5%	Waso et al. 2016
Caffeine (ng/L - µg/L)	Surface water and wastewater	73%	Glassmeyer et al. 2005
	Surface water	-	Heberer et al. 2002
	Lakes	-	Buerge et al. 2003
	Rivers	-	Buerge et al. 2003
	Mediterranean Sea	-	Buerge et al. 2003
	Groundwater	29%	Nakada et al. 2008
	Stormwater run-off	91%	Sidhu et al. 2013
	Harvested rainwater	100%	Waso et al. 2016
Carbamazepine (ng/L - µg/L)	Wastewater	100%	Clara et al. 2004
	Wastewater	-	Al-Rifai et al. 2007
	River water	-	Nakada et al. 2008
	Groundwater	36%	Nakada et al. 2008
	Groundwater	91%	Glassmeyer et al. 2005
	Surface water	-	Heberer et al. 2002
Cholestrol (µg/L)	Surface Water and Wastewater	91%	Glassmeyer et al. 2005
Coprostanol (µg/L)	Wastewater	-	Gregor et al. 2002
24-ethylcoprostanol (µg/L)	Wastewater	-	Gregor et al. 2002

Table 1.2 (Continued) Chemical Source Tracking Markers Applied to Various Water Sources as Indicators of Anthropogenic Contamination

Compound (Concentration)	Water Analysed	Detection Frequency in Samples Analysed	Reference
6-phenyldodecane (µg/L)	Wastewater	-	Gregor et al. 2002
Fluorescent whitening agents (µg/L)	Wastewater	-	Gregor et al. 2002
	Wastewater and surface water	-	Hayashi et al. 2002
Triclosan (ng/L - µg/L)	Wastewater	100%	Glassmeyer et al. 2005
	River water	-	Nakada et al. 2008

1.5.2.1 Faecal Sterols and Stanols

Sterols are steroid alcohols that occur naturally in fungi, plants and animals. The most common type of sterol is cholesterol. Faecal stanols are formed from sterols in the gastrointestinal tract of animals and humans during the digestion of food. Conversion of sterols to stanols is dependent on the diet, metabolism and the gut microbial flora of animals or humans (Gregor et al. 2002). Cholesterol, for example, is converted to coprostanol in humans but in the natural environment it is reduced to cholestanol (Gregor et al. 2002). Sterols and stanols have been studied in animal faeces and it was found that these compounds could be utilised to differentiate among faecal matter from different hosts. This was attributed to variable metabolic processes used as well as to physiological differences among various animal species (Leeming et al. 1996; Gregor et al. 2002). Coprostanol was found to be predominantly associated with human faeces whereas 24-ethylcoprostanol was specific to herbivores (Gregor et al. 2002). In a study by Gregor et al. (2002) it was concluded that septic tank effluents could be distinguished from community wastewater because of the higher concentrations of faecal stanols present in the former. However, further research is required for the elucidation of faecal sterols and stanols that are host-specific for a variety of host sources. Once this has been completed, to ensure accurate source tracking analyses, the markers need to be tested and validated in the field against different sample types (Gregor et al. 2002).

1.5.2.2 Compounds Associated with Food, Pharmaceuticals and Personal Care Products

Caffeine is an alkaloid that is found in more than 60 plants including the seeds of coffee, cacao and cola trees (Buerge et al. 2003). Caffeine is considered a pharmacologically active compound (PhAC) and it is thus found in various medications such as analgesics, where it enhances the effect of the drug. Caffeine also acts as a diuretic (Buerge et al. 2003). It is widely

consumed by humans as it is prevalent in cold beverages, certain medications, coffee and in certain foodstuffs such as chocolates and pastries (Heberer et al. 2002; Buerge et al. 2003). Subsequently this compound occurs in raw municipal wastewater at very high concentrations [microgram per litre (µg/L)]. Caffeine has therefore been proposed as a potential human associated source tracking marker (Heberer et al. 2002; Buerge et al. 2003; Sidhu et al. 2013).

Acesulfame is an artificial low-calorie sweetener and is used in a variety of foods, beverages, pharmaceuticals and personal care products such as mouthwash and toothpaste (Scheurer et al. 2009; Sidhu et al. 2013). This compound has been proposed as a promising source tracking marker as it is not removed during wastewater treatment and occurs in wastewater at levels as high as µg/L (Buerge et al. 2009; Scheurer et al. 2009; Sidhu et al. 2013). Buerge et al. (2009) investigated the prevalence and persistence of acesulfame in domestic wastewater (treated and untreated), surface water, groundwater and tap water samples. The study reported that acesulfame was constantly detected in treated and untreated wastewater samples at levels ranging from 10.0 to 46.0 µg/L (**Table 1.2**). It was also detected in surface water (up to 2.8 µg/L), groundwater (up to 4.7 µg/L) and tap water (up to 2.6 µg/L). However, the concentrations were approximately tenfold less than the concentrations detected in wastewater samples. The authors also noted a marked positive correlation between the concentration of acesulfame in surface water samples and the anthropogenic load in the vicinity of the surface water sources. Thus they concluded that acesulfame may be utilised to quantitatively assess contamination of environmental waters by domestic wastewater (Buerge et al. 2009).

Certain pharmaceuticals have also been identified as potential markers of sewage contamination in the environment. These include analgesics such as acetaminophen (also known as paracetamol) and salicylic acid (metabolic product of aspirin) (Sidhu et al. 2013). These pharmaceuticals are of particular interest as they may be obtained from a pharmacy without a prescription and as a result are widely used by the human population. Furthermore, salicylic acid is not only derived from aspirin, but may be found in personal care products such as face wash, face scrub, ointments for the treatment of acne and products used to treat warts. Salicylic acid and acetaminophen are biodegradable and have shown to have removal rates of up to 100% during wastewater treatment (Sidhu et al. 2013). This characteristic renders these compounds suitable for use as potential markers of raw sewage and household waste contamination in the environment.

In a study conducted in Australia by Sidhu et al. (2013) caffeine, acetaminophen, salicylic acid and acesulfame were utilised as CST markers in conjunction with adenovirus, polyomavirus, *Bacteroides* HF183 and *M. smithii nifH* as MST markers, to screen for human contamination in urban stormwater run-off. Caffeine was detected in 91% of the samples tested and acesulfame occurred more frequently, in 96% of the samples. Acetaminophen was present in 87% of the

samples analysed. Salicylic acid was the least prevalent marker and was detected in 78% of the samples. Caffeine, acesulfame, salicylic acid and acetaminophen were all detected at µg/L levels in the samples (**Table 1.2**). In addition, caffeine displayed a high concurrence with acesulfame (87%) and human adenovirus (83%). The study revealed a notable concurrence (> 80%) of the sewage associated marker *Bacteroides* HF183, human adenovirus, acesulfame, acetaminophen and caffeine, suggesting that these markers were promising indicators of wastewater sources in environmental waters and could possibly be utilised as a toolbox of markers to trace sewage pollution in environmental waters (Sidhu et al. 2013).

In a study conducted by Waso et al. (2016), caffeine, salicylic acid and acetaminophen were screened for in gutter debris and harvested rainwater samples collected from DRWH systems, together with the MST markers, HF183 and adenovirus. Throughout the study period, caffeine was the predominant CST marker detected in the harvested rainwater samples (100%) and salicylic acid was the most common CST marker detected in gutter debris samples (100%). In contrast, acetaminophen was only detected sporadically in both the tank water and gutter debris samples. The CST markers were all detected at µg/L levels (Waso et al. 2016). The study also revealed strong concurrence frequencies between *Bacteroides* HF183 and caffeine (80%) and HF183 and salicylic acid (95%) in gutter debris samples, as well as moderate concurrence frequencies between HF183 and salicylic acid (60%) in harvested rainwater samples. These compounds could then indicate household waste or anthropogenic contamination of domestically harvested rainwater.

1.5.2.3 Carbamazepine

Carbamazepine is an established pharmaceutical used for the treatment of grand mal seizures, psychomotor epilepsy and is prescribed for bipolar depression (Clara et al. 2004). This compound is primarily metabolised in the liver to form carbamazepine 10, 11 – epoxide and 2% to 3% of a given dose is excreted in an unchanged form in the faeces of patients using the drug (Clara et al. 2004). Carbamazepine has been detected at nanogram per litre (ng/L) levels in various surface water sources and it is believed that this compound enters surface waters, rivers and streams through wastewater treatment plant effluents as it is not degraded or removed during wastewater treatment processes (Clara et al. 2004). It therefore has potential for use as a source tracking marker and has been investigated in this regard.

Carbamazepine was utilised as a source tracking marker for sewage detection in 37 Japanese river systems and was detected at levels not exceeding 34.7 ng/L (Nakada et al. 2008) (**Table 1.2**). It showed a positive correlation with human population density (r^2 = 0.84) thereby increasing confidence for the use of this compound as a marker for the detection of wastewater contamination in the natural environment (Nakada et al. 2008). The authors also screened for

pharmaceuticals in groundwater and found carbamazepine in five of the 14 samples analysed, at an average concentration of 1.6 ng/L. They concluded that carbamazepine is a stable and persistent compound associated with human waste and the application of this compound as a CST marker in different water sources should be investigated (Nakada et al. 2008).

1.5.2.4 Fluorescent Whitening Agents

Stilbene-type fluorescent whitening agents (FWA) which are found in laundry detergents, have been investigated as possible source tracking markers and include the compounds: 4,4' – bis (2-sulfostyryl) biphenyl (DSBP) and 4,4' bis [(4 anilino 6 morpholino 1.3.5 triazin 2 yl) amino] stilbene 2,2' disulfonate (DAS1) (Hayashi et al. 2002). These compounds are water soluble and can be used to trace the movement of contaminants through a water source by using ultraviolet light (Hayashi et al. 2002). In addition, these compounds are not biodegradable and could therefore be useful sewage indicators in environmental waters (Hayashi et al. 2002).

In a study conducted in Tokyo by Hayashi et al. (2002), FWA were found to have removal rates of between 15 and 79% during wastewater treatment processes and were subsequently detected ubiquitously in rivers polluted by wastewater effluents. Dissolved DSBP was detected at levels ranging from 0.1 µg/L to 6.4 µg/L and DAS1 was detected at approximately 1 µg/L (**Table 1.2**). These compounds were further detected in Tokyo Bay at ng/L concentrations. It was concluded that FWA could be useful as source tracking markers of sewage effluents in river and coastal water sources. It was also shown that FWA can be accurately detected in as little as 100 mL of water.

1.5.2.5 Antimicrobials Associated with Personal Care Products

Triclosan is a synthetic antimicrobial agent with a broad spectrum of antimicrobial activity against bacteria and fungi (Yalavarthy et al. 2015). It inhibits these microorganisms by targeting the enocyl-acyl carrier protein reductase enzyme which is essential for fatty acid synthesis (Heath et al. 1999; Yalavarthy et al. 2015). Triclosan is a common constituent of a variety of personal care and household products. It may be found in toothpaste, hand soaps, dishwashing liquids, liquid soaps, cosmetics, shampoos, deodorants and mouthwash (Cox, 1987; Jones et al. 2000; Yalavarthy et al. 2015). In addition, triclosan has been incorporated into various household items including toys, mattresses, carpets, food storage containers and paints. This prevents bacterial and fungal growth and inhibits the development of unpleasant odours in these items (Glaser, 2004; Orhan et al. 2009; Yalavarthy et al. 2015). Research has however shown that triclosan exhibits endocrine disruptive properties and may influence normal hormone function in animals. Development and reproduction of aquatic and terrestrial organisms could therefore be adversely affected (Chalew & Halden, 2009; Yalavarthy et al. 2015). Triclosan is known to bio-accumulate and therefore it could be considered an emerging environmental

contaminant. In humans and animals, triclosan can accumulate in fatty tissue as it exhibits lipophilic characteristics. It has also been detected in urine, breast milk and blood (Yalavarthy et al. 2015). As a result of the widespread use of triclosan in household products, this marker may be valuable in tracing contamination originating from anthropogenic activities and domestic waste in environmental waters. In a study by Glassmeyer et al. (2005) the potential use of triclosan as a wastewater indicator in the environment was highlighted as it was found in high concentrations in wastewater and is relatively stable and persistent through wastewater treatment processes (**Table 1.2**).

Triclocarban is widely utilised as a topical antiseptic commonly found in soaps, detergents, cosmetics and other personal care products (Heidler et al. 2006). Similar to triclosan, triclocarban exhibits endocrine disrupting characteristics and may have adverse effects on the reproduction system of animals (Heidler et al. 2006). In addition, triclocarban may lead to disease in humans, including methemoglobinemia and cancer, due to the presence of harmful triclocarban transformation products (Heidler et al. 2006). As a result of the widespread use of triclocarban in everyday consumer products and the subsequent high concentrations of the compound in wastewater, it could also act as a suitable marker to trace contamination originating from anthropogenic activities in the environment. However, triclocarban has not been as extensively investigated as a source tracking marker as has triclosan and the prevalence of this compound in environmental waters should be investigated.

1.6 Study Site Description

In 2010, the Council for Scientific and Industrial Research (CSIR) in collaboration with the Overstrand Municipality, the Department of Science and Technology and Eskom initiated the development of the Kleinmond Housing Scheme (GPS co-ordinates: 34°20.11'81"S 19°00.59'74"E) in Kleinmond, a peri-urban coastal region in the Western Cape, South Africa (WRC Project K5/2124/3 Report, 2014). The housing scheme is comprised of three sites indicated by the circles in **Fig. 1.2**. The locations indicated by black circles consist of 25 (A) and 32 (B) houses respectively, and the site selected for the current study is indicated with the red circle (C) where there are 354 housing units (WRC Project K5/2124/3 Report, 2014).

Fig. 1.2. Aerial image of Kleinmond, Western Cape. The locations for the Kleinmond Housing Scheme (GPS co-ordinates: 34°20.11'81"S 19°00.59'74"E) sites are indicated by A, B and C, where the area in the red circle (C) was investigated during the current study (Adopted from WRC Project K5/2124/3 Report, 2014).

In order to decrease the dependency on municipal water and electricity supplies, each house (40 m^2) was fitted with sustainable technologies such as an above-ground DRWH tank (2000 L), a solar geyser and a photo-voltaic panel (**Fig. 1.3.**) (CSIR, 2011). The rooftop of each house acts as a catchment surface for the collection of rainwater which is then conveyed to a 2000 L polyethylene storage tank via a gutter system. Rooftops were constructed from Double Roman standard plus concrete roof tiles. The quality of the harvested rainwater at the Kleinmond Housing Scheme site (C) was previously investigated by Dobrowsky et al. (2014a; 2014b; 2014c). In these studies, it was found that although metal cation and anion concentrations were within acceptable limits as specified by drinking water guidelines, faecal indicator organisms and various pathogens exceeded acceptable levels in the harvested rainwater (Dobrowsky et al. 2014a; 2014b; 2014c). Thus the microbiological quality of the harvested water was compromised.

Fig. 1.3. Houses in the Kleinmond Housing Scheme site fitted with 2000 L DRWH tanks, solar geysers and photo-voltaic panels (Adopted from WRC Project K5/2124/3 Report, 2014).

A social perception study (68 respondents) was also conducted at the Kleinmond Housing Scheme site (C) (WRC Project K5/2124/3 Report, 2014). This study showed that the Kleinmond residents utilised the harvested rainwater for various domestic purposes. The majority of the home owners used the tank water for laundry purposes (92%), followed by house cleaning (70%), gardening (46%), bathing (44%) and cooking (19%). Of particular concern was the proportion of residents utilising the tank water for drinking purposes (24%) without any pre-consumption purification treatment. Previous screening of the Kleinmond Housing Scheme tank water as conducted by Dobrowsky et al. (2014a, 2014b) indicated the presence of pathogens including virulent *E. coli*, *Legionella* spp. and *Pseudomonas* spp. amongst others. The presence of these pathogens thus poses a serious health risk to the individuals who utilise the tank water for bathing, cooking and consumption.

Microbial contaminants probably enter the tanks by means of debris (which includes soil and faecal matter from animals) washing into the tanks from the rooftops. Sources of microbial contaminants are thought to include a variety of bird species frequently observed during sampling sessions, on the rooftops utilised as rainwater catchment surfaces and in the vicinity of the rainwater harvesting tanks. In addition, the residents' pets (cats and dogs) roam freely and often defecate around the rainwater harvesting tanks, which may serve as sources of animal faecal contamination (Waso et al. 2016). Some residents have also reported that they store garbage bags on top of the rainwater tanks to prevent dogs from scavenging the household waste for food. This refuse would also contaminate water in the tank. Others

reported that the lids of the rainwater tanks were damaged or missing thereby contributing to the contamination of the stored rainwater (WRC Project K5/2124/3 Report, 2014). As reported elsewhere it is likely that anthropogenic activities in close proximity to the tanks exert a marked influence on the quality of harvested rainwater (Mwenge Kahinda et al. 2007).

It is crucial to identify the dominant contamination sources influencing the quality of tank water and to subsequently identify markers that may be used to supplement indicator analysis such that the quality of the stored rainwater is accurately monitored. Such information will be of value for future studies where the potential health risks associated with the use of the harvested rainwater are assessed. Information on the dominant contamination sources of the Kleinmond study site may also assist the future implementation of preventative strategies or treatment systems to improve the microbiological quality of the harvested rainwater. A study conducted by Waso et al. (2016) at the Kleinmond Housing scheme site detected *Bacteroides* HF183, adenovirus, caffeine, salicylic acid and acetaminophen in the tank water and gutter debris samples. These results highlighted the potential for the use of MST and CST markers to supplement indicator organism analyses to monitor the quality of the harvested rainwater. Therefore, to compile a toolbox of MST and CST markers to supplement indicator organism analysis, the presence of other reported source tracking markers at this study site should be investigated.

1.7 Project Aims

Rainwater harvesting has been earmarked as a sustainable solution for the supply of freshwater directly to households, especially to those located in urban informal settlements and rural areas. The implementation of DRWH systems should aid in alleviating the increasing demand often experienced by municipal water supplies. The technology will also provide access to an on-site freshwater source for households in dispersed settlements and could improve sanitation practices. Rainwater is not a pure water source and various pathogens have been detected in harvested rainwater. Thus, it has limited domestic use (Uba & Aghogho, 2000; Ahmed et al. 2008a; Simmons et al. 2008; Ahmed et al. 2010a; 2012; De Kwaadsteniet et al. 2013; Dobrowsky et al. 2014a; 2014b).

Monitoring water sources such as rainwater for all known pathogens is not feasible and it is therefore standard practice to utilise indicator organisms and FIB for this purpose. However, this approach has limitations. These include the lack of correlations between indicator organisms and pathogens present in water sources, the fact that certain indicator strains are able to persist and proliferate in the environment and the lack of host-specificity noted for indicator organisms. Furthermore, culture-based techniques which are commonly used for indicator organism analysis introduce a bias for culturable organisms and ignore the viable but non-culturable microorganisms present in a particular water source.

Source tracking has thus emerged as a supplementary tool to monitor water quality and to identify the dominant sources of contaminants present in a water source. Furthermore, the use of molecular techniques such as the PCR and qPCR produce accurate results in a short space of time. These techniques may also detect viable but non-culturable and anaerobic microorganisms in water. The ability to identify and trace the origin of faecal contamination in water to specific sources offers a number of advantages. These include assisting with the development of strategies to limit or prevent contamination of the water source, implementing effective remediation strategies and accurately assessing the health risk associated with the use of the water. In addition, there is a need to identify the relevant source tracking markers that correlate with waterborne FIB and pathogens. By achieving this, the presence of faecal contamination would be reliably detected and the presence of pathogens accurately predicted. By the use of these methods, monitoring of water quality can be performed rapidly and water remediation efforts should improve.

The primary aim of the current study was thus to identify a toolbox of MST and CST markers present in DRWH systems which could be implemented to augment or supplement indicator organism analysis for future screenings of DRWH systems. This aim was achieved as outlined in the research chapters as follows:

Chapter 2: *Primary Microbial and Chemical Source Tracking Markers Associated with Domestic Rainwater Harvesting Systems: Correlation to Indicator Organisms*

- Harvested rainwater and gutter debris were screened for MST markers by means of conventional PCR assays for the detection of: human-specific *Bacteroides* HF183, *Bifidobacterium* spp., *Enterococcus esp* gene, Lachnospiraceae, adenovirus, polyomavirus, enterovirus, *Methanobrevibacter* spp., *Methanobrevibacter smithii nif*H, human, bovine and porcine mitochondrial DNA. This was done in order to identify the predominant MST markers associated with rainwater harvesting systems in Kleinmond, Western Cape.

- Harvested rainwater and gutter debris samples were screened for the presence of potential CST markers. This was done by using solid phase extraction (SPE) together with hydrophilic-lipophilic balanced (HLB) cartridges and high performance liquid chromatography tandem mass spectrometry (HPLC/MS/MS) to detect caffeine, acetaminophen, carbamazepine, salicylic acid, triclosan, triclocarban and methyl paraben. In this manner, the most frequently occurring CST markers associated with rainwater harvesting systems in Kleinmond, Western Cape were identified.

- Harvested rainwater was screened for the presence of traditional indicator organisms: *E. coli*, enterococci, total coliforms, faecal coliforms and heterotrophic bacteria using

culture based analysis in an attempt to correlate indicator counts with the source tracking markers.

- Quantitative PCR assays were optimised to detect and enumerate the predominant MST markers in the harvested rainwater and gutter debris.

- Quantitative PCR assays were also optimised for the detection and quantification of *E. coli* and enterococci in both harvested rainwater and gutter debris.

- Correlation analyses was performed in order to identify relationships between the MST and CST markers, indicator organism counts and the numbers of *E. coli* and enterococci present as determined by qPCR. This was performed in order to identify potential source tracking markers which could be used to supplement indicator analyses for future DRWH system monitoring.

Chapter 3: *The Development and Small-Scale Validation of a Novel Avian-Associated Mitochondrial DNA Source Tracking Marker for the Detection of Avian Faecal Contamination in Domestic Rainwater Harvesting Systems*

- Primers specific to avian spp. mitochondrial DNA were designed using the ClustalW software.

- Conventional PCR assays were optimised for use with the novel avian mitochondrial DNA primers.

- A small-scale validation study was carried out by screening a variety of avian faecal samples (chicken, seagull, goose, macaw, hadeda, pigeon) and non-host faecal samples (bovine, porcine, human, canine, feline) to determine the host-specificity and host-sensitivity of the novel avian mitochondrial DNA marker.

- The avian mitochondrial DNA markers were used to monitor harvested rainwater and gutter debris samples for the presence of avian faecal contamination in the DRWH systems in Kleinmond, Western Cape.

- Bayesian statistics were applied to determine the conditional probability of the designed avian mitochondrial DNA MST markers to detect true avian faecal contamination in the harvested rainwater and gutter debris samples.

1.8 References

Abbasi, T. & Abbasi, S. 2011. Sources of pollution in rooftop rainwater harvesting systems and their control. *Critical Reviews in Environmental Science and Technology.* 41(23):2097-2167.

Ahmed, W., Gardner, T. & Toze, S. 2011a. Microbiological quality of roof-harvested rainwater and health risks: A review. *Journal of Environmental Quality.* 40(1):13-21.

Ahmed, W., Goonetilleke, A. & Gardner, T. 2010b. Human and bovine adenoviruses for the detection of source-specific faecal pollution in coastal waters in Australia. *Water Research.* 44(16):4662-4673.

Ahmed, W., Goonetilleke, A., Powell, D. & Gardner, T. 2009. Evaluation of multiple sewage-associated *Bacteroides* PCR markers for sewage pollution tracking. *Water Research.* 43(19):4872-4877.

Ahmed, W., Hamilton, K., Gyawali, P., Toze, S. & Haas, C. 2016. Evidence of Avian and Possum Fecal Contamination in Rainwater Tanks as Determined by Microbial Source Tracking Approaches. *Applied and Environmental Microbiology.* DOI: 10.1128/AEM.00892-16.

Ahmed, W., Harwood, V., Nguyen, K., Young, S., Hamilton, K. & Toze, S. 2015. Utility of *Helicobacter* spp. associated GFD markers for detecting avian faecal pollution in natural waters of two continents. *Water Research.* 88:613-622.

Ahmed, W., Hodgers, L., Sidhu, J.P. & Toze, S. 2012. Faecal indicators and zoonotic pathogens in household drinking water taps fed from rainwater tanks in Southeast Queensland, Australia. *Applied and Environmental Microbiology.* 78(1):219-226.

Ahmed, W., Huygens, F., Goonetilleke, A. & Gardner, T. 2008a. Real-time PCR detection of pathogenic microorganisms in roof-harvested rainwater in Southeast Queensland, Australia. *Applied and Environmental Microbiology.* 74(17):5490-5496.

Ahmed, W., Sidhu, J. & Toze, S. 2011b. Evaluation of the *nif*H gene marker of *Methanobrevibacter smithii* for the detection of sewage pollution in environmental waters in Southeast Queensland, Australia. *Environmental Science & Technology.* 46(1):543-550.

Ahmed, W., Stewart, J., Powell, D. & Gardner, T. 2008b. Evaluation of the host-specificity and prevalence of Enterococci Surface Protein (ESP) marker in sewage and its application for sourcing human faecal pollution. *Journal of Environmental Quality.* 37(4):1583-1588.

Ahmed, W., Vieritz, A., Goonetilleke, A. & Gardner, T. 2010a. Health risk from the use of roof-harvested rainwater in Southeast Queensland, Australia, as potable or non-potable water,

determined using quantitative microbial risk assessment. *Applied and Environmental Microbiology.* 76(22):7382-7391.

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

Albinana-Gimenez, N., Miagostovich, M.P., Calgua, B., Huguet, J.M., Matia, L. & Girones, R. 2009. Analysis of adenoviruses and polyomaviruses quantified by qPCR as indicators of water quality in source and drinking water treatment plants. *Water Research.* 43(7):2011-2019.

Albrechtsen, H. 2002. Microbiological investigations of rainwater and graywater collected for toilet flushing. *Water Science & Technology.* 46(6-7):311-316.

Al-Rifai, J.H., Gabelish, C.L. & Schäfer, A.I. 2007. Occurrence of pharmaceutically active and non-steroidal estrogenic compounds in three different wastewater recycling schemes in Australia. *Chemosphere.* 69(5):803-815.

Anderson, K.L., Whitlock, J.E. & Harwood, V.J. 2005. Persistence and differential survival of faecal indicator bacteria in subtropical waters and sediments. *Applied and Environmental Microbiology.* 71(6):3041-3048.

Ballestè, E. & Blanch, A.R. 2011. Bifidobacterial diversity and the development of new microbial source tracking indicators. *Applied and Environmental Microbiology.* 77(10):3518-3525.

Beekwilder, J., Nieuwenhuizen, R., Havelaar, A. & Van Duin Leiden, J. 1996. An oligonucleotide hybridization assay for the identification and enumeration of F^+-specific RNA phages in surface water. *Journal of Applied Bacteriology.* 80(2):179-186.

Belay, N., Mukhopadhyay, B., Conway de Macario, E., Galask, R. & Daniels, L. 1990. Methanogenic bacteria in human vaginal samples. *Journal of Clinical Microbiology.* 28(7):1666-1668.

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

Bernhard, A.E. & Field, K.G. 2000b. Identification of nonpoint sources of faecal pollution in coastal waters by using host-specific 16S ribosomal DNA genetic markers from faecal anaerobes. *Applied and Environmental Microbiology.* 66(4):1587-1594.

Bonadonna, L., Briancesco, R., Ottaviani, M. & Veschetti, E. 2002. Occurrence of *Cryptosporidium* oocysts in sewage effluents and correlation with microbial, chemical and physical water variables. *Environmental Monitoring and Assessment.* 75(3):241-252.

Bond, J.H., Engel, R.R. & Levitt, M.D. 1971. Factors influencing pulmonary methane excretion in man: An indirect method of studying the *in situ* metabolism of the methane-producing colonic bacteria. *The Journal of Experimental Medicine.* 133(3):572-588.

Botes, M., De Kwaadsteniet, M. & Cloete, T.E. 2013. Application of quantitative PCR for the detection of microorganisms in water. *Analytical and Bioanalytical Chemistry.* 405:91-108.

Buerge, I.J., Buser, H., Kahle, M., Muller, M.D. & Poiger, T. 2009. Ubiquitous occurrence of the artificial sweetener acesulfame in the aquatic environment: An ideal chemical marker of domestic wastewater in groundwater. *Environmental Science & Technology.* 43(12):4381-4385.

Buerge, I.J., Poiger, T., Müller, M.D. & Buser, H. 2003. Caffeine, an anthropogenic marker for wastewater contamination of surface waters. *Environmental Science & Technology.* 37(4):691-700.

Caldwell, J.M. & Levine, J.F. 2009. Domestic wastewater influent profiling using mitochondrial real-time PCR for source tracking animal contamination. *Journal of Microbiological Methods.* 77(1):17-22.

Caldwell, J.M., Raley, M.E. & Levine, J.F. 2007. Mitochondrial multiplex real-time PCR as a source tracking method in faecal-contaminated effluents. *Environmental Science & Technology.* 41(9):3277-3283.

Chalew, T.E. & Halden, R.U. 2009. Environmental exposure of aquatic and terrestrial biota to triclosan and triclocarban. *Journal of the American Water Resources Association.* 45(1):4-13.

Centres for Disease Control and Prevention (CDC). 2014. *Non-polio enterovirus*. Available: http://www.cdc.gov/non-polio-enterovirus/about/index.html [2015, December 10].

Cimenti, M., Hubberstey, A., Bewtra, J. & Biswas, N. 2007. Alternative methods in tracking sources of microbial contamination in waters. *Water SA.* 33(2).

Clara, M., Strenn, B. & Kreuzinger, N. 2004. Carbamazepine as a possible anthropogenic marker in the aquatic environment: Investigations on the behaviour of carbamazepine in wastewater treatment and during groundwater infiltration. *Water Research.* 38(4):947-954.

Colson, P., Richet, H., Desnues, C., Balique, F., Moal, V., Grob, J., Berbis, P., Lecoq, H., Harlé, J., Berland, Y. & Raoult, D. 2010. Pepper Mild Mottle Virus, a plant virus associated with

specific immune responses, fever, abdominal pains, and pruritus in humans. *PloS One.* 5(4):e10041.

Cox, A.R. 1987. Efficacy of the antimicrobial agent triclosan in topical deodorant products: Recent developments *in vivo*. *Journal of the Society of Cosmetic Chemists.* 38:223-231.

Crabtree, K.D., Ruskin, R.H., Shaw, S.B. & Rose, J.B. 1996. The detection of *Cryptosporidium* oocysts and *Giardia* cysts in cistern water in the US Virgin Islands. *Water Research.* 30(1):208-216.

Council for Scientific and Industrial Research (CSIR), 2011. Sustainable Implementation, Monitoring and Evaluation. Available: http://www.csir.co.za/publications/pdfs/Sustainable_implementation_monitoring_evaluation.pdf [2016, July 28].

De Kwaadsteniet, M., Dobrowsky, P.H., Van Deventer, A., Khan, W. & Cloete, T. E. 2013. Domestic rainwater harvesting: microbial and chemical water quality and point-of-use treatment systems. *Water, Air, & Soil Pollution.* 224(7):1-19.

Department of Water Affairs and Forestry (DWAF). (1996). South African water quality guidelines 2nd Edition, Volume 1: Domestic water use. CSIR Environmental Services. Pretoria, South Africa.

Despins, C., Farahbakhsh, K. & Leidl, C. 2009. Assessment of rainwater quality from rainwater harvesting systems in Ontario, Canada. *Aqua.* 58(2):117.

Di Rosa, R., Creti, R., Venditti, M., D'Amelio, R., Arciola, C.R., Montanaro, L. & Baldassarri, L. 2006. Relationship between biofilm formation, the Enterococcal Surface Protein (ESP) and gelatinase in clinical isolates of *Enterococcus faecalis* and *Enterococcus faecium*. *FEMS Microbiology Letters.* 256(1):145-150.

Dobrowsky, P.H., Mannel, D., De Kwaadsteniet, M., Prozesky, H., Khan, W. & Cloete, T.E. 2014c. Quality assessment and primary uses of harvested rainwater in Kleinmond, South Africa. *Water SA.* 40(3):401-406.

Dobrowsky, P.H., De Kwaadsteniet, M., Cloete, T.E. & Khan, W. 2014b. Distribution of indigenous bacterial pathogens and potential pathogens associated with roof-harvested rainwater. *Applied and Environmental Microbiology.* 80(7):2307-2316.

Dobrowsky, P.H., Van Deventer, A., De Kwaadsteniet, M., Ndlovu, T., Khan, S., Cloete, T.E. & Khan, W. 2014a. Prevalence of virulence genes associated with pathogenic *Escherichia coli*

strains isolated from domestically harvested rainwater during low- and high-rainfall periods. *Applied and Environmental Microbiology.* 80(5):1633-1638.

Eaton, T.J. & Gasson, M.J. 2002. A variant Enterococcal Surface Protein (ESP) in *Enterococcus faecium*; distribution among food, commensal, medical, and environmental isolates. *FEMS Microbiology Letters.* 216(2):269-275.

Farreny, R., Morales-Pinzón, T., Guisasola, A., Taya, C., Rieradevall, J. & Gabarrell, X. 2011. Roof selection for rainwater harvesting: Quantity and quality assessments in Spain. *Water Research.* 45(10):3245-3254.

Field, K.G. & Samadpour, M. 2007. Fecal source tracking, the indicator paradigm, and managing water quality. *Water Research.* 41(16):3517-3538.

Fisher, K. & Phillips, C. 2009. The ecology, epidemiology and virulence of *Enterococcus*. *Microbiology (Reading, England).* 155(6):1749-1757.

Food and Agriculture Organization (FAO). 1985. *Irrigation water management: training manual No.1-Introduction to irrigation*. Available: http://www.fao.org/docrep/r4082e/r4082e05.htm [2014, December 13]

Gikas, G.D. & Tsihrintzis, V.A. 2012. Assessment of water quality of first-flush roof run-off and harvested rainwater. *Journal of Hydrology.* 466:115-126.

Glaser, A. 2004. The ubiquitous triclosan: A common antibacterial agent exposed. *Pesticides and You.* 24(3):12-17.

Glassmeyer, S.T., Furlong, E.T., Kolpin, D.W., Cahill, J.D., Zaugg, S.D., Werner, S.L., Meyer, M.T. & Kryak, D.D. 2005. Transport of chemical and microbial compounds from known wastewater discharges: Potential for use as indicators of human faecal contamination. *Environmental Science & Technology.* 39(14):5157-5169.

Gómez-Doñate, M., Payán, A., Cortés, I., Blanch, A.R., Lucena, F., Jofre, J. & Muniesa, M. 2011. Isolation of bacteriophage host strains of *Bacteroides* species suitable for tracking sources of animal faecal pollution in water. *Environmental Microbiology.* 13(6):1622-1631.

Gould, J. & Nissen-Petersen, E. 1999. *Rainwater catchment systems for domestic supply.* London: Intermediate Technology Publications.

Gourmelon, M., Caprais, M., Mieszkin, S., Marti, R., Wery, N., Jardé, E., Derrien, M., Jadas-Hécart, A., Communal, P.Y., Jaffrezic, A. & Pourcher, A.M. 2010. Development of microbial and

chemical MST tools to identify the origin of the faecal pollution in bathing and shellfish harvesting waters in France. *Water Research.* 44(16):4812-4824.

Gregor, J., Garrett, N., Gilpin, B., Randall, C. & Saunders, D. 2002. Use of classification and regression tree (CART) analysis with chemical faecal indicators to determine sources of contamination. *New Zealand Journal of Marine and Freshwater Research.* 36(2):387-398.

Hagedorn, C. & Weisberg, S.B. 2009. Chemical-based faecal source tracking methods: current status and guidelines for evaluation. *Reviews in Environmental Science and Bio/Technology.* 8(3):275-287.

Hamza, I.A., Jurzik, L., Überla, K. & Wilhelm, M. 2011. Evaluation of Pepper Mild Mottle Virus, human picornavirus and Torque Teno Virus as indicators of faecal contamination in river water. *Water Research.* 45(3):1358-1368.

Haramoto, E., Kitajima, M., Kishida, N., Konno, Y., Katayama, H., Asami, M. & Akiba, M. 2013. Occurrence of Pepper Mild Mottle Virus in drinking water sources in Japan. *Applied and Environmental Microbiology.* 79(23):7413-7418.

Harwood, V.J., Staley, C., Badgley, B.D., Borges, K. & Korajkic, A. 2014. Microbial source tracking markers for detection of faecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews.* 38(1):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(6):3163-3170.

Hayashi, Y., Managaki, S. & Takada, H. 2002. Fluorescent whitening agents in Tokyo Bay and adjacent rivers: Their application as anthropogenic molecular markers in coastal environments. *Environmental Science & Technology.* 36(16):3556-3563.

Heath, R.J., Rubin, J.R., Holland, D.R., Zhang, E., Snow, M.E. & Rock, C.O. 1999. Mechanism of triclosan inhibition of bacterial fatty acid synthesis. *The Journal of Biological Chemistry.* 274(16):11110-11114.

Heberer, T., Reddersen, K. & Mechlinski, A. 2002. From municipal sewage to drinking water: Fate and removal of pharmaceutical residues in the aquatic environment in urban areas. *Water Science & Technology.* 46(3):81-88.

Heidler, J., Sapkota, A. & Halden, R.U. 2006. Partitioning, persistence, and accumulation in digested sludge of the topical antiseptic triclocarban during wastewater treatment. *Environmental Science & Technology.* 40(11):3634-3639.

Hsu, F.C., Shieh, Y.S., van Duin, J., Beekwilder, M.J. & Sobsey, M.D. 1995. Genotyping male-specific RNA coliphages by hybridization with oligonucleotide probes. *Applied and Environmental Microbiology.* 61(11):3960-3966.

Hundesa, A., Bofill-Mas, S., de Motes, C.M., Rodriguez-Manzano, J., Bach, A., Casas, M. & Girones, R. 2010. Development of a quantitative PCR assay for the quantitation of bovine polyomavirus as a microbial source tracking tool. *Journal of Virological Methods.* 163(2):385-389.

Hundesa, A., Maluquer de Motes, C., Bofill-Mas, S., Albinana-Gimenez, N. & Girones, R. 2006. Identification of human and animal adenoviruses and polyomaviruses for determination of sources of faecal contamination in the environment. *Applied and Environmental Microbiology.* 72(12):7886-7893.

Huston, R., Chan, Y., Chapman, H., Gardner, T. & Shaw, G. 2012. Source apportionment of heavy metals and ionic contaminants in rainwater tanks in a subtropical urban area in Australia. *Water Research.* 46(4):1121-1132.

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

Jett, B.D., Huycke, M.M. & Gilmore, M.S. 1994. Virulence of enterococci. *Clinical Microbiology Reviews.* 7(4):462-478.

Jimenez-Clavero, M.A., Fernandez, C., Ortiz, J.A., Pro, J., Carbonell, G., Tarazona, J.V., Roblas, N. & Ley, V. 2003. Teschoviruses as indicators of porcine faecal contamination of surface water. *Applied and Environmental Microbiology.* 69(10):6311-6315.

Jofre, J., Blanch, A.R., Lucena, F. & Muniesa, M. 2014. Bacteriophages infecting *Bacteroides* as a marker for microbial source tracking. *Water Research.* 55:1-11.

Johnston, C., Byappanahalli, M.N., Gibson, J.M., Ufnar, J.A., Whitman, R.L. & Stewart, J.R. 2013. Probabilistic analysis showing that a combination of *Bacteroides* and *Methanobrevibacter* source tracking markers is effective for identifying waters contaminated by human faecal pollution. *Environmental Science & Technology.* 47(23):13621-13628.

Johnston, C., Ufnar, J.A., Griffith, J.F., Gooch, J.A. & Stewart, J.R. 2010. A real-time qPCR assay for the detection of the *nif*H gene of *Methanobrevibacter smithii*, a potential indicator of sewage pollution. *Journal of Applied Microbiology.* 109(6):1946-1956.

Jones, M.E., Draghi, D.C., Thornsberry, C., Karlowsky, J.A., Sahm, D.F. & Wenzel, R.P. 2004. Emerging resistance among bacterial pathogens in the intensive care unit – a European and North American surveillance study (2000-2002). *Annals of Clinical Microbiology and Antimicrobials.* 3(1):1.

Jones, M.P. & Hunt, W.F. 2010. Performance of rainwater harvesting systems in the South-Eastern United States. *Resources, Conservation and Recycling.* 54(10):623-629.

Jones, R.D., Jampani, H.B., Newman, J.L. & Lee, A.S. 2000. Triclosan: A review of effectiveness and safety in health care settings. *American Journal of Infection Control.* 28(2):184-196.

Jones, M.S., Harrach, B., Ganac, R.D., Gozum, M.M., Dela Cruz, W.P., Riedel, B., Pan, C., Delwart, E.L. & Schnurr, D.P. 2007. New adenovirus species found in a patient presenting with gastroenteritis. *Journal of Virology.* 81(11):5978-5984.

Jones-Lepp, T. 2006. Chemical markers of human waste contamination: Analysis of urobilin and pharmaceuticals in source waters. *Journal of Environmental Monitoring.* 8(4):472-478.

Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B. & 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(16):3701-3715.

Kitajima, M., Iker, B.C., Pepper, I.L. & Gerba, C.P. 2014. Relative abundance and treatment reduction of viruses during wastewater treatment processes—identification of potential viral indicators. *Science of the Total Environment.* 488:290-296.

Kobayashi, A., Sano, D., Hatori, J., Ishii, S. & Okabe, S. 2013. Chicken- and duck-associated *Bacteroides–Prevotella* genetic markers for detecting faecal contamination in environmental water. *Applied Microbiology and Biotechnology.* 97(16):7427-7437.

Kuroda, K., Nakada, N., Hanamoto, S., Inaba, M., Katayama, H., Do, A.T., Nga, T.T.V., Oguma, K., Hayashi, T. & Takizawa, S. 2015. Pepper Mild Mottle Virus as an indicator and a tracer of fecal pollution in water environments: Comparative evaluation with wastewater-tracer pharmaceuticals in Hanoi, Vietnam. *Science of the Total Environment.* 506:287-298.

La Rosa, G., Della Libera, S., Petricca, S., Iaconelli, M., Briancesco, R., Paradiso, R., Semproni, M., Di Bonito, P. & Bonadonna, L. 2015. First detection of papillomaviruses and polyomaviruses in swimming pool waters: Unrecognized recreational water-related pathogens. *Journal of Applied Microbiology.* 119(6):1683-1691.

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 faecal pollution in water. *Applied and Environmental Microbiology.* 72(6):4214-4224.

Leeming, R., Ball, A., Ashbolt, N. & Nichols, P. 1996. Using faecal sterols from humans and animals to distinguish faecal pollution in receiving waters. *Water Research.* 30(12):2893-2900.

Lemarchand, K. & Lebaron, P. 2003. Occurrence of *Salmonella* spp. and *Cryptosporidium* spp. in a French coastal watershed: Relationship with faecal indicators. *FEMS Microbiology Letters.* 218(1):203-209.

Levesque, B., Pereg, D., Watkinson, E., Maguire, J., Bissonnette, L., Gingras, S., Rouja, P., Bergeron, M. & Dewailly, E. 2008. Assessment of microbiological quality of drinking water from household tanks in Bermuda. *Canadian Journal of Microbiology.* 54(6):495-500.

Li, Z., Boyle, F. & Reynolds, A. 2010. Rainwater harvesting and greywater treatment systems for domestic application in Ireland. *Desalination.* 260(1):1-8.

Lin, C. & Miller, T.L. 1998. Phylogenetic analysis of *Methanobrevibacter* isolated from faeces of humans and other animals. *Archives of Microbiology.* 169(5):397-403.

Lopes, Maria de Fátima Silva, Ribeiro, T., Abrantes, M., Marques, J.J.F., Tenreiro, R. & Crespo, M.T.B. 2005. Antimicrobial resistance profiles of dairy and clinical isolates and type strains of enterococci. *International Journal of Food Microbiology.* 103(2):191-198.

Lund, V. 1996. Evaluation of *E. coli* as an indicator for the presence of *Campylobacter jejuni* and *Yersinia enterocolitica* in chlorinated and untreated oligotrophic lake water. *Water Research.* 30(6):1528-1534.

Malema, S., Abia, L.K.A., Mwenge Kahinda, J. & Ubomba-Jaswa, E. 2016. Gaining a better understanding of the factors that influence the quality of harvested rainwater in South Africa – a review. *WISA 2016 Biennial Conference*. 16 -19 May 2016. Durban, South Africa.

Maluquer de Motes, C., Clemente-Casares, P., Hundesa, A., Martin, M. & Girones, R. 2004. Detection of bovine and porcine adenoviruses for tracing the source of faecal contamination. *Applied and Environmental Microbiology.* 70(3):1448-1454.

Martinson, B. & Thomas, T. 2005. Quantifying the first-flush phenomenon. *Proceedings of the 12th International Rainwater Catchment Systems Conference*. November 2005. New Delhi, India.

McLellan, S.L., Newton, R.J., Vandewalle, J.L., Shanks, O.C., Huse, S.M., Eren, A.M. & Sogin, M.L. 2013. Sewage reflects the distribution of human faecal Lachnospiraceae. *Environmental Microbiology.* 15(8):2213-2227.

McQuaig, S., Griffith, J. & Harwood, V.J. 2012. Association of faecal indicator bacteria with human viruses and microbial source tracking markers at coastal beaches impacted by nonpoint source pollution. *Applied and Environmental Microbiology.* 78(18):6423-6432.

McQuaig, S.M., Scott, T.M., Harwood, V.J., Farrah, S.R. & Lukasik, J.O. 2006. Detection of human-derived faecal pollution in environmental waters by use of a PCR-based human polyomavirus assay. *Applied and Environmental Microbiology.* 72(12):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 faecal samples. *Applied and Environmental Microbiology.* 75(11):3379-3388.

Meays, C.L., Broersma, K., Nordin, R. & Mazumder, A. 2004. Source tracking faecal bacteria in water: A critical review of current methods. *Journal of Environmental Management.* 73(1):71-79.

Meehan, C.J. & Beiko, R.G. 2014. A phylogenomic view of ecological specialization in the Lachnospiraceae, a family of digestive tract-associated bacteria. *Genome Biology and Evolution.* 6(3):703-713.

Miller, T.L. & Wolin, M.J. 1983. Stability of *Methanobrevibacter smithii* populations in the microbial flora excreted from the human large bowel. *Applied and Environmental Microbiology.* 45(1):317-318.

Mohamed, J.A. & Huang, D.B. 2007. Biofilm formation by enterococci. *Journal of Medical Microbiology.* 56(12):1581-1588.

Muscillo, M., Pourshaban, M., Iaconelli, M., Fontana, S., Di Grazia, A., Manzara, S., Fadda, G., Santangelo, R., La Rosa, G. 2008. Detection and quantification of human adenoviruses in surface waters by nested PCR, TaqMan real-time PCR and cell culture assays. *Water, Air, & Soil Pollution.* 191(1-4):83-93.

Mwenge Kahinda, J. & Taigbenu, A.E. 2011. Rainwater harvesting in South Africa: Challenges and opportunities. *Physics & Chemistry of the Earth - Parts A/B/C.* 36(14):968-976.

Mwenge Kahinda, J., Taigbenu, A. & Boroto, R. 2010. Domestic rainwater harvesting as an adaptation measure to climate change in South Africa. *Physics and Chemistry of the Earth, Parts A/B/C.* 35(13):742-751.

Mwenge Kahinda, J., Taigbenu, A.E. & Boroto, J.R. 2007. Domestic rainwater harvesting to improve water supply in rural South Africa. *Physics and Chemistry of the Earth, Parts A/B/C.* 32(15):1050-1057.

Nakada, N., Kiri, K., Shinohara, H., Harada, A., Kuroda, K., Takizawa, S. & Takada, H. 2008. Evaluation of pharmaceuticals and personal care products as water-soluble molecular markers of sewage. *Environmental Science & Technology.* 42(17):6347-6353.

Newton, R.J., Vandewalle, J.L., Borchardt, M.A., Gorelick, M.H. & McLellan, S.L. 2011. Lachnospiraceae and *Bacteroidales* alternative faecal indicators reveal chronic human sewage contamination in an urban harbor. *Applied and Environmental Microbiology.* 77(19):6972-6981.

Noble, R.T., Allen, S.M., Blackwood, A.D., Chu, W., Jiang, S.C., Lovelace, G.L., Sobsey, M.D., Stewart, J.R. & Wait, D.A. 2003. Use of viral pathogens and indicators to differentiate between human and non-human faecal contamination in a microbial source tracking comparison study. *Journal of Water and Health.* 1(4):195-207.

Okabe, S., Okayama, N., Savichtcheva, O. & Ito, T. 2007. Quantification of host-specific *Bacteroides–Prevotella* 16S rRNA genetic markers for assessment of faecal pollution in freshwater. *Applied Microbiology and Biotechnology.* 74(4):890-901.

Orhan, M., Kut, D. & Gunesoglu, C. 2009. Improving the antibacterial activity of cotton fabrics finished with triclosan by the use of 1, 2, 3, 4-butanetetracarboxylic acid and citric acid. *Journal of Applied Polymer Science.* 111(3):1344-1352.

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(4):351-356.

Rosario, K., Symonds, E.M., Sinigalliano, C., Stewart, J. & Breitbart, M. 2009. Pepper Mild Mottle Virus as an indicator of faecal pollution. *Applied and Environmental Microbiology.* 75(22):7261-7267.

Rusinõl, M., Fernadez-Cassi, X., Hundesa, A., Vieira, C., Kern, A., Eriksson, I., Ziros, P., Kay, D., Miagostovich, M., Vargha, M., Vantarakis, A., Wyn-Jones, P., Bofill-Mas, S. & Girones, R. 2014. Application of human and animal viral microbial source tracking tools in fresh and marine waters from five different geographical areas. *Water Research.* 59: 119-129

Sauer, E.P., VandeWalle, J.L., Bootsma, M.J. & McLellan, S.L. 2011. Detection of the human-specific *Bacteroides* genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. *Water Research.* 45(14):4081-4091.

Scheurer, M., Brauch, H. & Lange, F.T. 2009. Analysis and occurrence of seven artificial sweeteners in German waste water and surface water and in soil aquifer treatment (SAT). *Analytical and Bioanalytical Chemistry.* 394(6):1585-1594.

Schill, W.B. & Mathes, M.V. 2008. Real-time PCR detection and quantification of nine potential sources of faecal contamination by analysis of mitochondrial cytochrome B targets. *Environmental Science & Technology.* 42(14):5229-5234.

Scott, T.M., Jenkins, T.M., Lukasik, J. & Rose, J.B. 2005. Potential use of a host associated Molecular marker in *Enterococcus faecium* as an index of human fecal pollution. *Environmental Science & Technology.* 39(1):283-287.

Sercu, B., Van De Werfhorst, Laurie C, Murray, J.L. & Holden, P.A. 2011. Terrestrial sources homogenize bacterial water quality during rainfall in two urbanized watersheds in Santa Barbara, CA. *Microbial Ecology.* 62(3):574-583.

Seurinck, S., Defoirdt, T., Verstraete, W. & Siciliano, S.D. 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(2):249-259.

Shankar, V., Baghdayan, A.S., Huycke, M.M., Lindahl, G. & Gilmore, M.S. 1999. Infection-derived *Enterococcus faecalis* strains are enriched in *esp*, a gene encoding a novel surface protein. *Infection and Immunity.* 67(1):193-200.

Shanks, O.C., White, K., Kelty, C.A., Sivaganesan, M., Blannon, J., Meckes, M., Varma, M. & Haugland, R.A. 2010. Performance of PCR-based assays targeting Bacteroidales genetic markers of human faecal pollution in sewage and faecal samples. *Environmental Science & Technology.* 44(16):6281-6288.

Sidhu, J., 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.* 463:488-496.

Simmons, G., Hope, V., Lewis, G., Whitmore, J. & Gao, W. 2001. Contamination of potable roof-collected rainwater in Auckland, New Zealand. *Water Research.* 35(6):1518-1524.

Simmons, G., Jury, S., Thornley, C., Harte, D., Mohiuddin, J. & Taylor, M. 2008. A Legionnaires' Disease outbreak: A water blaster and roof-collected rainwater systems. *Water Research.* 42(6):1449-1458.

Simpson, J.M., Santo Domingo, J.W. & Reasoner, D.J. 2002. Microbial source tracking: State of the science. *Environmental Science & Technology.* 36(24):5279-5288.

Staggemeier, R., Bortoluzzi, M., da Silva Heck, Tatiana Moraes, da Luz, R.B., Fabres, R.B., Soliman, M.C., Rigotto, C., Baldasso, N.A., Spilki, F.R. & Almeida, S.E.M. 2015. Animal and human enteric viruses in water and sediment samples from dairy farms. *Agricultural Water Management.* 152:135-141.

Steffen, J., Jensen, M., Pomeroy, C.A. & Burian, S.J. 2013. Water supply and stormwater management benefits of residential rainwater harvesting in US cities. *Journal of the American Water Resources Association.* 49(4):810-824.

Stewart-Pullaro, J., Daugomah, J., Chestnut, D., Graves, D., Sobsey, M. & Scott, G. 2006. F^+ RNA coliphage typing for microbial source tracking in surface waters. *Journal of Applied Microbiology.*101(5):1015-1026.

Stoeckel, D.M. & Harwood, V.J. 2007. Performance, design, and analysis in microbial source tracking studies. *Applied and Environmental Microbiology.* 73(8):2405-2415.

Uba, B.N. & Aghogho, O. 2000. Rainwater quality from different roof catchments in the Port Harcourt District, Rivers State, Nigeria. *Aqua.* 49:281-288.

Ufnar, J., Wang, S.Y., Christiansen, J., Yampara-Iquise, H., Carson, C. & Ellender, R. 2006. Detection of the *nif*H gene of *Methanobrevibacter smithii*: A potential tool to identify sewage pollution in recreational waters. *Journal of Applied Microbiology.* 101(1):44-52.

United Nations (UN). 2015a. *The Millenium Development Goals Report 2015*. Available: http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20(July%201).pdf [2015, June 05]

United Nations (UN). 2015b. *Sustainable development goals - water and sanitation.* Available: http://www.un.org/sustainabledevelopment/water-and-sanitation/ [2015, June 05]

Uygur, N., Karaca, F. & Alagha, O. 2010. Prediction of sources of metal pollution in rainwater in Istanbul, Turkey using factor analysis and long-range transport models. *Atmospheric Research.* 95(1):55-64.

Van Heerden, J., Ehlers, M.M., Heim, A. & Grabow, W.O. 2005. Prevalence, quantification and typing of adenoviruses detected in river and treated drinking water in South Africa. *Journal of Applied Microbiology.* 99(2):234-242.

Van Heerden, J., Ehlers, M.M., Van Zyl, W.B. & Grabow, W.O. 2003. Incidence of adenoviruses in raw and treated water. *Water Research.* 37(15):3704-3708.

Verrinder, G. & Keleher, H. 2001. Domestic drinking water in rural areas: Are water tanks on farms a health hazard. *Environmental Health.* 1(3):51.

Villemur, R., Imbeau, M., Vuong, M.N., Masson, L. & Payment, P. 2015. An environmental survey of surface waters using mitochondrial DNA from human, bovine and porcine origin as faecal source tracking markers. *Water Research.* 69:143-153.

Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K. & Whitman, W.B. 2009. The Firmicutes. In *Bergey's Manual of Systematic Bacteriology: Volume 3*. New York: Springer.

Waso, M., Ndlovu, T., Dobrowsky, P.H., Khan, S. & Khan, W. 2016. Presence of microbial and chemical source tracking markers in roof-harvested rainwater and catchment systems for the detection of fecal contamination. *Environmental Science and Pollution Research*. 23 (17), 16987–17001.

Water Research Commission (WRC). 2014. *Domestic rainwater harvesting: Survey of perceptions of users in Kleinmond* (WRC Project K5/2124/3 Report). Pretoria, South Africa.

World Health Organisation/United Nations International Children's Emergency Fund (WHO/UNICEF). 2015. *World Health Organisation (WHO)/United Nations International Children's Emergency Fund (UNICEF) Progress on Sanitation and Drinking Water – 2015 Update* and *MDG Assessment*. Available: http://files.unicef.org/publications/files/Progress_on_Sanitation_and_Drinking_Water_2015_Update_.pdf [2015, June 05].

Wolf, S., Hewitt, J. & Greening, G.E. 2010. Viral multiplex quantitative PCR assays for racking sources of faecal contamination. *Applied and Environmental Microbiology.* 76(5):1388-1394.

Yalavarthy, D.P., Kumar, K.R., Rohini, P. & Md, A.R. 2015. A review on occurrence, fate and toxicity of triclosan. *World Journal of Pharmacy and Pharmaceutical Sciences*. 4(7):336-369.

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

Zhang, T., Breitbart, M., Lee, W.H., Run, J., Wei, C.L., Soh, S.W.L., Hibberd, M.L., Liu, E.T., Rohwer, F. & Ruan, Y. 2005. RNA viral community in human faeces: Prevalence of plant pathogenic viruses. *PLoS Biology.* 4(1):e3.

Chapter 2:

(Chapter 2 is compiled in the format of the Science of the Total Environment journal and US spelling is employed)

Primary Microbial and Chemical Source Tracking Markers Associated with Domestic Rainwater Harvesting Systems: Correlation to Indicator Organisms

Waso, M.[1], Ndlovu, T. [1], Dobrowsky, P.H. [1], Khan, S.[2] and Khan, W.[1]*

[1] Department of Microbiology, Faculty of Science, Stellenbosch University, Private Bag X1, Stellenbosch, 7602, South Africa.

[2] Faculty of Health and Applied Sciences, Namibia University of Science and Technology, 13 Storch Street, Private Bag 13388, Windhoek, Namibia.

Abstract

Stored harvested rainwater (tank water) and rooftop debris samples from domestic rainwater harvesting (DRWH) systems were screened for the presence of a panel of microbial (MST) and chemical source tracking (CST) markers in order to determine which markers could be utilized to supplement traditional indicator analysis. The tank water samples were thus also screened for traditional indicator organisms utilizing culture based analysis. Utilizing conventional PCR, the MST markers most frequently detected were *Bacteroides* HF183, adenovirus, Lachnospiraceae and human mtDNA. The most readily detected MST markers were quantified utilizing quantitative PCR (qPCR) assays with the HF183 marker detected at a mean concentration of 5.1×10^3 and 4.7×10^3 gene copies/µL, adenovirus detected at a mean concentration of 3.2×10^2 and 6.4×10^3 gene copies/µL, Lachnospiraceae detected at a mean concentration of 3.0×10^4 and 6.9×10^3 gene copies/µL and human mtDNA detected at 1.1×10^6 (90%) and 3.0×10^5 gene copies/µL (91.7%) in the tank water and rooftop debris samples, respectively. The concentrations of the CST markers (caffeine, salicylic acid, methylparaben, triclosan, triclocarban, acetaminophen and carbamazepine) were determined utilizing high-performance liquid chromatography tandem mass spectrometry analysis. All the CST markers (except carbamazepine) were then detected at µg/L levels in the tank water and rooftop debris samples. The quantitative data was subjected to statistical analysis in order to identify correlations between the various indicator organisms and source tracking markers detected. In the tank water samples, significant positive correlations were observed for adenovirus versus *Eschericia coli* (*E. coli*) (enumerated with the culturing techniques) ($r = 0.983$; $p = 0.000$), the HF183 marker versus *E. coli* (quantified by qPCR) ($r = 0.303$; $p = 0.023$), Lachnospiraceae versus heterotrophic bacteria ($r = 0.682$; $p = 0.000$) and human mtDNA versus enterococci (enumerated with the culturing techniques) ($r = 0.297$; $p = 0.026$). In addition, significant positive correlations were observed for caffeine versus enterococci (quantified by qPCR) ($r = 0.863$; $p = 0.000$); fecal coliforms ($r = 0.447$; $p = 0.001$); total coliforms ($r = 0.483$; $p = 0.000$) and enterococci (enumerated with culturing techniques) ($r = 0.399$; $p = 0.002$), respectively. Salicylic acid also positively correlated with total coliforms ($r = 0.300$; $p = 0.024$) in the tank water samples. For the rooftop debris samples, significant positive correlations were then observed for *E. coli* (quantified by qPCR) versus methylparaben ($r = 0.623$; $p = 0.000$) and salicylic acid ($r = 0.273$; $p = 0.042$). Based on the correlations observed for the MST and CST markers versus the indicator organisms, the HF183 marker, Lachnospiraceae, human mtDNA, adenovirus, caffeine, salicylic acid and methylparaben may be utilized to supplement traditional indicator organism analysis to monitor the quality of harvested rainwater.

Keywords: Indicator organisms; microbial source tracking; chemical source tracking; correlation analysis; harvested rainwater

2.1 Introduction

Indicator organisms are utilized globally to monitor water quality and predict the presence of pathogens in contaminated environmental waters. However, there is growing evidence that most indicator organisms are capable of proliferating in water sources (Field & Samadpour, 2007) and certain strains have been shown to survive, grow and establish populations in other natural environments such as plant cavities, algal mats, beach sands, soils and sediments (Fujioka et al., 1998; Solo-Gabriele et al., 2000; Whitman et al., 2003; Byappanahalli & Fujioka, 2004; Anderson et al., 2005; Whitman et al., 2005; Byappanahalli et al., 2006a; 2006b; Ishii et al., 2006; Olapade et al., 2006; Field & Samadpour, 2007). Moreover, research has indicated that the presence of indicator organisms generally exhibits a poor correlation with the presence of pathogens in contaminated water and the detection of these organisms does not provide information on the specific sources of fecal contamination in water bodies (Harwood et al., 2005; Field & Samadpour, 2007; Harwood et al., 2014). To compensate for these pitfalls, it is essential that alternative indicators be identified to supplement current indicator organism analysis. These supplementary indicators could be utilized to monitor for fecal contamination in environmental waters. This may in turn indicate the potential health risk associated with the use of the water source, as contaminated waters may harbor pathogens such as enteric viruses, *Salmonella enterica*, and *Pseudomonas aeruginosa* (Liang et al., 2015), amongst other organisms.

Microbial (MST) and chemical source tracking (CST) markers are being investigated as potential alternative indicators of water quality (Field & Samadpour, 2007; Harwood et al., 2014). The premise of source tracking is that certain fecal microorganisms may be strongly associated with specific hosts and may therefore be employed to indicate host-specific contamination of environmental waters (Harwood et al., 2014). For CST analysis, chemical compounds associated with household waste, anthropogenic activity and wastewater may then be utilised to identify these sources of contamination in environmental waters. A few common microbial and chemical source tracking markers include: human-specific *Bacteroides* HF183 and *Methanobrevibacter smithii nif*H (*M. smithii nif*H) (Seurinck et al., 2005; Ufnar et al., 2006; Sercu et al., 2011; Sidhu et al., 2013); *Bifidobacterium* spp. (Gourmelon et al., 2010); the *Enterococcus esp* gene (Ahmed et al., 2008a); human adenovirus and polyomavirus (Muscillo et al., 2008; Sauer et al., 2011; Sidhu et al., 2013); enterovirus (Wolf et al., 2010); compounds associated with pharmaceuticals and personal care products such as acetaminophen and salicylic acid (Hagedorn & Weisberg, 2009; Sidhu et al., 2013); fecal sterols and stanols (metabolic by-products of cholesterol); optical brighteners and caffeine (Hagedorn & Weisberg, 2009). These source tracking markers have diverse applications ranging from monitoring beach water quality (Brownell et al., 2007; Abdelzaher et al., 2010; Korajkic et al., 2011; Harwood et al., 2014) to food quality and have potential applications in the legal arena (for example

identifying sources of untreated human waste discharge into environmental waters) (Harwood et al., 2014). In addition, source tracking markers have been applied to rivers (Seurinck et al., 2005; Kobayashi et al., 2013), lakes (Jones-Lepp, 2006; Ahmed et al., 2010b), seawater (Muscillo et al., 2008) and stormwater run-off (Sidhu et al., 2013), amongst others, to identify the dominant contamination sources. In a study conducted by Staley et al. (2016) the MST markers, general (GenBactF3) *Bacteroides* as well as human (HF183), ruminant (CF128) and canine (DG37) *Bacteroides* and a gull marker (*Catelicoccus marimammalium*– qGull4), were utilized in combination with the anthropogenic-associated CST markers caffeine, carbamazepine, cotinine, codeine, acetaminophen and acesulfame, to monitor the quality of the Humber River watershed in Toronto, Canada. Based on the detection of the human-associated markers in the river water samples, it was concluded that sewage was the major source of fecal pollution of the watershed. In addition, Kirs et al. (2016) investigated the concentrations of human polyomavirus and *Bacteroides* HF183 in streams and marine water in Oahu, Hawaii. Based on the results obtained, it was concluded that sewage was also the main source of the HF183 marker and human polyomavirus detected in the marine water and that the streams and beaches may thus be impacted by anthropogenic activities.

Furthermore, identifying source tracking markers that correlate well with both waterborne pathogens and indicator organisms may improve their predictive capability in indicating fecal contamination and the presence of pathogens in a water source (Harwood et al., 2014; Liang et al., 2015). Bradshaw et al. (2016) applied the general (GenBac), ruminant (CowM3 and Rum-2-bac) and human *Bacteroides* (HF183) MST markers to river water and sediment samples collected from the South Fork Broad River in Georgia, United States of America (USA). Correlations between these markers and *E. coli*, *Listeria*, *Campylobacter*, *Salmonella* and the specific virulence gene encoding for the Shiga toxin (stx_2) were subsequently determined. This study then detected significant positive correlations between the ruminant MST markers and the Shiga toxin gene. It was deduced that the presence of the Shiga toxin gene could be attributed to agricultural land use as cattle pastures were observed around the river system. *Campylobacter* was also positively correlated with the ruminant markers and thus the presence of the *Campylobacter* in the river water was once again attributed to the agricultural activities along the river system. In addition, *Listeria* positively correlated with the human-associated HF183 marker and this correlation indicated that the presence of the *Listeria* in the river water could possibly be attributed to sewage contamination. The authors also noted that other MST markers should be included in future screenings of the river water samples in order to fully elucidate all the sources of the pathogens detected in the river water samples (Bradshaw et al., 2016).

Rainwater harvesting (RWH) is currently being utilized worldwide as an alternative fresh water source, however numerous studies have indicated that the microbial quality of harvested

rainwater does not adhere to drinking water guidelines as indicator organisms and various pathogens have been detected in stored rainwater (Crabtree et al., 1996; Verrinder & Keleher, 2001; Handia, 2005; Field & Samadpour, 2007; Ahmed et al., 2008b; Simmons et al., 2008; Despins et al., 2009; Ahmed et al., 2010a; 2011; 2012; Dobrowsky et al., 2014a; 2014b; 2014c). Limited research on the application of source tracking markers for the screening of RWH systems is however, available (Ahmed et al., 2016; Waso et al., 2016). Correlations between source tracking markers and indicator organisms in harvested rainwater have also not been extensively studied. The aim of the current study was thus to identify the primary MST and CST markers associated with RWH catchment systems. A secondary aim was to identify correlations between the primary source tracking markers detected and indicator organisms. This was performed in order to identify the markers that may be utilized to supplement conventional indicator organism analysis in future screenings of RWH systems. These aims were achieved by completing the following objectives: i) screen harvested rainwater and rooftop debris from the catchment surface (rooftops) for a panel of MST markers that have been shown in literature to exhibit host-specific distributions; ii) screen harvested rainwater and rooftop debris for a panel of CST markers that have been shown to be associated with anthropogenic activities, iii) screen the harvested rainwater for indicator organisms; *E. coli*, enterococci, total and fecal coliforms and heterotrophic bacteria, utilizing traditional culturing techniques, iv) optimize and apply quantitative real-time PCR (qPCR) assays for the quantification of the frequently detected MST markers, *E. coli* and enterococci in the harvested rainwater and rooftop debris samples and v) perform statistical analysis to identify correlations between the source tracking markers and indicator organisms in the RWH systems.

2.2 Materials and Methods

2.2.1 Sampling Site

All harvested rainwater (henceforth referred to as tank water samples) and rooftop debris (e.g. dust, fecal material and soil) samples were collected from ten domestic rainwater harvesting (DRWH) systems connected to ten houses (**Fig. 2.1.** – red markers denote houses utilized in the current study) located in the Kleinmond Housing Scheme site in Kleinmond, a peri-urban coastal town situated in the Western Cape, South Africa (GPS co-ordinates: 34°20’11.81"S 19°00’59.74"E). The ten houses were selected from a pool of houses utilized in previous studies conducted by Dobrowsky et al. (2014a, b, c) The Kleinmond Housing Scheme site consists of a cluster of 411 pilot scale houses (40 m^2 each) which were all fitted with DRWH tanks, photovoltaic cells and solar geysers to investigate technologies that will improve the sustainability and quality of low-income subsidized housing in South Africa [Council for Scientific and Industrial Research (CSIR), 2011; Dobrowsky et al., 2014a]. The DRWH tanks connected to the houses in Kleinmond have a capacity of 2000 L and no first flush diverters were installed.

Fig. 2.1. Aerial image of the Kleinmond Housing Scheme site (GPS co-ordinates: 34°20’11.81"S 19°00’59.74"E) with the locations of the sampled DRWH tanks and corresponding gutter systems indicated by red markers.

Sampling was conducted one to four days after a rain event during low (October 2015; April 2016) and high (August 2015; September 2015; March 2016; June 2016) rainfall periods. Tank water samples were collected in 5 L sterile polypropylene bottles which were rinsed with distilled water, sterilized with 70% ethanol and dried overnight in an oven at 60 °C. Rooftop debris samples (~20 g) were collected in sterile Falcon™ 50 mL high-clarity polypropylene tubes. After collection, all samples were stored at 4 °C until further processing.

In total, six sampling occasions were conducted with data on the total rainfall and ambient temperature for each sampling session obtained from the South African Weather Services (Pretoria, South Africa). Thus, while ten tank water and ten rooftop debris samples were collected per sampling occasion, a total of 60 tank water samples and 60 rooftop debris samples were collected overall. The pH and temperature of the collected rainwater samples were measured using a Martini Instruments pH55 meter (Milwuakee Instruments Inc., Rocky Mount, USA) on site during sample collection.

In addition, three municipal tap water samples (2 L) were collected during sampling occasions three and four, while four municipal tap water samples were collected during sampling occasion five from various houses utilized as study sites (**Fig. 2.1.**). These samples were collected to

compare the concentrations of the CST markers detected in the tank water samples to the concentrations of the CST markers detected in the municipal tap water samples (Section 2.2.4.1).

Henceforth, the collected tank water samples will be denoted by the letter T followed by letters which indicate the specific tank (1 to 10 represented by A to J) sampled, while the sampling occasion is indicated by the numbers 1 to 6. For example, tank water samples collected during sampling four from tank B will be denoted as TB4. The same rule will apply to the rooftop debris samples, however rooftop debris will be denoted by the indicator RD. Lastly, all the municipal tap water samples will be denoted as MT followed by a number indicating the sampling session during which the samples were collected i.e. either 3 indicating sampling three or 4 indicating sampling four, etc. An additional number is included indicating the specific sample collected from a municipal tap during the sampling occasion. For example, three municipal tap water samples were collected during sampling three and will be denoted as MT3.1, MT3.2 and MT3.3.

2.2.2 Enumeration of Indicator Organisms

Various culture media and incubation conditions were utilized to enumerate *E. coli*, total coliforms, fecal coliforms, enterococci and heterotrophic bacteria in all tank water samples collected during the entire sampling period. Each medium utilized was prepared according to the manufacturer's instructions. Samples were processed within 4 hours (h) of sample collection in order to accurately enumerate *E. coli* and total coliforms. It should be noted that the guidelines stipulated for indicator organism analysis by national and international organizations focus on water and water uses, whereas no guidelines or methods for indicator organism analysis on solid matrices, such as rooftop debris, are stipulated. Culture based analysis for indicator organisms in the rooftop debris samples was thus not conducted.

Membrane filtration was utilized to enumerate *E. coli* and total coliforms and the procedure was performed in duplicate. Briefly, for each tank water sample, 100 mL undiluted tank water and 10^{-1} and 10^{-2} diluted tank water were filtered through a sterile GN-6 Metricel® S-Pack membrane disc filter (Pall Life Sciences, Ann Harbor, USA) with a pore size of 0.45 µm and a diameter of 47 mm at a flow rate of approximately 65 mL/min/cm^2 at 70 kPa. The filters were then placed on Membrane Lactose Glucuronide Agar [MLGA (Oxoid, Hampshire, England)] and were incubated for 18 to 24 h at 37 °C [United States Environmental Protection Agency (US EPA), 2008].

To enumerate enterococci and fecal coliforms, 100 µL of undiluted tank water was spread plated onto Slanetz and Bartley agar (Oxoid, Hampshire, England) and m-FC agar (Merck, Kennilworth, USA), respectively, in duplicate. The plates were then incubated for 44 - 48 h and 22 - 24 h at 37 °C, for enterococci and fecal coliform enumeration, respectively.

To enumerate heterotrophic bacteria, 100 µL of undiluted, 10^{-1} and 10^{-2} diluted tank water were spread plated onto Reasoner's 2A (R2A) agar (BD Difco, Michigan, USA) in duplicate. The plates were incubated for up to 72 h at 37 °C.

2.2.3 Microbial Source Tracking

2.2.3.1 Concentration and Nucleic Acid Extractions from Tank Water Samples

One liter (1 L) of each 5 L tank water sample collected throughout the sampling period was concentrated as previously described by Waso et al. (2016) for whole DNA extractions. Briefly, 2 mL of 1 M calcium chloride ($CaCl_2$) and 2 mL of 1 M di-sodium hydrogen orthophosphate anhydrous (Na_2HPO_4) were added to each liter of tank water to allow for flocculation. The samples were then stirred on a magnetic stirrer for 5 min, and were filtered through a non-charged mixed ester membrane (0.45 µm; Merck, Millipore, Billerica, USA) at a flow rate of approximately 65 mL/min/cm^2 at 70 kPa. The membrane filter was then placed in 4 mL citrate buffer (0.3 M, pH 3.5) and was soaked for 3 min to remove cells from the membrane filters. The 4 mL concentrate from each tank water sample was centrifuged at 15 000 × *g* for 15 min and each pellet was subsequently resuspended in 200 µL tris ethylenediaminetetraacetic acid (TE) buffer (pH 8). Whole DNA was extracted from the resuspended pellets utilizing the Zymo Research Soil Microbe DNA MiniPrep™ kit (Zymo Research, Irvine, USA) as per manufacturer's instructions.

In addition, 1 L of each of the 5 L tank water sample collected throughout the sampling period, was concentrated for viral nucleic acid extractions as outlined above. However, the resuspended pellet from each tank water sample was subjected to viral nucleic acid extraction utilizing the QIAamp® Ultrasens® Virus Kit (Qiagen, Hilden, Germany) as per manufacturer's instructions. After the viral DNA/RNA extractions were completed, a 10 µL aliquot of each extract was stored at 4 °C for adenovirus and polyomavirus detection, and 8 µL aliquots were stored at -20 °C for enterovirus detection (Saayman et al., 2012).

2.2.3.2 Nucleic Acid Extractions from Rooftop Debris Samples

To remove cells and viral particles from the rooftop debris matrix, 10 g of each debris sample was sonicated for 10 min in 5 mL sterile distilled water in a Branson 5510 sonication bath (Bransonic® Ultrasonic Cleaner) (Jackson et al., 2009). After sonication, two aliquots (2 mL each) of the supernatant were transferred into two separate sterile Eppendorf tubes and these samples were centrifuged at 15 000 × *g* for 15 min. After centrifugation, the supernatant from each tube was discarded. This procedure was repeated three times for both Eppendorf tubes. The resulting pellets were then resuspended in 200 µL TE buffer (pH 8). Whole DNA was extracted from one of the resuspended pellets utilizing the Zymo Research Soil Microbe DNA

MiniPrep™ kit as per manufacturer's instructions. From the remaining resuspended pellet, viral nucleic acids were extracted utilizing the QIAamp® Ultrasens® Virus Kit as per manufacturer's instructions. After the viral DNA/RNA extractions, a 10 µL aliquot of each extract was stored at 4 °C for adenovirus and polyomavirus detection, and 8 µL aliquots were stored at -20 °C for enterovirus detection (Saayman et al., 2012).

2.2.3.3 Enterovirus cDNA Synthesis

Enterovirus is a single stranded RNA virus, and therefore complementary DNA (cDNA) was synthesized prior to the conventional PCR assays for the detection of this virus in the tank water and rooftop debris samples. Complementary DNA was synthesized as previously described by Saayman et al. (2012) utilizing the Improm-II™ Reverse Transcription System (Promega Corp, Madison, USA). Briefly, 4 µL of DNA/RNA, which was extracted using the QIAamp® Ultrasens® Virus Kit from each tank water and rooftop debris sample as outlined in Sections 2.2.3.1 and 2.2.3.2, respectively, was added to 1 µL EP1 primer (final concentration of 2 µM) (EP1: ATTGTCCACCATAAGCAGCCA; **Table 2.1**) and was subsequently denatured at 70 °C for 5 min. The denatured samples were then placed on ice for 5 min. In separate tubes, a reaction mixture containing: 3.75 mM $MgCl_2$, 0.5 µM dNTP mix, 1X Improm-II™ reaction buffer, 20 U RNasin ribonuclease inhibitor and 1 µL Improm-II™ Reverse Transcriptase was prepared to a final volume of 15 µL per reaction. The 5 µL denatured RNA-primer mix was then added to the 15 µL reaction mix for each sample and reverse transcription was completed at 42 °C for 60 min. The Improm-II™ Reverse Transcriptase enzyme was then inactivated at 70 °C for 15 min. The samples were immediately utilized for the enterovirus PCR assay (**Table 2.1**; Section 2.2.3.4) or stored at 4 °C until further analysis.

2.2.3.4 Conventional PCR Assays

Conventional PCR was performed on the nucleic acid extractions from all the collected tank water and rooftop debris samples with the respective primer sets and cycling parameters as outlined in **Table 2.1**. For the bacterial, mitochondrial and archaeal PCR assays, whole DNA extracted using the Zymo Research Soil Microbe DNA MiniPrep™ kit was utilized as template DNA, while for the adenovirus and polyomavirus assays, DNA extracted with the QIAamp® Ultrasens® Virus Kit was utilized as template DNA. Synthesized cDNA (Section 2.2.3.3) was then utilized as template for the nested PCR assay for the detection of enterovirus.

For the detection of the HF183 marker, each PCR mixture consisted of 1X Green GoTaq® Reaction Buffer (Promega Corp, Madison, USA), 2 mM $MgCl_2$, 0.2 mM dNTP mix, 0.25 µM of each primer (**Table 2.1**), 2.5 U GoTaq® Flexi DNA polymerase (Promega Corp, Madison, USA) and 4 µL of template DNA (Seurinck et al., 2005) in a final volume of 25 µL.

For the detection of *Bifidobacterium* spp., each PCR mixture consisted of 1X Green GoTaq® Reaction Buffer, 2 mM $MgCl_2$, 0.2 mM dNTP mix, 0.3 µM of each primer (**Table 2.1**), 1.5 U GoTaq® Flexi DNA polymerase and 2.5 µL of template DNA (Gourmelon et al., 2010) in a final volume of 25 µL.

For the detection of the *Enterococcus esp* marker, each PCR mixture contained 1X Green GoTaq® Reaction Buffer, 2 mM $MgCl_2$, 0.2 mM dNTP mix, 0.2 µM of each primer (**Table 2.1**), 1.5 U GoTaq® Flexi DNA polymerase and 2 µL of template DNA (Ahmed et al., 2008a) in a final volume of 25 µL.

For the detection of Lachnospiraceae, each PCR mixture consisted of 1X Green GoTaq® Reaction Buffer, 2 mM $MgCl_2$, 0.2 mM dNTP mix, 0.2 µM of each primer (**Table 2.1**), 1.5 U GoTaq® Flexi DNA polymerase and 2 µL of template DNA (Newton et al., 2011) in a final volume of 25 µL.

For the detection of *Methanobrevibacter* spp. and the *M. smithii nif*H marker, each PCR mixture consisted of 1X Green GoTaq® Reaction Buffer, 2 mM $MgCl_2$, 0.2 mM dNTP mix, 0.5 µM of each primer (**Table 2.1**), 0.5 U GoTaq® Flexi DNA polymerase and 2 µL of template DNA (Ufnar et al., 2006) in a final volume of 25 µL.

For the detection of polyomavirus, each PCR mixture consisted of 1X Green GoTaq® Reaction Buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTP mix, 0.5 µM of each primer (**Table 2.1**), 1.5 U GoTaq® Flexi DNA polymerase and 4 µL of template DNA (McQuaig et al., 2009) in a final volume of 25 µL.

Each PCR mixture for the detection of adenovirus consisted of 1X Green GoTaq® Reaction Buffer, 2.5 mM $MgCl_2$, 0.2 mM dNTP mix, 0.3 µM of each primer (**Table 2.1**), 1.25 U GoTaq® Flexi DNA polymerase and 6 µL of template DNA (Saayman et al., 2012) in a final volume of 25 µL.

For the detection of enterovirus, a nested PCR was utilized and each PCR mixture consisted of 1X Green GoTaq® Reaction Buffer, 3.6 mM $MgCl_2$, 0.2 mM dNTP mix, 0.5 µM of each primer (**Table 2.1**), 1 U GoTaq® Flexi DNA polymerase and 2 µL of cDNA in a final volume of 25 µL. For the first PCR, the EP1/EP2 primer pair was utilized with 2 µL synthesized cDNA (Section 2.2.3.3) and for the second PCR, the EP3/EP4 primer pair was utilized with 2 µL of product from the first PCR (Saayman et al., 2012).

For the detection of human, bovine and porcine mitochondrial DNA (mtDNA) markers, each PCR mixture consisted of 1X Green GoTaq® Reaction Buffer, 5.5 mM $MgCl_2$, 0.2 mM dNTP mix, 0.3 µM of each primer (**Table 2.1**), 1.5 U GoTaq® Flexi DNA polymerase and 4 µL of template DNA (Caldwell et al., 2007) in a final volume of 25 µL.

Table 2.1 Conventional PCR assay primers, cycling parameters and PCR product size of the MST markers screened for in the tank water and rooftop debris samples.

Microbial Marker	Primer	Primer Sequence (5' - 3')	Cycling Parameters	Product Size	Reference
Bacteroides HF183	HF183F	ATCATGAGTTCACATGTCCG	95 °C for 4 min; 40 cycles of 95 °C for 30 s, 53 °C for 1 min, 72 °C for 2 min; final elongation at 72 °C for 10 min	86 bp	Seurinck et al., 2005
	HF183R	TACCCCGCCTACTATCTAATG			
Bifidobacterium spp.	W257F	GGGTGGTAATGCCGGATG	95 °C for 4 min; 40 cycles of 95 °C for 15 s, 60 °C for 1 min, 72 °C for 1 min; final elongation at 72 °C for 6 min	325 bp	Gourmelon et al., 2010
	W255R	GGTGCTTATTCGAAAGGTACACTCA			
Enterococcus esp	*esp*FWD	TATGAAAGCAACAGCACAAGTT	95 °C for 5 min; 30 cycles of 94 °C for 30 s, 59 °C for 1 min, 72 °C for 1 min; final elongation at 72 °C for 10 min	680 bp	Ahmed et al., 2008a
	*esp*REV	ACGTCGAAAGTTCGATTTCC			
Lachnospiraceae	Lachno2FWD	TTCGCAAGAATGAAACTCAAAG	95 °C for 10 min; 40 cycles of 95 °C for 15 s, 60 °C for 1 min, 72 °C for 1 min; final elongation at 72 °C for 10 min	144 bp	Newton et al., 2011
	Lachno2REV	AAGGAAAGATCCGGTTAAGGATC			

Table 2.1 (Continued) Conventional PCR assay primers, cycling parameters and PCR product size of the MST markers screened for in the tank water and rooftop debris samples.

Microbial Marker	Primer	Primer Sequence (5' - 3')	Cycling Parameters	Product Size	Reference
Methanobrevibacter spp.	MET-105f	TGGGAAACTGGGGATAATACTG	92 °C for 2 min; 30 cycles of 92 °C for 1 min, 55.1 °C for 30 s, 72 °C for 1 min; final elongation at 72 °C for 6 min	282 bp	Ufnar et al., 2006
	MET-386r	AATGAAAAGCCATCCCGTTAAG			
*Methanobrevibacter smithii nif*H	Mnif-342f	AACAGAAAACCCAGTGAAGAG	92 °C for 2 min; 30 cycles of 92 °C for 1 min, 55.1 °C for 30 s, 72 °C for 1 min; final elongation at 72 °C for 6 min	222 bp	Ufnar et al., 2006
	Mnif-363r	ACGTAAAGGCACTGAAAAACC			
Adenovirus	AQ1	GCCACGGTGGGGTTTCTAAACTT	94 °C for 2 min; 35 cycles of 94 °C for 30 s, 55 °C for 1 min, 72 °C for 1 min; final elongation at 72 °C for 7 min	110 bp	Heim et al., 2003
	AQ2	GCCCCAGTGGTCTTACATGCACATC			
Polyomavirus	SM2	AGTCTTTAGGGTCTTCTACCTTT	95 °C for 5 min; 40 cycles of 95 °C for 15 s, 55 °C for 15 s, 60 °C for 1 min; final elongation at 72 °C for 10 min	173 – 176 bp	McQuaig et al., 2009
	P6	GGTGCCAACCTATGGAACAG			

Table 2.1 (Continued) Conventional PCR assay primers, cycling parameters and PCR product size of the MST markers screened for in the tank water and rooftop debris samples.

Microbial Marker	Primer	Primer Sequence (5' - 3')	Cycling Parameters	Product Size	Reference
Enterovirus	EP1	ATTGTCCACCATAAGCAGCCA	Nested PCR: 94 °C for 4 min; 35 cycles: 94 for 1 min, 42 for 40 s, 72 for 2 min; final elongation at 72 °C for 10 min	513 bp	Kuan, 1997; Saayman et al., 2012
	EP2	ACCTTTGTACGCCTGTT			
	EP3	AAGCACTTCTGTTTCCC		297 bp	
	EP4	ATTCAGGGGCCGGAGGA			
Human Mitochondrial DNA	mitoHUf	CAGCAGCCATTCAAGCAATGC	95 °C for 2 min; 40 cycles of 94 °C for 10 s, 60 °C for 12 s, 72 °C for 10 s; final elongation at 72 °C for 10 min	195 bp	Caldwell at al., 2007
	mitoHUr	GGTGGAGACCTAATTGGGCTGATTAG			
Bovine Mitochondrial DNA	mitoBOf	CAGCAGCCCTACAAGCAATGT	95 °C for 2 min; 40 cycles of 94 °C for 10 s, 57 °C for 12 s, 72 °C for 10 s; final elongation at 72 °C for 10 min	191 bp	
	mitoBOr	GAGGCCAAATTGGGCGGATTAT			
Porcine Mitochondrial DNA	mitoPOf	ACAGCTGCACTACAAGCAATGC		196 bp	
	mitoPOr	GGATGTAGTCCGAATTGAGCTGATTAT			

A corresponding positive control and a negative control (sterile milliQ water) were included for each respective PCR assay. For the adenovirus assay, an attenuated adenovirus positive control (lyophilized) was obtained from Coris Bioconcept (Gembloux, Belgium). For the bovine and porcine mtDNA marker assays, positive control DNA was extracted as previously described by Okuma and Hellburg (2014) from beef and pork meat obtained from a local supermarket. Briefly, 30 g of beef or pork meat was placed in a homogenizer bag with 60 mL sterile water. The products were then incubated for 1 h at room temperature and were homogenized in a Seward Stomacher® 400 Circulator (Seward, West Sussex, UK) at 230 rpm for 60 s. The homogenized products were hand-mixed for a further 60 s. Ten milligrams (10 mg) of homogenized meat was utilized for DNA extractions utilizing the QIAamp® DNA Mini Kit (Qiagen, Hilden, Germany) as per manufacturer's instructions (Okuma & Hellburg, 2014). For all the other PCR assays: *Bacteroides* HF183, *Bifidobacterium* spp., *Enterococcus esp,* Lachnospiraceae, *Methanobrevibacter* spp. and *M. smithii nif*H, polyomavirus, enterovirus and human mtDNA, a wastewater sample (1 L) was collected at the influent point from the Stellenbosch Wastewater Treatment Plant (GPS co-ordinates: 33°59'21.13"S 18°47'47.75"E). Briefly, 4 mL (2 × 2 mL) aliquots of the wastewater was centrifuged at 15 000 × *g* for 15 min. The pellets were subsequently resuspended in 200 µL TE buffer and were processed for nucleic acids utilizing the Zymo Research Soil Microbe DNA MiniPrep™ (for whole DNA) and the QIAamp® Ultrasens® Virus Kit (for viral DNA and RNA) as per manufacturer's instructions. Conventional PCR assays were then performed on these nucleic acid extractions for the detection of the specific markers. The PCR products were subjected to sequencing as outlined in Section 2.2.3.4.1. Once the sequencing results confirmed the detection of the correct markers, the nucleic acid extractions were utilized as positive controls for the respective PCR assays. For the enterovirus positive control, RNA was extracted from the wastewater, cDNA was synthesized (Section 2.2.3.3) and the cDNA was subsequently utilized as a positive control in all enterovirus PCR assays.

2.2.3.4.1 Gel Electrophoresis, Sequencing and Data Analysis

The PCR products were electrophoresed on an agarose gel (1.5%) stained with ethidium bromide (0.5 µg/mL) in 1X tris borate ethylenediaminetetraacetic acid (TBE) buffer, for 1 h 20 min at 80 V. The products were then visualized using the Vilber Lourmat gel documentation system (Vilber Lourmat, Collégien, France) to confirm the presence of the desired amplicon.

Once the presence of the desired PCR product was confirmed by gel electrophoresis, representative PCR products were cleaned and concentrated using the Zymo Research DNA Clean & Concentrator 5 kit™ (Zymo Research, Irvine, USA) as per manufacturer's instruction. The purified PCR products and the respective forward primers (**Table 2.1**) were sent to the

Central Analytical Facility (CAF) at Stellenbosch University for sequencing with the BigDye Terminator v3.1 Sequencing Kit (Applied Biosystems, Foster City, USA). Sequencing was also performed for the positive controls to confirm the detection of the correct markers using the conventional PCR assays. All sequence chromatograms were examined utilizing FinchTV version 1.4.0 software. The sequencing data were then analyzed using the online Basic Local Alignment Search Tool (BLAST) program (Altschul et al., 1990) of the National Centre for Biotechnology Information (NCBI) to find the closest match based on local similarity to known sequences on international databases: GenBank, EMBL, DDBJ, and PDB (Altschul et al., 1997). The DNA sequences of representative isolates that showed > 97% similarity (< 3% diversity) to organisms on the database were recorded.

2.2.3.5 Quantitative Real-Time PCR

Quantitative real-time PCR was utilized to quantify the MST markers frequently detected in the tank water and rooftop debris samples. In addition, two indicator organisms (*E. coli* and enterococci) were screened for utilizing qPCR assays. This was done in order to perform correlation analysis with the quantified MST markers, CST markers (as described in Section 2.2.4) and the plate counts obtained from the tank water samples (as described in Section 2.2.2). The qPCR assays were thus utilized to screen the tank water and rooftop debris samples for; *Bacteroides* HF183, adenovirus, Lachnospiraceae, human mtDNA, *E. coli* and *Enterococcus*, with a LightCycler® 96 Instrument (Roche Diagnostics, Mannheim, Germany) (**Table 2.2)**.

For *Bacteroides* HF183, adenovirus, human mtDNA, *E. coli* and *Enterococcus,* each qPCR reaction mixture consisted of: 10 µL (1X) FastStart Essential DNA Green Master (Roche Diagnostics, Mannheim, Germany), 0.4 µL (0.2 µM) of each primer (**Table 2.2**), 4.2 µL PCR-grade water and 5 µL template DNA (DNA extracted from the tank water samples, rooftop debris samples or positive control DNA) in a total volume of 20 µL. For the Lachnospiraceae qPCR assay, the reaction mixture consisted of: 10 µL (1X) FastStart Essential DNA Probes Master (Roche Diagnostics, Mannheim, Germany), 2 µL primer-probe mixture [1 µM of each primer and 0.08 µM of the probe (**Table 2.2**)], 3 µL PCR-grade water and 5 µL template DNA in a total volume of 20 µL. All tank water and rooftop debris DNA samples were diluted (1:9) prior to analysis with the qPCR assay and all DNA samples were analyzed in duplicate.

Table 2.2 Quantitative real-time PCR primers and optimized cycling parameters utilized to screen for: *Bacteroides* HF183, human mtDNA, adenovirus, *E. coli*, *Enterococcus* and Lachnospiraceae.

Marker / Organism (Gene)	Primers	Primer Sequences (5' – 3')	Optimized Cycling Parameters	Reference
Bacteroides HF183 (16S rRNA)	HF183F	ATCATGAGTTCACATGTCCG	95 °C for 10 min; 40 cycles of 95 °C for 30 s, 53 °C for 1 min, 60 °C for 1 min; Melting: 95 °C for 10 s, 65 °C for 1 min, 97 °C for 1 s	Seurinck et al., 2005
	HF183R	TACCCCGCCTACTATCTAATG		
Human mtDNA (NADH dehydrogenase)	mitoHUf	CAGCAGCCATTCAAGCAATGC	95 °C for 2 min; 50 cycles of 94 °C for 10 s, 60 °C for 12 s, 72 °C for 10 s; Melting: 95 °C for 10 s, 65 °C for 1 min, 97 °C for 1 s	Caldwell et al., 2007
	mitoHUr	GGTGGAGACCTAATTGGGCTGATTAG		
Adenovirus (Hexon)	AQ1	GCCACGGTGGGGTTTCTAAACTT	95 °C for 10 min; 55 cycles of 95 °C for 3 s, 55 °C for 10 s, 65 °C for 1 min; Melting: 95 °C for 10 s, 65 °C for 1 min, 97 °C for 1 s	Heim et al., 2003
	AQ2	GCCCCAGTGGTCTTACATGCACATC		

Table 2.2 (Continued) Quantitative real-time PCR primers and optimized cycling parameters utilized to screen for: *Bacteroides* HF183, human mtDNA, adenovirus, *E. coli*, *Enterococcus* and Lachnospiraceae.

Marker / Organism (Gene)	Primers	Primer Sequences (5' – 3')	Optimized Cycling Parameters	Reference
E. coli (*uidA*)	784F	GTGTGATATCTACCCGCTTCGC	95 °C for 10 min; 50 cycles of 95 °C for 15 s, 60 °C for 1 min; Melting: 95 °C for 10 s, 65 °C for 1 min, 97 °C for 1 s	Frahm & Obst, 2003
	866R	AGAACGGTTTGTGGTTAATCAGGA		
Enterococcus (23S rRNA)	ECST784F	AGAAATTCCAAACGAACTTG	95 °C for 10 min; 50 cycles of 95 °C for 15 s, 60 °C for 1 min; Melting: 95 °C for 10 s, 65 °C for 1 min, 97 °C for 1 s	Frahm & Obst, 2003
	ENC854R	CAGTGCTCTACCTCCATCATT		
Lachnospiraceae (16S rRNA V6 region)	Lachno2 FWD	TTCGCAAGAATGAAACTCAAAG	50 °C for 2 min, 95 °C for 10 min; 55 cycles of 95 °C for 15 s, 60 °C for 1 min; Melting: 95 °C for 10 s, 65 °C for 1 min, 97 °C for 1 s	Newton et al., 2011
	Lachno2 REV	AAGGAAAGATCCGGTTAAGGATC		
	Lachno2 probe	6-carboxyfluoroscein (6-FAM)-ACCAAGTCTTGACATCCG – minor groove binder (MGB)		

All the qPCR assays were first optimized by analyzing the amplification of a standard curve. Thereafter, for each qPCR assay, a no template (negative) control (PCR-grade water) was included and a standard curve was included for quantification purposes. The standard curve and no template control were analyzed in duplicate on each 96-well plate utilized for the respective qPCR assays. To generate a standard curve for *Bacteroides* HF183, adenovirus, human mtDNA and Lachnospiraceae, purified conventional PCR products from positive control DNA for each of the markers was utilized (Section 2.2.3.4).

For *E. coli* and enterococci, conventional PCR was performed using the parameters outlined in **Table 2.2** on positive control DNA extracted from a clinical *Enterococcus faecium* isolate (showed sequence similarity to accession number: CP011281.1) and *E. coli* DH5α. The reaction mixture for the detection of both these organisms with the conventional PCR assays, consisted of: 1X Green GoTaq® Reaction Buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTP mix, 0.5 µM of each primer (**Table 2.2**), 1.25 U GoTaq® Flexi DNA polymerase and 5 µL of template DNA in a final volume of 25 µL per reaction. These conventional PCR products were purified and then utilized for the standard curve for the respective qPCR assays.

The concentration of the purified PCR products was determined with a NanoDrop 1000 system (Thermo Fisher Scientific, Carlsbad, USA) in triplicate. The sizes of the respective PCR products were utilized to calculate the dilution needed to obtain a final DNA concentration of 1×10^9 gene copies/µL for the dilution with the highest copy numbers. The DNA was subsequently serially diluted from 1×10^9 to 1×10^0 gene copies/µL, to generate the quantification curve. Lastly, a melting curve analysis was included for all of the SYBR Green qPCR assays in order to verify the specificity of the assay. The temperature was thus increased from 60 to 99 °C at a rate of 0.2 °C/s with continuous fluorescent signal acquisition of 5 readings/°C, to generate the melting curve.

All the qPCR performance characteristics were analyzed using the Roche LightCycler® 96 Software Version 1.1 and Microsoft Excel 2016. In addition, the lower limit of detection (LLOD) was determined as the lowest gene copies/µL consistently and accurately detected per qPCR assay.

2.2.4 Chemical Source Tracking

2.2.4.1 Detection of Chemical Markers in Tank and Municipal Tap Water Samples

For the detection of acetaminophen, salicylic acid, caffeine, methylparaben, triclosan, triclocarban and carbamazepine in all tank water and municipal tap water (collected sporadically throughout the sampling period) samples, 1 L aliquots of the respective water samples were processed as outlined in the US EPA method 1694 (US EPA, 2007). Briefly, 500 mg

ethylenediaminetetraacetic acid tetrasodium salt hydrate ($Na_4EDTA \cdot 2H_2O$) was added to each sample aliquot prior to Solid Phase Extraction (SPE). The aliquots were mixed and allowed to equilibrate for 1 h at room temperature. Hereafter, the chemical markers (acetaminophen, salicylic acid, caffeine, methylparaben, triclosan, triclocarban and carbamazepine) were extracted by SPE with hydrophilic-lipophilic balanced (HLB) cartridges (Supel™ Select HLB SPE 1 g/20 mL; Sigma-Aldrich, St. Louis, USA). As outlined in the US EPA method 1694 (US EPA, 2007), the cartridges were pre-conditioned by eluting 20 mL of methanol (LiChrosolv®, Merck, Johannesburg, SA) followed by the addition of 6 mL reagent water (commercially available bottled water) and then 6 mL reagent water with an adjusted pH of 2. Each 1 L sample aliquot was filtered through the SPE HLB cartridge at a flow rate of approximately 5 - 10 mL/min. A wash step was then conducted with 10 mL reagent water to remove the ethylenediaminetetraacetic acid (EDTA). The analytes were subsequently eluted in 12 mL of methanol. Two hundred microliters (200 µL) of each extract was concentrated under a gentle stream of nitrogen whereafter 100 µL dansyl chloride (2 mg/mL) (Sigma-Aldrich, St. Louis, USA) and 100 µL sodium bicarbonate ($NaHCO_3$; 0.1 M; pH 10.5) were added to each sample. The samples were derivatized in an oven at 60 °C for 12 min. The derivatized aqueous extracts (200 µL) were sent to the CAF for high-performance liquid chromatography tandem mass spectrometry (HPLC/MS/MS) analysis. In addition, a standard curve was generated for analysis, which consisted of 1 mL of methanol spiked with 10 mg of acetaminophen, salicylic acid, caffeine, methylparaben, triclosan, triclocarban and carbamazepine (Sigma-Aldrich, St. Louis, USA), respectively, which was then serially diluted (10^{-4} – 10^{-8}; 1 ppm - 0.0001 ppm) to create a standard curve for quantification purposes. The HPLC/MS/MS analysis was performed with the Waters Xevo TQ MS with UPLC (quadrupole tandem mass spectrometer; Waters, Milford, USA) coupled with Waters BEH C18, 2.1x100 mm, 1.7 µm columns (Waters, Milford, USA).

2.2.4.2 Detection of Chemical Markers in Rooftop Debris Samples

For the detection of acetaminophen, salicylic acid, caffeine, methylparaben, triclosan, triclocarban and carbamazepine, in all the rooftop debris samples, 10 g of each sample was processed as outlined in the US EPA method 1694 (US EPA, 2007). Briefly, 15 mL of phosphate buffer (pH 2) was added to each debris sample. The samples were vortexed and 20 mL acetonitrile (Emsure®, Merck, Kennilworth, USA) was added to each sample. Each sample was then sonicated (Branson 5510 Sonicator) for 30 min followed by centrifugation for 5 min at 1 200 × *g*. The supernatant was subsequently decanted into a separate, sterile 250 mL Schott bottle. This procedure was repeated thrice. The decanted supernatant represented the extract and contained the compounds to be analyzed. The samples were then concentrated by freeze drying and not rotary evaporation as stated in the US EPA method 1694 (US EPA, 2007). The lyophilized samples were reconstituted in 200 mL reagent water and 500 mg

$Na_4EDTA \cdot 2H_2O$ was added to each sample. The extracts were subjected to SPE with HLB cartridges, were concentrated and analyzed as described in Section 2.2.4.1.

2.2.5 Statistical Analysis

To compare the mean gene copies/µL of each MST marker, *E. coli* and enterococci detected in the tank water samples with the mean gene copies/µL detected in the rooftop debris samples, a *t*-test for equal means with a two-tailed *p* value was conducted in Microsoft Excel 2016. The same *t*-test was also utilized to compare the concentrations of the CST markers detected in the rooftop debris with the concentrations of the CST markers detected in the tank water samples by HPLC/MS/MS analysis. In addition, parametric Pearson's correlation (significant when $p < 0.05$) analyses were conducted to determine significant correlations between the rainfall observed during the sampling period and the concentration of the MST and CST markers detected in the tank water and rooftop debris samples. All the correlation analyses were performed using Statistica™ 64 Version 13 (2016) software.

For the municipal tap water samples, *t*-tests for independent samples between groups were conducted with Statistica™ 64 Version 13 (2016) software, to determine significant differences between the concentrations of the CST markers detected in the municipal tap water samples collected during the three different sampling occasions. In addition, *t*-test for equal means were conducted to determine significant differences between the concentrations of the CST markers detected in the tank water samples versus the concentrations of the CST markers detected in the municipal tap water samples. Moreover, as surface water is primarily utilized as a source of municipal tap water for the Kleinmond region, parametric Pearson's correlation (significant when $p < 0.05$) analyses were conducted to determine significant correlations between the rainfall observed during the sampling period and the concentrations of the markers detected in the municipal tap water samples.

Pearson's correlation (significant when $p < 0.05$) analyses were also used to investigate relationships between the indicator organisms enumerated with culturing, the quantified MST markers, *E. coli* and enterococci investigated with the respective qPCR assays and the concentrations of the CST markers as determined by HPLC/MS/MS analysis. The correlation analyses were further investigated using Eigenvalues and a Cattell Scree plot to reduce the number of variables and to identify strongly related variables. The variables then refer to the various source tracking markers and indicator organisms screened for in this study. This analysis was followed by a Factor Loadings (Varimax normalized) analysis with 8 factors and a marked loadings value equal to or greater than 0.55. The Factor Analysis identifies strongly related variables that may share a single factor (not necessarily defined) that may cause the relatedness observed for the variables. Lastly, Cluster Analysis with Ward's method was utilized

to illustrate/visualize the relatedness of the indicator organisms, MST markers and CST markers detected in the tank water and rooftop debris samples. The Cluster Analysis is based on the Pearson's correlations observed in the data set for all the variables analyzed. The stronger the correlation observed between two variables the more closely related two variables will be and the shorter the linkage distance observed between the variables.

2.3 Results

2.3.1 Physico-Chemical Parameters

Overall 60 tank water samples were collected from ten DRWH tanks located in the Kleinmond Housing Scheme site (Western Cape, South Africa), while 60 rooftop debris samples were collected during the sampling period from the gutters connected to each sampled rainwater storage tank. Samples were collected on six occasions after a rain event, during high and low rainfall periods. The total rainfall recorded per month was 40.4 mm in August 2015 (sampling one), 50.2 mm in September 2015 (sampling two), 13.4 mm in October 2015 (sampling three), 39.4 mm in March 2016 (sampling four), 26.6 mm in April 2016 (sampling five) and 88.6 mm in June 2016 (sampling six) (South African Weather Services, Pretoria, South Africa). The average minimum and maximum ambient temperatures were also recorded per month and ranged from 11.4 °C to 17.6 °C in August 2015, 12.1 °C to 20.2 °C in September 2015, 14.2 °C to 21.7 °C in October 2015, 15.7 °C to 22.6 °C in March 2016, 14.7 °C to 21.9 °C in April 2016 and 10.8 °C to 18.7 °C in June 2016.

In addition, the average temperature of the tank water was recorded as 16.8 °C for the entire sampling period. The tank water temperatures then ranged from 15.3 °C to 18.0 °C in August 2015, 13.2 °C to 14.5 °C in September 2015, 21.9 °C to 23.8 °C in October 2015, 16.7 °C to 18.7 °C in March 2016, 16.7 °C to 19.0 °C in April 2016 and 11.6 °C to 12.3 °C in June 2016. The average pH of the tank water was recorded as 7.5. The pH of the tank water ranged from 8.2 to 8.3 in August 2015, 8.2 to 8.4 in September 2015, 8.4 to 8.5 in October 2015, 6.9 to 7.2 in March 2016, 6.9 to 7.6 in April 2016 and 5.4 to 7.5 in June 2016.

2.3.2 Enumeration of Indicator Organisms in Tank Water Samples

The traditional indicator organisms *E. coli*, enterococci, total coliforms, fecal coliforms and heterotrophic bacteria, were enumerated for all 60 tank water samples collected throughout the sampling period (August 2015 to June 2016). These parameters were compared to national and international drinking water guidelines as there are currently no guidelines stipulated for harvested rainwater. Throughout the sampling period, the *E. coli* counts recorded in 57 (95%) of the tank water samples analyzed exceeded the recommended drinking water guideline of < 1 colony forming units (CFU)/100 mL as outlined by the South African National Standard

(SANS) 241 [South African Bureau of Standards (SABS, 2005)], the World Health Organization (WHO) (2011), the Department of Water Affairs and Forestry (DWAF) (1996) and the Australian Drinking Water Guidelines (ADWG) [National Health and Medical Research Council (NHMRC) and Natural Resource Management Ministerial Council (NRMMC), 2011]. The results obtained indicated that the *E. coli* counts ranged from 3.6×10^1 CFU/100 mL in the fifth sampling occasion to 6.1×10^2 CFU/100 mL in the first sampling occasion. Throughout the sampling period, the lowest *E. coli* count of < 1 CFU/100 mL [below the detection limit (BDL)] was recorded for three tank water samples (TH1, TB4 and TH5), while the highest *E. coli* count of 5.9×10^3 CFU/100 mL was recorded for TE1, with an overall mean *E. coli* count of 2.8×10^2 CFU/100 mL obtained for the entire sampling period (**Fig. 2.2.**).

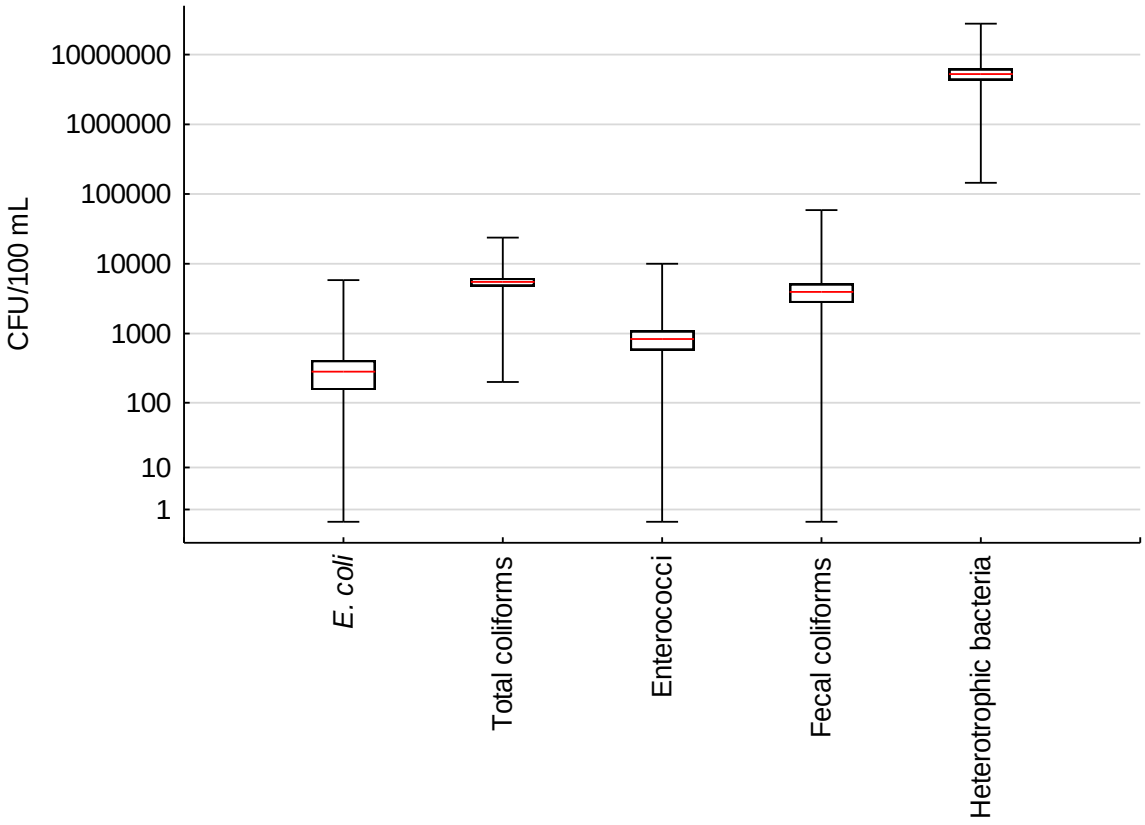


Fig. 2.2. Box and whiskers plot of the indicator organisms detected in the tank water samples from sampling one to six with the whiskers illustrating the minimum and maximum CFU/100 mL, the outer box illustrating the mean CFU/100 mL ± the standard error and the red inner line illustrating the mean CFU/100 mL.

The total coliform counts for all the tank water samples (n = 60; 100%) collected throughout the sampling period exceeded the recommended drinking water guideline of 5 CFU/100 mL and 10 CFU/100 mL as stipulated by DWAF (1996) and SANS 214 (SABS, 2005), respectively. The results obtained for the total coliform counts then ranged from 3.1×10^3 CFU/100 mL in the first sampling occasion to 9.5×10^3 CFU/100 mL in the sixth sampling occasion. For the entire sampling period, the lowest total coliform count of 2.0×10^2 CFU/100 mL was recorded in TH5,

while the highest total coliform count of 2.4×10^4 CFU/100 mL was recorded for TB2 (**Fig. 2.2.**). An overall mean total coliform count of 5.5×10^3 CFU/100 mL was obtained for the entire sampling period.

Enterococci counts sporadically exceeded the drinking water guideline of < 1 CFU/100 mL as stipulated by ADWG (NHMRC and NRMMC, 2011) with enterococci detected in 32% (n = 19) of the tank water samples collected during the sampling period. The results obtained indicated that the enterococci counts ranged from 2.5×10^2 CFU/100 mL in the first sampling occasion to 1.8×10^3 CFU/100 mL in the fifth sampling occasion. For the entire sampling period, the lowest enterococci count of < 1 CFU/100 mL (BDL) was recorded in various tanks, while the highest count of 1.0×10^4 CFU/100 mL was recorded for TB4 and TB5, with an overall mean enterococci count of 8.3×10^2 CFU/100 mL obtained for the entire sampling period (**Fig. 2.2.**).

In addition, the fecal coliform counts exceeded the recommended drinking water guideline of < 1 CFU/100 mL as stipulated by SANS 241 (SABS, 2005), WHO (2011), DWAF (1996) and ADWG (NHMRC and NRMMC, 2011), in 85% (n = 51) of the tank water samples analyzed. The results then indicated that the fecal coliform counts ranged from 1.9×10^3 CFU/100 mL in the first sampling occasion to 9.6×10^3 CFU/100 mL in the sixth sampling occasion. For the entire sampling period, the lowest fecal coliform count of < 1 CFU/100 mL (BDL) was recorded for various tanks, while the highest count of 5.9×10^4 CFU/100 mL was recorded for TD6, with an overall mean fecal coliform count of 4.0×10^3 CFU/100 mL recorded for the entire sampling period (**Fig. 2.2.**).

The heterotrophic bacteria counts obtained then exceeded the drinking water guideline of 100 CFU/mL (1×10^4 CFU/100 mL) and 1000 CFU/mL (1×10^5 CFU/100 mL) as stipulated by DWAF (1996) and SANS 241 (SABS, 2005), respectively, in all the tank water samples (n = 60; 100%) analyzed during the sampling period. The results for the heterotrophic plate counts indicated that the counts ranged from 6.5×10^5 CFU/100 mL in the fourth sampling occasion to 1.7×10^7 CFU/100 mL in the third sampling occasion. For the entire sampling period, the lowest heterotrophic plate count of 1.5×10^5 CFU/100 mL was recorded for TA4, while the highest count of 2.8×10^7 CFU/100 mL was recorded for TD3, with an overall mean heterotrophic plate count of 5.3×10^6 CFU/100 mL obtained for the entire sampling period (**Fig. 2.2.**).

2.3.3 Detection of MST Markers in Tank Water and Rooftop Debris Samples with Conventional PCR

Conventional PCR assays were performed on the nucleic acids extracted from all the tank water and rooftop debris samples collected during the entire sampling period to determine the most prevalent MST markers. For the tank water samples, overall the *Bacteroides* HF183 (Genbank accession number CP011531.1) MST marker exhibited the highest frequency of detection

(86.7%; n = 52), followed by adenovirus (Genbank accession number K01264.1), which was detected in 66.7% (n = 40) of the tank water samples (**Fig. 2.3.**). Lachnospiraceae (Genbank accession number KF374935.1) was then detected in 55% (n = 33) of the tank water samples, while bovine mtDNA (Genbank accession number KT343749.1) was detected in 37% (n = 22) and human mtDNA (Genbank accession number KX055572.1) was detected in 35% (n = 21) of the tank water samples, respectively (**Fig. 2.3.**). Furthermore, *Enterococcus esp* (Genbank accession number AH013270.2) was detected in 33% (n = 20), porcine mtDNA (Genbank accession number KT279760.1) was detected in 32% (n = 19), and *Bifidobacterium* (Genbank accession number AB470329.1) and enterovirus (Genbank accession number HM209152.1) were both detected in 28% (n = 17) of the tank water samples, respectively (**Fig. 2.3.**). In contrast polyomavirus, *Methanobrevibacter* spp. and *M. smithii nif*H were not detected in any of the tank water samples collected throughout the sampling period (results not shown).

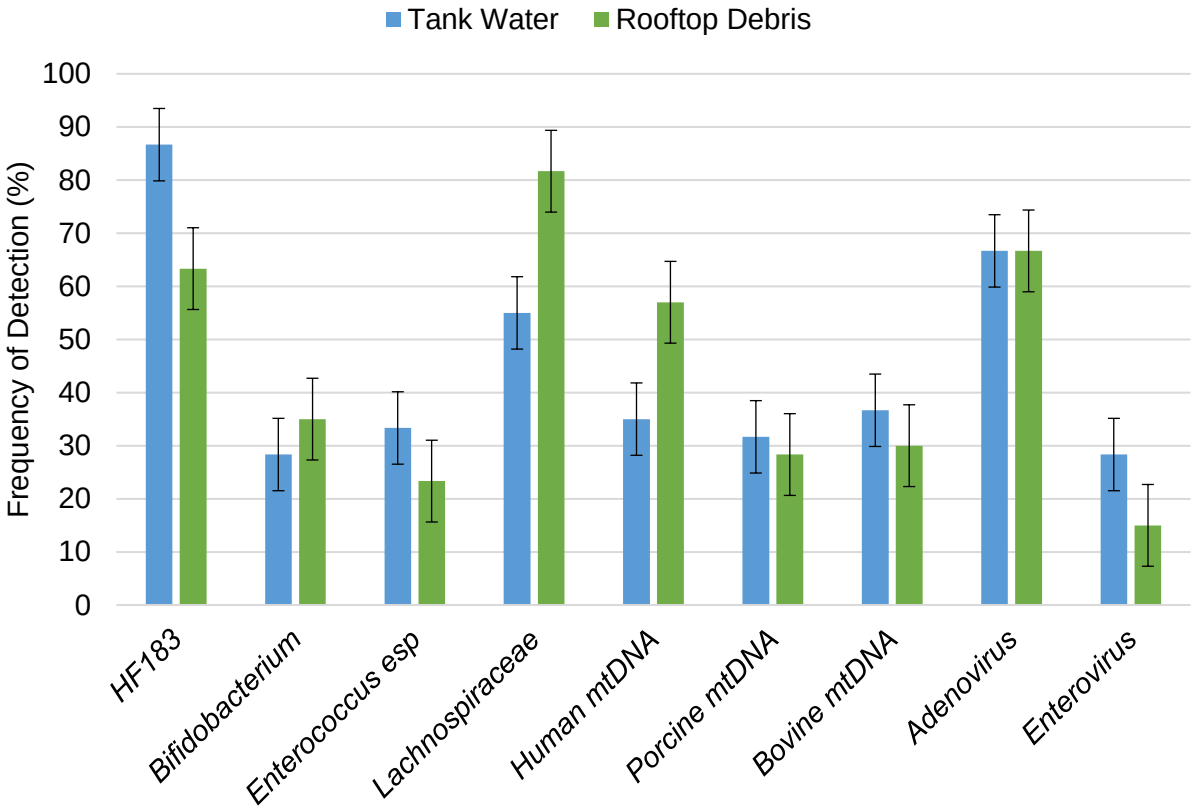


Fig. 2.3. The frequency of detection percentages of the MST markers in the tank water and rooftop debris samples for sampling one to six with standard error indicated by the error bars.

For the rooftop debris samples, overall Lachnospiraceae exhibited the highest frequency of detection (81.7%; n = 49), followed by adenovirus which was detected in 66.7% (n = 40) of the rooftop debris samples. *Bacteroides* HF183 was then detected in 63.3% (n = 38) of the rooftop debris samples, while human mtDNA was detected in 57% (n = 34) of the rooftop debris samples (**Fig. 2.3.**). Furthermore, *Bifidobacterium* was detected in 35% (n = 21), bovine mtDNA was detected in 30% (n = 18), porcine mtDNA was detected in 28% (n = 17), *Enterococcus esp*

was detected in 23% (n = 14) and enterovirus was detected in 15% (n = 9) of the rooftop debris samples, respectively (**Fig. 2.3.**). Similar to the results obtained for the tank water samples, polyomavirus, *Methanobrevibacter* spp. and *M. smithii nif*H were not detected in any rooftop debris samples collected during the sampling period (results not shown).

2.3.4 Quantification of the MST Markers Frequently Detected in the Tank Water and Rooftop Debris Samples

The MST markers frequently detected in the tank water and rooftop debris samples included *Bacteroides* HF183, Lachnospiraceae, adenovirus and human mtDNA (**Fig. 2.3.**). These markers were subsequently quantified using qPCR analysis. In addition, qPCR was utilized to quantify *E. coli* and enterococci in the tank water and rooftop debris samples in order to perform correlation analysis with the indicator organisms enumerated by culture based analysis and the quantified MST and CST markers. All the qPCR performance characteristics are summarized in the **Appendix Table A1**.

2.3.4.1 *Bacteroides* HF183

The qPCR assays for the detection of the HF183 marker (16S rRNA) had a mean amplification efficiency (*E*) of 99.4% and a correlation coefficient (r^2) of 0.98 (**Appendix Table A1**). The LLOD was determined to range from 1.44 to 1.33 × 10^1 gene copies/µL. The HF183 marker was detected in all tank water samples (100%; n = 60) at a mean concentration of 5.1 × 10^3 gene copies/µL for the entire sampling period (**Fig. 2.4A**). The concentration of the HF183 marker ranged from 1.8 × 10^2 gene copies/µL obtained during the third sampling occasion to 1.9 × 10^4 gene copies/µL recorded during the fourth sampling occasion. The lowest HF183 copy number of 9.3 × 10^1 gene copies/µL was detected in TD3, while the highest copy number of 3.3 × 10^4 gene copies/µL was detected in TC4 (**Fig. 2.4A**).

Similarly, all the rooftop debris samples (100%; n = 60) tested positive for the HF183 marker at a mean concentration of 4.7 × 10^3 gene copies/µL for the entire sampling period (**Fig. 2.4B**). The concentration of the HF183 marker in the rooftop debris samples then ranged from 6.6 × 10^2 gene copies/µL obtained during the first sampling session to 1.1 × 10^4 gene copies/µL recorded during the fourth sampling session. The lowest HF183 copy number of 2.3 × 10^2 gene copies/µL was detected in rooftop debris collected from the gutter system of tank B during sampling occasion one (RD-B1), while the highest HF183 copy number of 2.3 × 10^4 gene copies/µL was detected in RD-F4 (**Fig. 2.4B**).

The *t*-test analysis indicated that the mean concentration (5.1 × 10^3 gene copies/µL) of the HF183 marker detected in the tank water samples did not differ significantly from the mean

concentration (4.7×10^3 gene copies/µL) of the HF183 marker detected in the rooftop debris samples ($p = 0.61$) (**Fig. 2.4A** and **Fig. 2.4B**).

2.3.4.2 Adenovirus

The qPCR assays for the detection of the adenovirus (Hexon gene) had a mean *E* of 94.8% and a r^2 of 0.99 (**Appendix Table A1**). The LLOD was determined to range from 2.33 to 4.28 gene copies/µL. Adenovirus was detected in all tank water samples (100%; n = 60) at a mean concentration of 3.2×10^2 gene copies/µL for the entire sampling period (**Fig. 2.4A**). The concentration of the adenovirus then ranged from 3.0×10^1 gene copies/µL obtained during the fourth sampling occasion to 1.4×10^3 gene copies/µL recorded during the first sampling occasion. The lowest adenovirus copy number of 7.3 gene copies/µL was detected in TE6, while the highest copy number of 1.2×10^4 gene copies/µL was detected in TE1 (**Fig. 2.4A**).

All the rooftop debris samples (100%; n = 60) also tested positive for adenovirus at a mean concentration of 6.4×10^3 gene copies/µL for the entire sampling period (**Fig. 2.4B**). The concentration of adenovirus in the rooftop debris samples ranged from 2.9×10^1 gene copies/µL obtained during the third sampling occasion to 2.8×10^4 gene copies/µL recorded during the fifth sampling occasion. The lowest adenovirus copy number of 3.4 gene copies/µL was detected in RD-A4, while the highest adenovirus copy number of 4.2×10^4 gene copies/µL was detected in RD-G5 (**Fig. 2.4B**).

The *t*-test analysis indicated that the mean concentration (3.2×10^2 gene copies/µL) of adenovirus detected in the tank water samples differed significantly from the mean concentration (6.4×10^3 gene copies/µL) of adenovirus detected in the rooftop debris samples ($p = 0.00025$), as the mean adenovirus gene copies/µL detected in the rooftop debris samples were significantly higher than the mean gene copies/µL detected in the tank water samples (**Fig. 2.4A** and **Fig. 2.4B**).

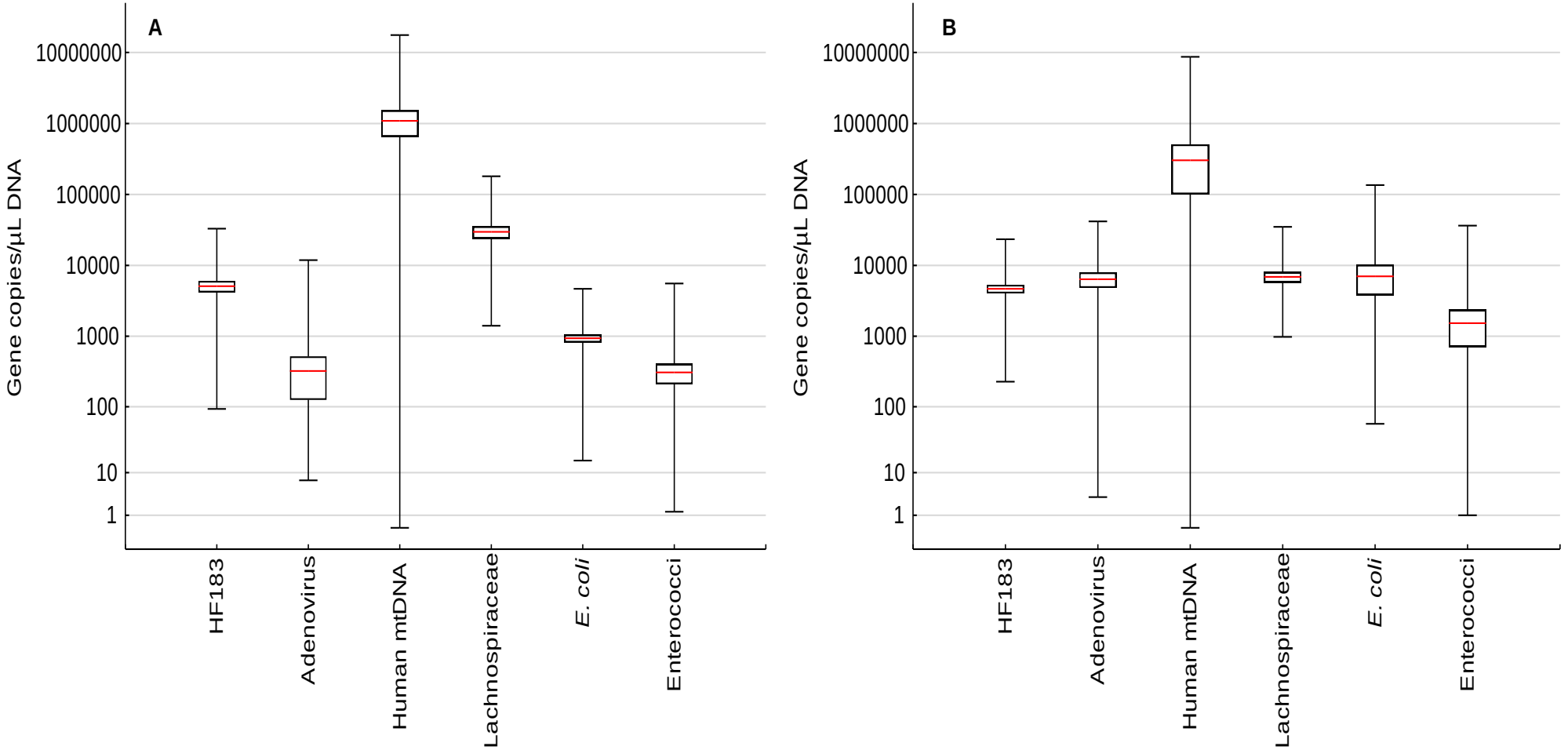


Fig. 2.4. Box and whiskers plot of the gene copies/µL for HF183, adenovirus, human mtDNA, Lachnospiraceae, *E. coli* and enterococci detected in the tank water (**A**) and rooftop debris (**B**) samples for sampling one to six, with the whiskers illustrating the minimum and maximum gene copies/µL, the outer box illustrating the mean gene copies/µL ± standard error and the inner red line illustrating the mean gene copies/µL.

2.3.4.3 Human Mitochondrial DNA

The qPCR assays for the detection of human mtDNA (NADH dehydrogenase subunit 5) had a mean *E* of 92.8% and a r^2 of 0.98 (**Appendix Table A1**). The LLOD was determined to range from 2.41 to 1.10 × 10^1 gene copies/µL. Human mtDNA was detected in all tank water samples collected during sampling occasions three to six (n = 40) and 70% (n = 14) of the tank water samples collected during sampling sessions one and two. Human mtDNA was detected at a mean concentration of 1.1 × 10^6 gene copies/µL for the entire sampling period (**Fig. 2.4A**). The concentrations then ranged from 6.6 × 10^1 gene copies/µL obtained during the second sampling occasion to 4.4 × 10^6 gene copies/µL recorded during the fifth sampling occasion. The lowest human mtDNA copy number of BDL (< 2.41 - 1.10 × 10^1 gene copies/µL) was detected in various tank water samples, while the highest human mtDNA copy number of 1.8 × 10^7 gene copies/µL was detected in TI5 (**Fig. 2.4A**).

Human mtDNA was detected in all the rooftop debris samples collected during sampling two, five and six (n = 30), in 90% (n = 9) of the rooftop debris samples collected during sampling one and in 80% (n = 16) of the rooftop debris samples collected during sampling three and four. Human mtDNA was detected at a mean concentration of 3.0 × 10^5 gene copies/µL for the entire sampling period (**Fig. 2.4B**). The concentrations in the rooftop debris samples then ranged from 5.8 × 10^1 gene copies/µL obtained during the third sampling occasion to 8.7 × 10^5 gene copies/µL recorded during the first sampling occasion. The lowest human mtDNA copy number of BDL (< 2.41 - 1.10 × 10^1 gene copies/µL) was detected in various rooftop debris samples, while the highest human mtDNA copy number of 8.7 × 10^6 gene copies/µL was detected in RD-C1 (**Fig. 2.4B**).

The *t*-test analysis indicated that the mean concentration (1.1 × 10^6 gene copies/µL) of the human mtDNA detected in the tank water samples did not differ significantly from the mean concentration (3.0 × 10^5 gene copies/µL) of human mtDNA detected in the rooftop debris samples (p = 0.11) (**Fig. 2.4A** and **Fig. 2.4B**).

2.3.4.4 Lachnospiraceae

The qPCR assays for the detection of Lachnospiraceae (16S rRNA V6 region) had a mean *E* of 93.3% and a r^2 of 0.99 (**Appendix Table A1**). The LLOD was determined to range from 7.66 to 6.4 × 10^1 gene copies/µL. All the tank water samples (100%; n = 60) tested positive for Lachnospiraceae at a mean concentration of 3.0 × 10^4 gene copies/µL for the entire sampling period (**Fig. 2.4A**). The Lachnospiraceae concentration ranged from 4.1 × 10^3 gene copies/µL obtained during the fifth sampling occasion to 1.0 × 10^5 gene copies/µL recorded during the third sampling occasion. The lowest Lachnospiraceae concentration of 1.4 × 10^3 gene copies/µL

was detected in TE5, while the highest Lachnospiraceae concentration of 1.8×10^5 gene copies/µL was detected in TD4 (**Fig. 2.4A**).

All the rooftop debris samples (100%; n = 60) also tested positive for Lachnospiraceae at a mean concentration of 6.9×10^3 gene copies/µL for the entire sampling period (**Fig. 2.4B**). The Lachnospiraceae concentration then ranged from 1.3×10^3 gene copies/µL obtained during sampling occasion one to 1.7×10^4 gene copies/µL recorded during sampling occasion six. The lowest Lachnospiraceae copy number of 9.8×10^2 gene copies/µL was detected in RD-E1, while the highest Lachnospiraceae copy number of 3.5×10^4 gene copies/µL was detected in RD-J6 (**Fig. 2.4B**).

The *t*-test analysis indicated that the mean concentration (3.0×10^4 gene copies/µL) of Lachnospiraceae detected in the tank water samples differed significantly from the mean concentration (6.9×10^3 gene copies/µL) of Lachnospiraceae detected in the rooftop debris samples ($p = 0.00061$), with the mean gene copies/µL detected in the tank water samples higher than the mean gene copies/µL concentration detected in the rooftop debris samples (**Fig. 2.4A** and **Fig. 2.4B**).

2.3.4.5 *Escherichia coli*

The qPCR assays for the detection of *E. coli* (*uidA*) had a mean *E* of 96.9% and a r^2 of 0.95 (**Appendix Table A1**). The LLOD was determined to range from 3.67 to 3.6×10^1 gene copies/µL. *Escherichia coli* was detected in all tank water samples (100%; n = 60) at a mean concentration of 9.4×10^2 gene copies/µL for the entire sampling period (**Fig. 2.4A**). The concentration of *E. coli* ranged from 3.5×10^2 gene copies/µL obtained during the second sampling occasion to 1.7×10^3 gene copies/µL recorded during the fifth sampling occasion. The lowest *E. coli* copy number of 1.6×10^1 gene copies/µL was detected in TE2, while the highest *E. coli* copy number of 4.7×10^3 gene copies/µL was detected in TB5 (**Fig. 2.4A**).

Escherichia coli was also detected in all the rooftop debris samples (100%; n = 60) collected during the sampling period at a mean concentration of 7.0×10^3 gene copies/µL (**Fig. 2.4B**). The *E. coli* concentration in the rooftop debris samples then ranged from 1.5×10^2 gene copies/µL, which was recorded during sampling occasion five to 3.2×10^4 gene copies/µL obtained during sampling occasion two. The lowest *E. coli* copy number of 5.6×10^1 gene copies/µL was detected in RD-E2, while the highest *E. coli* copy number of 1.4×10^5 gene copies/µL was detected in RD-C2 (**Fig. 2.4B**).

The *t*-test analysis indicated that the mean concentration of *E. coli* (9.4×10^2 gene copies/µL) detected in the tank water samples did not differ significantly from the mean concentration of

E. coli (7.0×10^3 gene copies/µL) detected in the rooftop debris samples ($p = 0.06$) (**Fig. 2.4A** and **Fig. 2.4B**).

2.3.4.6 Enterococci

The qPCR assays for the detection of enterococci (23S rRNA) had a mean *E* of 98.3% and a r^2 of 0.98 (**Appendix Table A1**). The LLOD was determined to be equal to 1.51 to 2.8×10^1 gene copies/µL. Enterococci was detected in all the tank water samples (100%; n = 60) collected during the sampling period at a mean concentration of 3.1×10^2 gene copies/µL (**Fig. 2.4A**). The concentration of the enterococci ranged from 1.6×10^1 gene copies/µL obtained during sampling five to 7.7×10^2 gene copies/µL recorded during sampling two. The lowest enterococci copy number of 2.01 gene copies/µL was detected in TC5, while the highest enterococci copy number of 5.6×10^3 gene copies/µL was detected in TB2 (**Fig. 2.4A**).

Enterococci was also detected in all the rooftop debris samples (100%; n = 60) collected during the sampling period at a mean concentration of 1.5×10^3 gene copies/µL (**Fig. 2.4B**). The enterococci copy numbers in the rooftop debris samples then ranged from 4.6×10^1 gene copies/µL recorded for sampling occasion five to 3.8×10^3 gene copies/µL recorded for the fourth sampling occasion. The lowest enterococci copy number of 2.03 gene copies/µL was detected in RD-G3, while the highest enterococci copy number of 3.6×10^4 gene copies/µL was detected in RD-D4 (**Fig. 2.4B**).

The *t*-test analysis indicated that the mean concentration of enterococci (3.1×10^2 gene copies/µL) in the tank water samples did not differ significantly from the enterococci (1.5×10^3 gene copies/µL) detected in the rooftop debris samples ($p = 0.14$) (**Fig. 2.4A** and **Fig. 2.4B**).

2.3.5 Detection of the CST Markers in the Tank Water and Rooftop Debris Samples

All tank water and rooftop debris samples were screened for caffeine, salicylic acid, methylparaben, triclosan, triclocarban, acetaminophen and carbamazepine as indicators of anthropogenic activity impacting on the DRWH systems. Carbamazepine was however, BDL (< 0.0001 ppm or < 0.1 µg/L) for all tank water and rooftop debris samples collected during the entire sampling period (100%; n = 60). The concentrations of the CST markers detected per tank water sample are summarized in **Appendix Table A2** and the concentrations of the CST markers detected per rooftop debris samples are summarized in **Appendix Table A3**.

Caffeine was detected in all the tank water samples (100%; n = 60) collected during the sampling period at an overall mean concentration of 2.50 µg/L (**Table 2.3**). The lowest caffeine concentration of 0.9 µg/L was detected in TC3, while the highest caffeine concentration of

8.0 µg/L was detected in TB5 (**Appendix Table A2**). Caffeine was then detected at a mean concentration of 1.78 µg/L during sampling one, 2.06 µg/L during sampling two and 1.62 µg/L during sampling three. Caffeine was also detected at a mean concentration of 2.70 µg/L during the fourth sampling occasion, 3.61 µg/L during the fifth sampling occasion and 3.22 µg/L during the sixth sampling occasion.

Table 2.3 The average concentration of each CST marker for the six sampling occasions in µg/L, detected in the tank water and rooftop debris samples obtained from the respective rainwater storage tanks or gutter systems connected to each sampled rainwater storage tank.

Tank Water (µg/L)						
Tank	**Caffeine**	**Salicylic acid**	**Methylparaben**	**Triclosan**	**Triclocarban**	**Acetaminophen**
A	2.67	13.4	3.63	10.2	1.70	8.93
B	4.53	14.5	2.50	6.27	0.933	9.42
C	1.92	12.2	2.73	5.57	0.983	9.43
D	2.32	13.3	4.19	4.80	0.933	8.82
E	3.02	12.9	2.10	4.33	0.883	8.93
F	2.38	14.2	2.48	3.43	0.867	8.88
G	2.07	10.5	1.95	5.70	0.900	8.40
H	2.13	12.3	4.90	4.25	0.850	8.32
I	2.07	13.4	2.23	2.52	0.883	8.65
J	1.88	11.1	2.87	1.67	0.833	7.42
Mean	2.50	12.8	2.96	4.87	0.977	8.72
Rooftop Debris (µg/L)						
Tank	**Caffeine**	**Salicylic acid**	**Methylparaben**	**Triclosan**	**Triclocarban**	**Acetaminophen**
A	1.80	11.4	2.83	2.98	1.18	8.27
B	1.57	10.4	3.13	2.73	1.27	9.05
C	1.83	10.8	4.77	2.75	1.30	8.60
D	1.73	10.9	4.30	1.93	1.33	9.08
E	2.22	10.4	2.73	1.63	1.67	8.48
F	2.33	10.4	4.32	3.67	1.50	8.62
G	1.85	10.1	3.62	1.47	1.17	8.57
H	1.62	10.1	3.27	2.00	1.13	8.13
I	1.57	11.4	2.73	1.95	1.08	7.80
J	1.97	12.6	4.10	2.03	1.67	9.07
Mean	1.85	10.8	3.58	2.32	1.33	8.57

Similarly, caffeine was detected in all the rooftop debris samples (100%; n = 60) collected during the sampling period, at a mean concentration of 1.85 µg/L (**Table 2.3**). The lowest caffeine concentration of 1.0 µg/L was detected in RD-H2, while the highest caffeine concentration of 3.4 µg/L was detected in RD-J4 (**Appendix Table A3**). Caffeine was then detected at a mean concentration of 1.75 µg/L during sampling one and three and 1.99 µg/L during sampling two. Caffeine was also detected at a mean concentration of 1.78 µg/L during

the fourth sampling occasion, 1.80 µg/L during the fifth sampling occasion and 2.02 µg/L during the sixth sampling occasion.

Salicylic acid was detected in 98% of the tank water samples analyzed (n = 59), at a mean concentration of 12.8 µg/L (**Table 2.3**). The lowest salicylic acid concentration of BDL (< 0.0001 ppm; < 0.1 µg/L) was detected in TG5, while the highest salicylic acid concentration of 22.6 µg/L was detected in TF2 (**Appendix Table A2**). Salicylic acid was detected at a mean concentration of 14.1 µg/L during sampling one, 16.4 µg/L during sampling two and 12.2 µg/L during sampling three. In addition, salicylic acid was detected at a mean concentration of 10.9 µg/L during the fourth sampling occasion, 12.4 µg/L during the fifth sampling occasion and 10.7 µg/L during the sixth sampling occasion.

Salicylic acid was detected in all the rooftop debris samples (100%; n = 60) collected during the sampling period, at a mean concentration of 10.8 µg/L (**Table 2.3**). The lowest salicylic acid concentration of 6.0 µg/L was detected in RD-E5, while the highest salicylic acid concentration of 16.4 µg/L was detected in RD-I2 (**Appendix Table A3**). Salicylic acid was detected at a mean concentration of 11.4 µg/L during sampling one, 12.6 µg/L during sampling two and 10.5 µg/L during sampling three. In addition, salicylic acid was detected at a mean concentration of 10.3 µg/L during the fourth sampling occasion, 9.76 µg/L during the fifth sampling occasion and 10.4 µg/L during the sixth sampling occasion.

Methylparaben was detected in all of the tank water samples (100%; n = 60) collected during the sampling period, at a mean concentration of 2.96 µg/L (**Table 2.3**). The lowest methylparaben concentration of 0.2 µg/L was detected in TB5 and TE6, while the highest methylparaben concentration of 17.4 µg/L was detected in TH2 (**Appendix Table A2**). Methylparaben was detected at a mean concentration of 3.60 µg/L during sampling one, 5.64 µg/L during sampling two and 2.30 µg/L during sampling three. In addition, methylparaben was detected at a mean concentration of 2.11 µg/L during the fourth sampling occasion, 2.00 µg/L during the fifth sampling occasion and 2.10 µg/L during the sixth sampling occasion.

Methylparaben was also detected in all the rooftop debris samples (100%; n = 60) collected during the sampling period, at a mean concentration of 3.58 µg/L (**Table 2.3**). The lowest methylparaben concentration of 0.4 µg/L was detected in RD-E5, while the highest methylparaben concentration of 19.5 µg/L was detected in RD-C2 (**Appendix Table A3**). Methylparaben was detected at a mean concentration of 2.03 µg/L during sampling one, 12.9 µg/L during sampling two and 1.61 µg/L during sampling three. In addition, methylparaben was detected at a mean concentration of 1.64 µg/L during the fourth sampling occasion, 1.54 µg/L during the fifth sampling occasion and 1.81 µg/L during the sixth sampling occasion.

Triclosan was detected in all of the tank water samples (100%; n = 60) collected during the sampling period, at a mean concentration of 4.87 µg/L (**Table 2.3**). The lowest triclosan concentration of 0.4 µg/L was detected in TJ6, while the highest triclosan concentration of 26.1 µg/L was detected in TA5 (**Appendix Table A2**). Triclosan was detected at a mean concentration of 0.76 µg/L during sampling one, 1.09 µg/L during sampling two and 1.06 µL during sampling three. Triclosan was then also detected at a mean concentration of 7.35 µg/L during the fourth sampling occasion, 9.97 µg/L during the fifth sampling occasion and 9.01 µg/L during the sixth sampling occasion.

Triclosan was also detected in all the rooftop debris samples (100%; n = 60) collected during the sampling period, at a mean concentration of 2.32 µg/L (**Table 2.3**). The lowest triclosan concentration of 0.3 µg/L was detected in RD-E5, while the highest triclosan concentration of 15.0 µg/L was detected in RD-F2 (**Appendix Table A3**). Triclosan was detected at a mean concentration of 1.19 µg/L during sampling one, 3.46 µg/L during sampling two and 1.06 µg/L during sampling three. Triclosan was then also detected at a mean concentration of 1.05 µg/L during the fourth sampling occasion, 3.31 µg/L during the fifth sampling occasion and 3.82 µg/L during the sixth sampling occasion.

Triclocarban was detected in all of the tank water samples (100%; n = 60) collected during the sampling period, at a mean concentration of 0.977 µg/L (**Table 2.3**). The lowest triclocarban concentration of 0.3 µg/L was detected in TJ4, while the highest triclocarban concentration of 4.9 µg/L was detected in TA5 (**Appendix Table A2**). Triclocarban was detected at a mean concentration of 0.95 µg/L during sampling one, 1.14 µg/L during sampling two and 1.26 µg/L during sampling three. Triclocarban was then also detected at a mean concentration of 0.66 µg/L during the fourth sampling occasion, 1.18 µg/L during the fifth sampling occasion and 0.67 µg/L during the sixth sampling occasion.

In addition, triclocarban was detected in all of the rooftop debris samples (100%; n = 60) collected during the sampling period, at a mean concentration of 1.33 µg/L (**Table 2.3**). The lowest triclocarban concentration of 0.8 µg/L was detected in RD-I2, RD-F3, RD-F4 and RD-G5, while the highest triclocarban concentration of 3.3 µg/L was detected in RD-E6 (**Appendix Table A3**). Triclocarban was detected at a mean concentration of 1.25 µg/L during sampling one, 1.22 µg/L during sampling two and 1.18 µg/L during sampling occasion three and four, respectively. Triclocarban was also detected at a mean concentration of 1.43 µg/L during the fifth sampling occasion and 1.72 µg/L during the sixth sampling occasion.

Acetaminophen was detected in all tank water samples (100%; n = 60) collected during the sampling period, at a mean concentration of 8.72 µg/L (**Table 2.3**). The lowest acetaminophen concentration of 1.00 µg/L was detected in TJ6, while the highest acetaminophen concentration

of 10.6 µg/L was detected in TC5 (**Appendix Table A2**). Acetaminophen was detected at a mean concentration of 9.12 µg/L during sampling one, 9.14 µg/L during sampling two and 8.94 µg/L during sampling three. Acetaminophen was then also detected at a mean concentration of 7.49 µg/L during the fourth sampling occasion, 9.59 µg/L during the fifth sampling occasion and 8.04 µg/L during the sixth sampling occasion.

Similarly, acetaminophen was detected in all the rooftop debris samples (100%; n = 60) collected during the sampling period, at a mean concentration of 8.57 µg/L (**Table 2.3**). The lowest acetaminophen concentration of 5.2 µg/L was detected in RD-E5, while the highest acetaminophen concentration of 10.9 µg/L was detected in RD-J4 (**Appendix Table A3**). Acetaminophen was detected at a mean concentration of 8.76 µg/L during sampling one, 8.73 µg/L during sampling two and 9.24 µg/L during sampling three. Acetaminophen was then also detected at a mean concentration of 9.02 µg/L during the fourth sampling occasion, 7.96 µg/L during the fifth sampling occasion and 7.69 µg/L during the sixth sampling occasion.

The *t*-test analysis indicated that for methylparaben ($p = 0.57$) and acetaminophen ($p = 0.60$) there was no significant difference between the mean concentrations of these markers detected in the tank water and rooftop debris samples. In contrast, significant differences were observed for the concentrations of caffeine ($p = 0.015$), salicylic acid ($p = 0.00025$), triclosan ($p = 0.005$) and triclocarban ($p = 0.001$) detected in the tank water versus the rooftop debris samples. Caffeine, salicylic acid and triclosan were then detected at higher concentrations in the tank water samples, while triclocarban was detected at higher concentrations in the rooftop debris samples.

2.3.6 Detection of the CST Markers in the Municipal Tap Water Samples

The prevalence of the CST markers analyzed in the current study, was then investigated in municipal tap water sampled intermittently from houses in the Kleinmond Housing scheme site. Overall, ten municipal tap water samples (denoted as MT) were collected, with three samples collected during sampling three (MT3.1, MT3.2 and MT3.3), three samples were collected during sampling four (MT4.1, MT4.2 and MT4.3) and four samples were collected during sampling five (MT5.1, MT5.2, MT5.3 and MT5.4) (**Fig. 2.5.**). Carbamazepine was BDL (< 0.1 µg/L) for all ten municipal tap water samples analyzed.

Caffeine was detected in all ten (100%) municipal tap water samples at an overall mean concentration of 2.50 µg/L. Caffeine then ranged from the mean concentration of 1.37 µg/L obtained for the samples collected during the third sampling occasion (MT3.1, MT3.2 and MT3.3) to 11.4 µg/L recorded for the samples collected during the fifth sampling occasion (MT5.1, MT5.2, MT5.3 and MT5.4). The lowest caffeine concentration of 0.90 µg/L was

detected in MT4.1, while the highest caffeine concentration of 4.70 µg/L was detected in MT5.1 (**Fig. 2.5.**).

Salicylic acid was also detected in all ten (100%) municipal tap water samples at an overall mean concentration of 13.4 µg/L. The salicylic acid concentration then ranged from a mean concentration of 11.6 µg/L obtained for the samples collected during the third sampling occasion (MT3.1, MT3.2 and MT3.3) to 15.4 µg/L recorded for the samples collected during the fourth sampling occasion (MT4.1, MT4.2 and MT4.3). The lowest salicylic acid concentration of 10.6 µg/L was detected in MT3.1, while the highest salicylic acid concentration of 18.4 µg/L was detected in MT4.3 (**Fig. 2.5.**).

Methylparaben was detected in all ten (100%) municipal tap water samples at an overall mean concentration of 6.32 µg/L. The methylparaben concentration then ranged from a mean concentration of 1.37 µg/L obtained for the samples collected during the third sampling occasion (MT3.1, MT3.2 and MT3.3) to 11.4 µg/L recorded for the samples collected during the fifth sampling occasion (MT5.1, MT5.2, MT5.3 and MT5.4). The lowest methylparaben concentration of 0.90 µg/L was detected in MT4.1, while the highest methylparaben concentration of 15.5 µg/L was detected in MT5.3 (**Fig. 2.5.**).

Triclosan was detected in all ten (100%) municipal tap water samples at an overall mean concentration of 1.58 µg/L. The triclosan concentration then ranged from a mean concentration of 0.80 µg/L obtained for the samples collected during the third sampling occasion (MT3.1, MT3.2 and MT3.3) to 2.30 µg/L recorded for the samples collected during the fifth sampling occasion (MT5.1, MT5.2, MT5.3 and MT5.4). The lowest triclosan concentration of 0.60 µg/L was detected in MT4.1, while the highest triclosan concentration of 2.80 µg/L was detected in MT5.3 (**Fig. 2.5.**).

Triclocarban was detected in all ten (100%) municipal tap water samples at an overall mean concentration of 0.74 µg/L. Triclocarban then ranged from a mean concentration of 0.70 µg/L obtained for the samples collected during the fifth sampling occasion (MT5.1, MT5.2, MT5.3 and MT5.4) to 0.77 µg/L recorded for the samples collected during the third (MT3.1, MT3.2 and MT3.3) and fourth (MT4.1, MT4.2 and MT4.3) sampling occasions. The lowest triclocarban concentration of 0.60 µg/L was detected in MT5.2 and MT5.4, while the highest triclocarban concentration of 0.90 µg/L was detected in MT5.3 (**Fig. 2.5.**).

Acetaminophen was detected in all ten (100%) municipal tap water samples at an overall mean concentration of 9.21 µg/L. Acetaminophen then ranged from a mean concentration of 8.53 µg/L obtained for the samples collected during the fourth sampling occasion (MT4.1, MT4.2 and MT4.3) to 9.73 µg/L recorded for the samples collected during the fifth sampling occasion (MT5.1, MT5.2, MT5.3 and MT5.4). The lowest acetaminophen concentration of 7.30 µg/L was

detected in MT4.2, while the highest acetaminophen concentration of 10.1 µg/L was detected in MT5.1 (**Fig. 2.5.**).

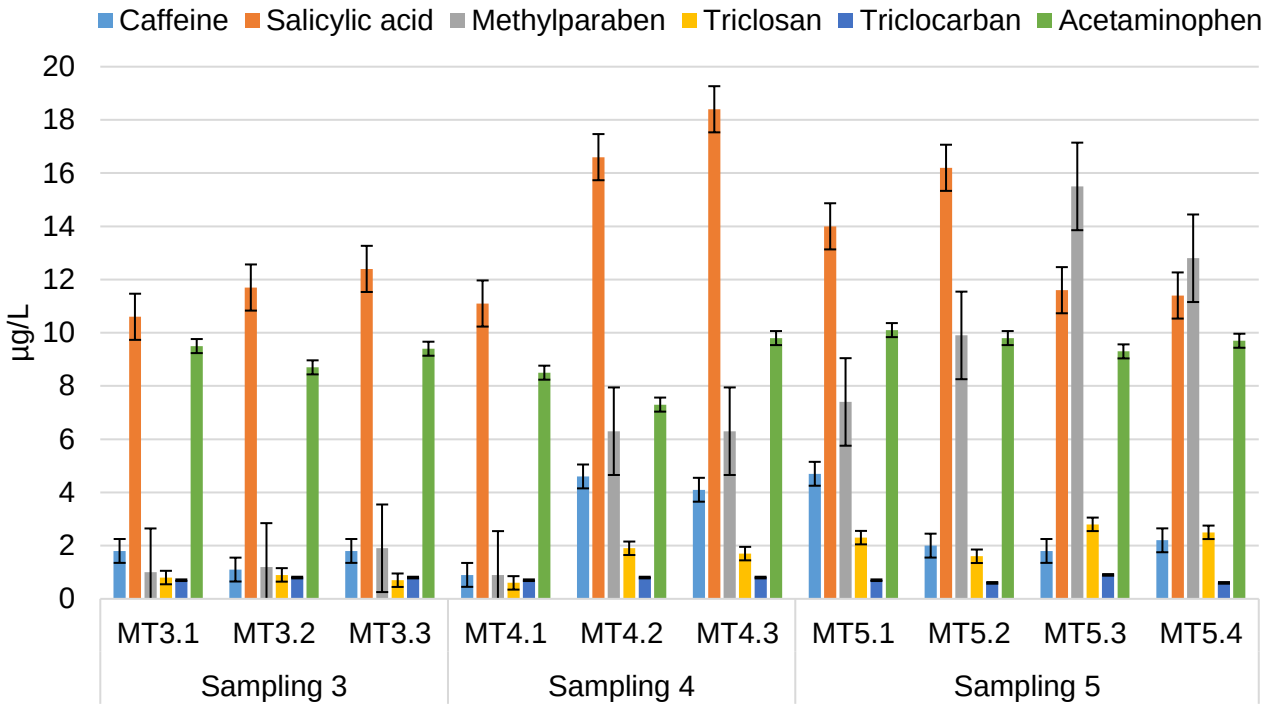


Fig. 2.5. Concentrations of the CST markers detected in the municipal tap water samples (MT3.1 – MT3.3; MT4.1 – MT4.3; MT5.1 – MT5.4) collected during sampling three, four and five with the standard error indicated with error bars.

The *t*-test analysis then indicated a significant difference between the mean methylparaben concentrations detected during the third and fourth sampling occasions versus the fifth sampling occasion, with the concentration of methylparaben recorded during sampling five (11.4 µg/L) significantly higher than the concentrations detected during sampling three (1.37 µg/L; $p = 0.0049$) and sampling four (4.50 µg/L; $p = 0.043$). In addition, a significant difference between the mean concentration of triclosan detected during the third sampling occasion versus the fifth sampling occasion was observed, with the mean concentration of triclosan being significantly higher ($p = 0.0044$) during the fifth sampling occasion (2.30 µg/L) in comparison to the concentration detected during sampling three (0.80 µg/L). No other significant differences between the concentrations of the other CST markers detected during the three sampling occasions were observed.

2.3.7 Correlations Between the MST, CST Markers and Indicator Organisms

Correlation analysis was performed for the indicator organisms enumerated with culturing techniques, the MST markers, the CST markers, *E. coli* (quantified by qPCR) and enterococci (quantified by qPCR) detected in the tank water samples. Correlation analysis was also

performed for the indicator organisms quantified with the qPCR assays (*E. coli* and enterococci), the MST and CST markers detected in the rooftop debris samples. Carbamazepine was BDL (< 0.1 µg/L) in all the rooftop debris and tank water samples, hence this marker was not included in the correlation analysis. For the tank water samples, all correlation analyses between the various MST markers and indicator organisms are summarized in **Table 2.4,** all correlation analyses between the various CST markers and indicator organisms are summarized in **Appendix Table A4** and all the correlation analyses conducted between the MST and CST markers are summarized in **Appendix Table A5**. For the rooftop debris samples, all correlation analyses between the various MST markers, the various CST markers and indicator organisms are summarized in **Appendix Table A6.**

No significant correlations were observed between the various MST markers analyzed in the tank water samples (collected throughout the study period) (**Table 2.4**). In contrast, for the correlation analysis conducted between the CST markers, significant positive correlations were observed for; salicylic acid versus acetaminophen (r = 0.316; p = 0.018) (**Appendix Table A4**), and salicylic acid versus triclocarban (r = 0.339; p = 0.011) (**Appendix Table A4**) as well as for triclocarban versus triclosan (r = 0.352; p = 0.008) (**Appendix Table A4**). Correlation analysis for the MST markers versus the CST markers detected in the tank water samples then revealed a significant positive correlation between *Bacteroides* HF183 versus triclosan (r = 0.282; p = 0.036) (**Appendix Table A5**). No other significant correlations were observed for the MST markers versus the CST markers identified in the tank water samples (**Appendix Table A5**).

Correlation analysis was also conducted between the indicator organisms *E. coli*, total and fecal coliforms, enterococci and heterotrophic bacteria (enumerated using culture based analysis) versus the qPCR analysis of *E. coli* and enterococci detected in the tank water samples. These correlations are presented in **Table 2.4** and **Fig. 2.6**. For the culture based analysis, significant positive correlations were observed for the total coliforms versus enterococci (r = 0.609; p = 0.000); total coliforms versus fecal coliforms (r = 0.64; p = 0.000); and fecal coliforms versus enterococci (r = 0.589; p = 0.000) (**Table 2.4; Fig. 2.6.**). In addition, positive correlations were observed for enterococci quantified by qPCR versus; total coliforms (r = 0.473; p = 0.000), enterococci (enumerated with culturing techniques) (r = 0.401; p = 0.002) and fecal coliforms (r = 0.437; p = 0.001) (**Table 2.4; Fig. 2.6.**).

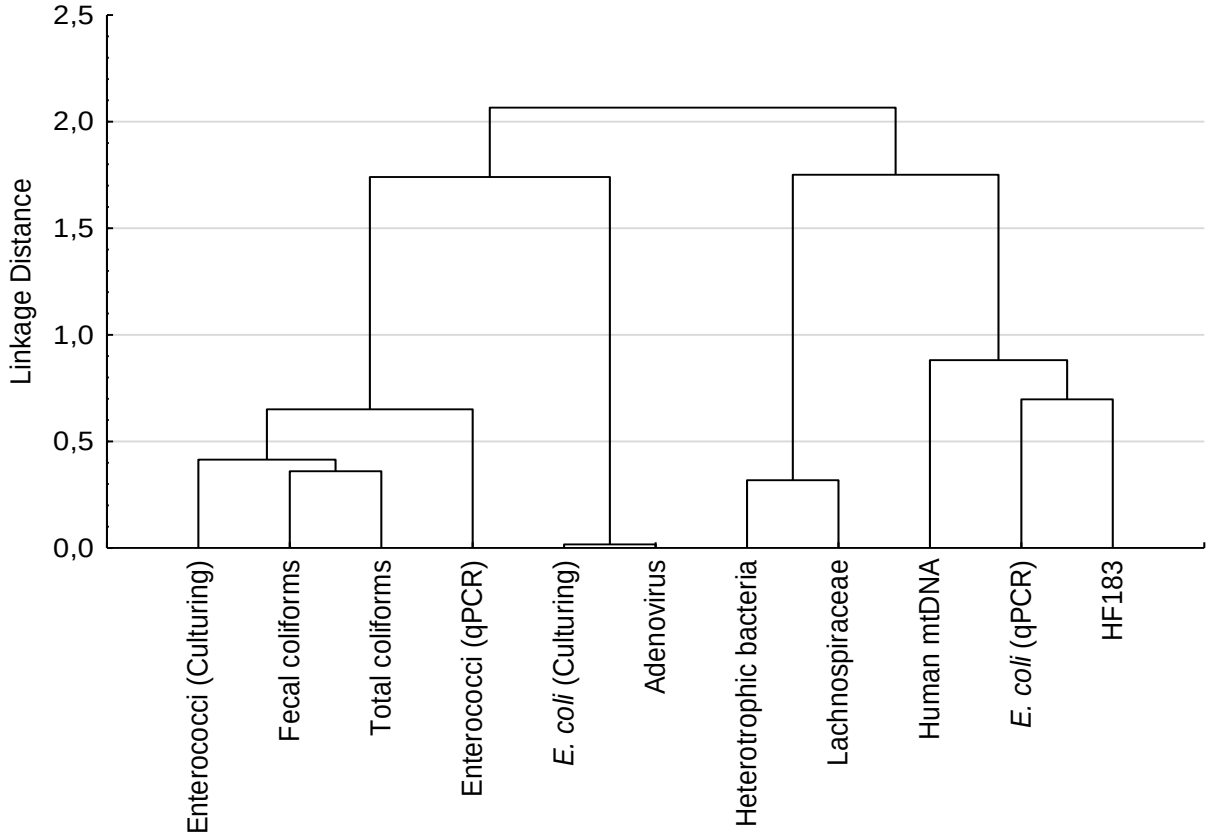


Fig. 2.6. Dendrogram based on Cluster Analysis with Ward's Methods of the MST markers versus the indicator organisms detected in the tank water samples.

For the correlation analysis of the MST markers versus the indicator organisms enumerated utilizing culturing techniques and the indicator organisms quantified with the qPCR assays in the tank water samples, significant positive correlations were observed for adenovirus versus *E. coli* (enumerated with culturing techniques) (r = 0.983; p = 0.000) and for Lachnospiraceae versus the heterotrophic bacteria count (r = 0.682; p = 0.000) (**Table 2.4; Fig. 2.6.**). In addition, a positive correlation was observed between *Bacteroides* HF183 and *E. coli* (quantified by qPCR) (r = 0.303; p = 0.023) (**Fig. 2.6.**) and human mtDNA versus enterococci (enumerated with culturing techniques) (r = 0.297; p = 0.026) (**Table 2.4**). Human mtDNA then clustered with *E. coli* (quantified by qPCR) and *Bacteroides* HF183, however human mtDNA did not significantly correlate with *Bacteroides* HF183 (r = 0.0875; p = 0.522) or *E. coli* (quantified by qPCR) (r = 0.242; p = 0.073) (**Table 2.4; Fig. 2.6.**).

Correlations analysis for the CST markers versus the indicator organisms enumerated using culture based analysis (*E. coli*, total and fecal coliforms, enterococci and heterotrophic bacteria) and the indicator organisms quantified with the qPCR assays (*E. coli* and enterococci) in the tank water samples, are presented in **Appendix Table A4** and **Fig. 2.7**. Significant positive correlations were then observed for caffeine versus; enterococci (quantified by qPCR) (r = 0.863; p = 0.000); total coliforms (r = 0.483; p = 0.000); fecal coliforms (r = 0.447; p = 0.001)

and enterococci (enumerated with culturing techniques) (r = 0.399; p = 0.002) (**Fig. 2.7.**). In addition, salicylic acid positively correlated with total coliforms (r = 0.301; p = 0.024) (**Appendix Table A4**). Negative correlations between the CST markers and indicator organisms were also observed for salicylic acid versus *E. coli* (quantified by qPCR) (r = -0.295; p = 0.028); acetaminophen versus fecal coliforms (r = -0.267; p = 0.047); triclocarban versus *E. coli* (quantified by qPCR) (r = -0.291; p = 0.029); and triclosan versus heterotrophic bacteria (r = -0.359; p = 0.007) (**Appendix Table A4**). Moreover, no significant correlations were observed for methylparaben versus heterotrophic bacteria (r = -0.0664; p = 0.627), salicylic acid (r = 0.133; p = 0.33) or acetaminophen (r = 0.100; p = 0.463). Methylparaben and heterotrophic bacteria however clustered with salicylic acid and acetaminophen [salicylic acid and acetaminophen correlated significantly (r = 0.316; p = 0.018)] (**Fig. 2.7.**).

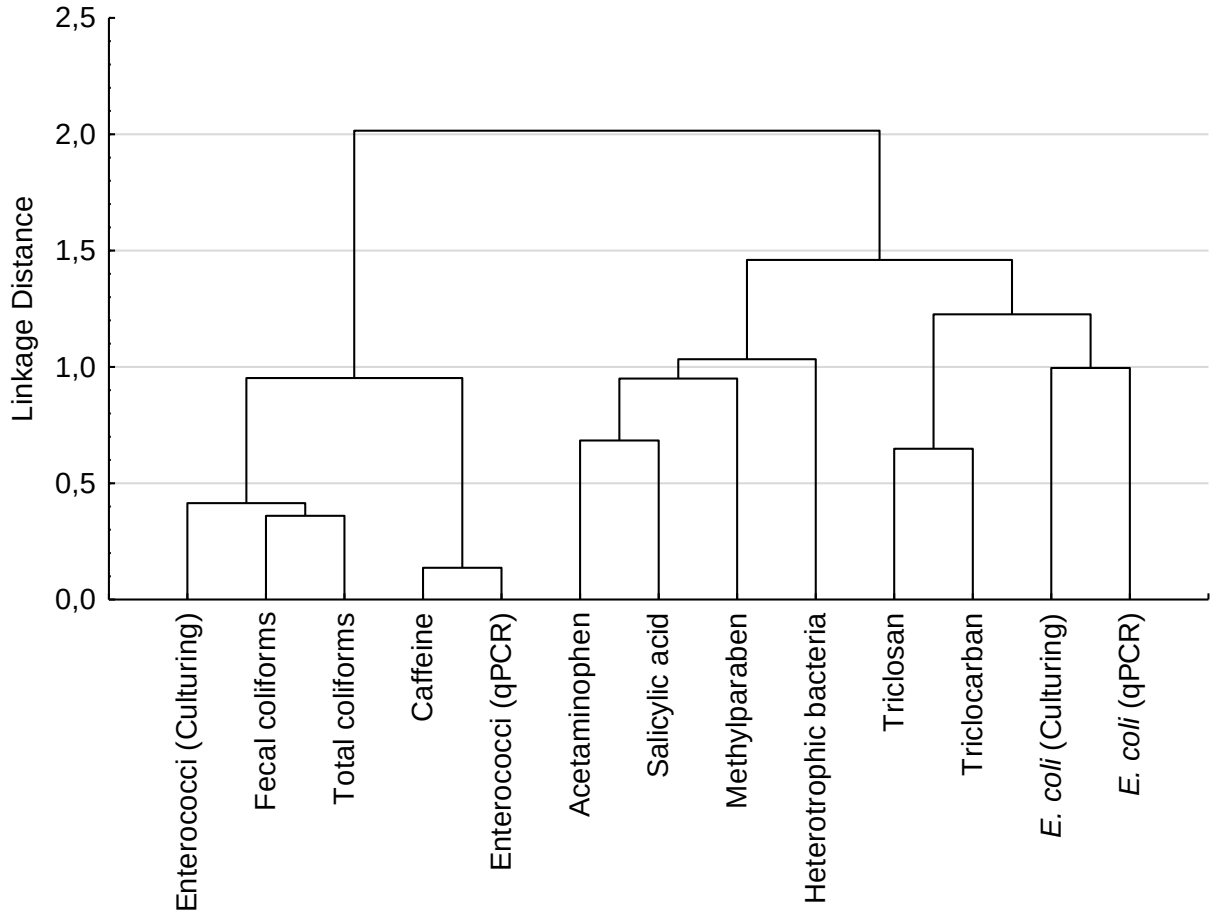


Fig. 2.7. Dendrogram based on Cluster Analysis with Ward's Methods of the CST markers versus the indicator organisms detected in the tank water samples.

Table 2.4 Summary of the correlations observed between the indicator organisms and MST markers detected in the tank water samples with the significant correlations ($p < 0.05$) indicated in red.

	HF183	Adenovirus	Human mtDNA	Lachno-spiraceae	*E. coli* (qPCR)	Enterococci (qPCR)	*E. coli* (culturing)	Total coliforms	Enterococci (culturing)	Fecal coliforms	Hetero-trophic bacteria
HF183	1	-0.0927	0.0875	0.0233	0.303	0.0278	-0.0608	-0.2184	-0.0115	-0.147	-0.319
	p = ---	*p* = 0.497	*p* = 0.522	*p* = 0.865	*p* = 0.023	*p* = 0.839	*p* = 0.656	*p* = 0.106	*p* = 0.933	*p* = 0.278	*p* = 0.017
Adenovirus	-0.0927	1	-0.0536	0.0588	-0.0237	0.0439	0.983	0.188	0.199	0.169	-0.0646
	p = 0.497	*p* = ---	*p* = 0.695	*p* = 0.667	*p* = 0.862	*p* = 0.748	*p* = 0.000	*p* = 0.166	*p* = 0.142	*p* = 0.214	*p* = 0.636
Human mtDNA	0.0875	-0.0536	1	-0.0335	0.242	-0.0979	-0.0471	0.03	0.297	-0.102	-0.0372
	p = 0.522	*p* = 0.695	*p* = ---	*p* = 0.806	*p* = 0.073	*p* = 0.473	*p* = 0.730	*p* = 0.826	*p* = 0.026	*p* = 0.453	*p* = 0.786
Lachno-spiraceae	0.0233	0.0588	-0.0335	1	0.0734	0.0638	0.0701	-0.0824	-0.147	-0.0846	0.682
	p = 0.865	*p* = 0.667	*p* = 0.806	*p* = ---	*p* = 0.591	*p* = 0.640	*p* = 0.608	*p* = 0.546	*p* = 0.280	*p* = 0.535	*p* = 0.000
E. coli (qPCR)	0.303	-0.0237	0.242	0.0734	1	-0.004	0.0045	-0.0487	0.074	0.154	-0.14
	p = 0.023	*p* = 0.862	*p* = 0.073	*p* = 0.591	*p* = ---	*p* = 0.977	*p* = 0.974	*p* = 0.722	*p* = 0.588	*p* = 0.259	*p* = 0.303
Enterococci (qPCR)	0.0278	0.0439	-0.0979	0.0638	-0.004	1	0.0386	0.473	0.401	0.437	0.137
	p = 0.839	*p* = 0.748	*p* = 0.473	*p* = 0.640	*p* = 0.977	*p* = ---	*p* = 0.778	*p* = 0.000	*p* = 0.002	*p* = 0.001	*p* = 0.313
E. coli (culturing)	-0.0608	0.983	-0.0471	0.0701	0.0045	0.0386	1	0.223	0.242	0.216	-0.0377
	p = 0.656	*p* = 0.00	*p* = 0.730	*p* = 0.608	*p* = 0.974	*p* = 0.778	*p* = ---	*p* = 0.098	*p* = 0.073	*p* = 0.109	*p* = 0.783
Total coliforms	-0.218	0.188	0.03	-0.0824	-0.0487	0.473	0.223	1	0.609	0.64	0.0856
	p = 0.106	*p* = 0.166	*p* = 0.826	*p* = 0.546	*p* = 0.722	*p* = 0.000	*p* = 0.098	*p* = ---	*p* = 0.000	*p* = 0.000	*p* = 0.531
Enterococci (culturing)	-0.0115	0.199	0.297	-0.1469	0.074	0.401	0.242	0.609	1	0.589	0.0164
	p = 0.933	*p* = 0.142	*p* = 0.026	*p* = 0.280	*p* = 0.588	*p* = 0.002	*p* = 0.073	*p* = 0.000	*p* = ---	*p* = 0.000	*p* = 0.904
Fecal coliforms	-0.147	0.169	-0.102	-0.0846	0.154	0.437	0.216	0.64	0.589	1	0.0436
	p = 0.278	*p* = 0.214	*p* = 0.453	*p* = 0.535	*p* = 0.259	*p* = 0.001	*p* = 0.109	*p* = 0.000	*p* = 0.000	*p* = ---	*p* = 0.750
Heterotrophic bacteria	-0.319	-0.0646	-0.0372	0.682	-0.14	0.137	-0.0377	0.0856	0.0164	0.0436	1
	p = 0.017	*p* = 0.636	*p* = 0.786	*p* = 0.000	*p* = 0.303	*p* = 0.313	*p* = 0.783	*p* = 0.531	*p* = 0.904	*p* = 0.750	*p* = ---

For the rooftop debris samples, the correlations between the MST markers, CST markers and the indicator organisms (*E. coli* and enterococci quantified by qPCR) are presented in **Appendix Table A6**. For the MST markers versus the MST markers, significant positive correlations were observed for adenovirus versus the HF183 marker (r = 0.40; p = 0.002) and adenovirus versus Lachnospiraceae (r = 0.318; p = 0.017) (**Fig. 2.8.; Appendix Table A6**). Human mtDNA was then related to *E. coli* and enterococci based on the Cluster Analysis (**Fig. 2.8.**), however significant correlations were not observed for human mtDNA versus *E. coli* (r = 0.104; p = 0.447) or human mtDNA versus enterococci (r = -0.0495; p = 0.717) (**Appendix Table A6**).

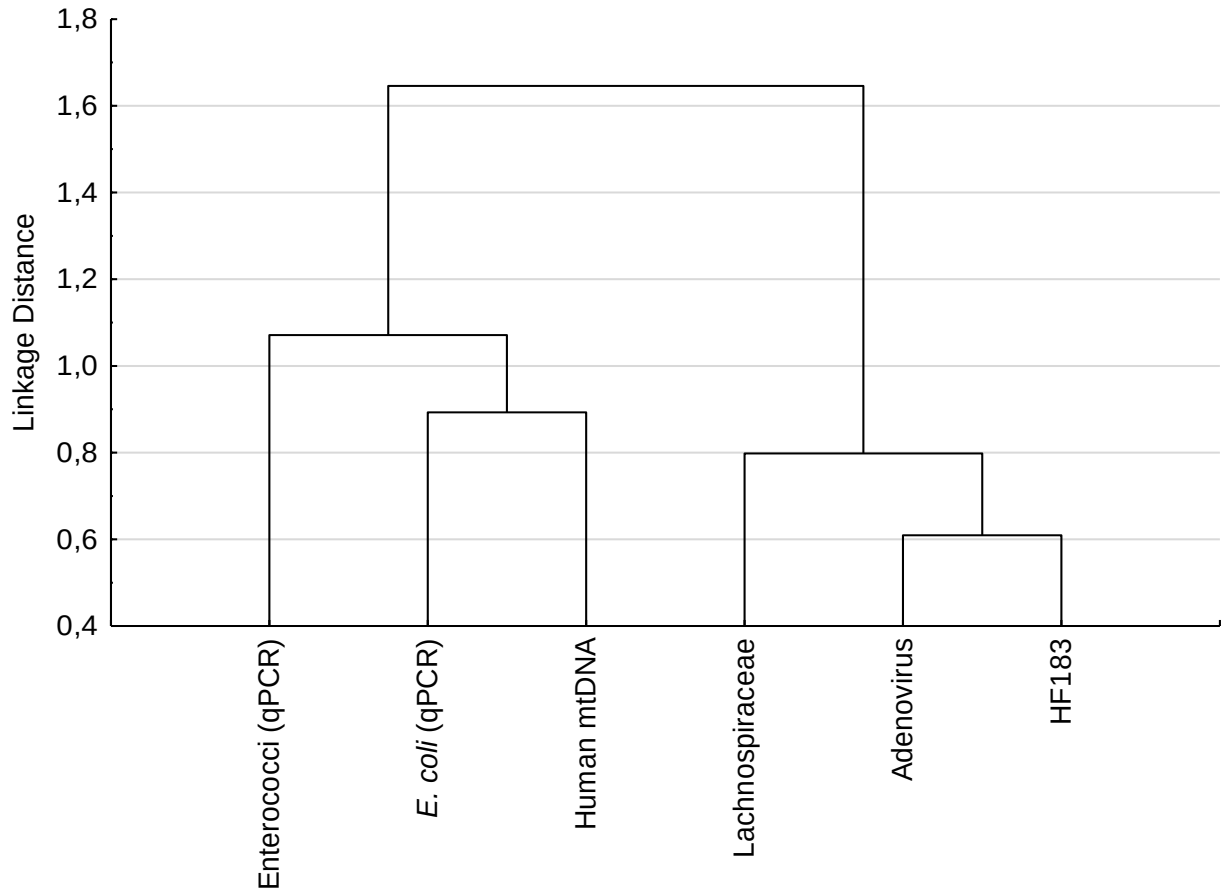


Fig. 2.8. Dendrogram based on Cluster Analysis with Ward's Methods of the MST markers versus the indicator organisms detected in the rooftop debris samples.

For the CST markers detected in the rooftop debris samples, significant positive correlations were observed for salicylic acid versus; caffeine (r = 0.267; p = 0.047); methylparaben (r = 0.470; p = 0.000) and acetaminophen (r = 0.278; p = 0.038) (**Appendix Table A6; Fig. 2.9.**). In addition, significant positive correlations were observed for acetaminophen versus caffeine (r = 0.358; p = 0.007) (**Appendix Table A6; Fig. 2.9.**). Moreover, triclocarban positively correlated with caffeine (r = 0.486; p = 0.000) and triclosan (r = 0.381; p = 0.004), while triclosan

also positively correlated with methylparaben (r = 0.324; p = 0.015) (**Appendix Table A6; Fig. 2.9**).

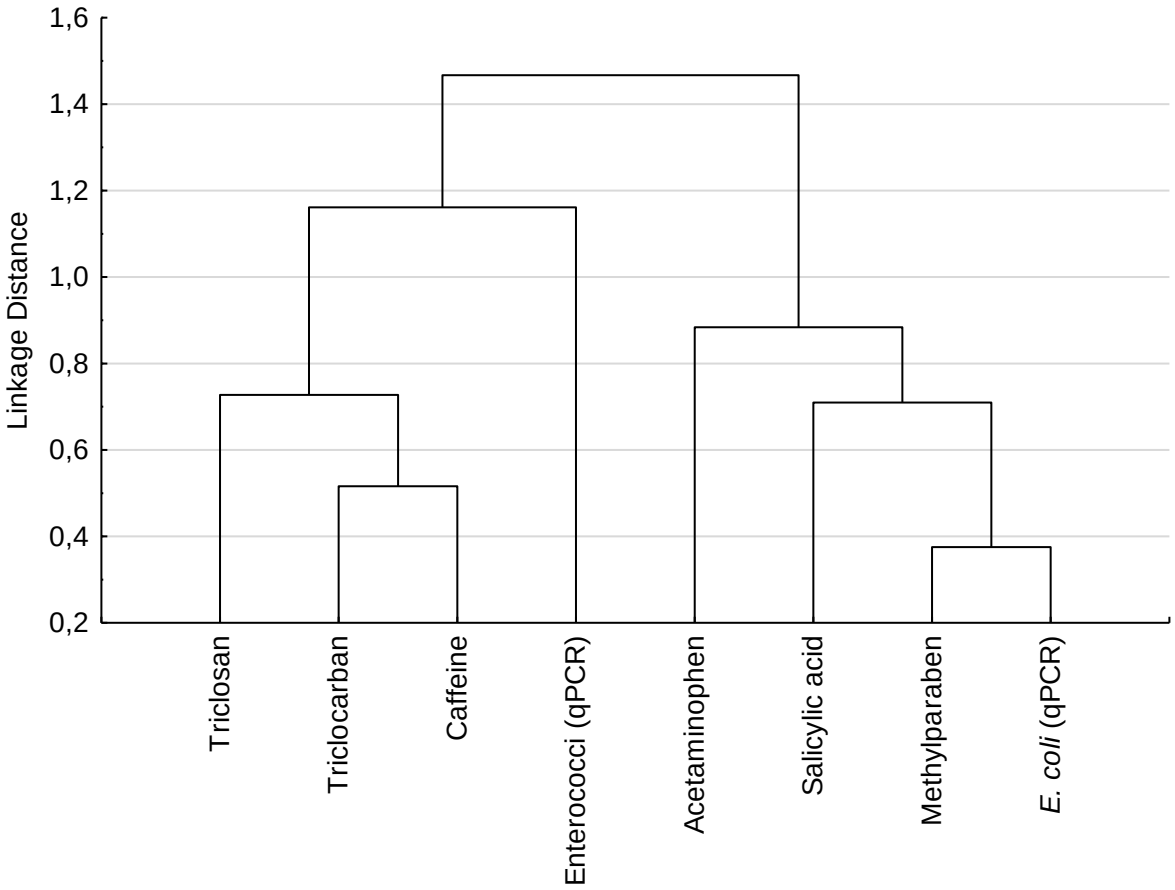


Fig. 2.9. Dendrogram based on Cluster Analysis with Ward's Methods of the CST markers versus the indicator organisms detected in the rooftop debris samples.

For the MST markers versus the CST markers detected in the rooftop debris samples, significant positive correlations were observed for Lachnospiraceae versus triclocarban (r = 0.508; p = 0.000) and Lachnospiraceae versus triclosan (r = 0.299; p = 0.025) (**Appendix Table A6**). In addition, significant negative correlations were observed for *Bacteroides* HF183 versus methylparaben (r = -0.294; p = 0.028) and Lachnospiraceae versus acetaminophen (r = -0.293; p = 0.028) (**Appendix Table A6**).

Indicator organism analysis utilizing culture based methods was not conducted on the rooftop debris samples and therefore correlations for indicator organisms quantified with the qPCR assays (*E. coli* and enterococci) only are presented for the rooftop debris samples. The results then indicated that no significant correlations were observed between *E. coli* and enterococci (r = -0.0116; p = 0.933) quantified with the qPCR assays in the rooftop debris samples (**Appendix Table A6**). In addition, none of the MST markers quantified in the rooftop debris samples exhibited significant correlations with *E. coli* (quantified by qPCR) or enterococci

(quantified by qPCR) (**Appendix Table A6**). For the CST markers versus the two indicator organisms detected by qPCR analysis in the rooftop debris samples, the only significant positive correlations were observed between *E. coli* (quantified by qPCR) versus methylparaben ($r = 0.623$; $p = 0.000$) and *E. coli* (quantified by qPCR) versus salicylic acid ($r = 0.273$; $p = 0.042$) (**Fig. 2.9.; Appendix Table A6**). The Cluster Analysis then revealed that acetaminophen clustered with salicylic acid, methylparaben and *E. coli* as a result of the positive correlations observed for acetaminophen versus caffeine ($r = 0.358$; $p = 0.007$) and salicylic acid ($r = 0.278$; $p = 0.038$), however acetaminophen did not significantly correlate with *E. coli* (quantified by qPCR) ($r = 0.220$; $p = 0.104$). The Cluster Analysis also showed that caffeine, triclocarban and triclosan clustered with enterococci, however no significant correlations between these CST markers and enterococci were observed (**Fig. 2.9.; Appendix Table A6**).

Pearson's correlation analysis was also performed to determine if there were any relationships between the individual MST markers, CST markers, *E. coli* (quantified by qPCR) and enterococci (quantified by qPCR) detected in the tank water samples versus the detection of the same marker or indicator organism in the rooftop debris samples (**Appendix Table A7**). Of the MST markers, only HF183 detected in the tank water samples displayed a significant positive correlation ($r = 0.50$; $p = 0.000$) with the HF183 marker detected in the rooftop debris samples (**Appendix Table A7**). In contrast, Lachnospiraceae detected in the tank water samples displayed a significant negative correlation ($r = -0.29$; $p = 0.028$) with the Lachnospiraceae detected in the rooftop debris samples. Adenovirus ($r = 0.09$; $p = 0.528$), human mtDNA ($r = -0.04$; $p = 0.774$), *E. coli* ($r = -0.19$; $p = 0.166$) and enterococci ($r = 0.15$; $p = 0.283$) were not significantly correlated between the tank water and rooftop debris samples. Moreover, the CST markers caffeine ($r = -0.09$; $p = 0.489$), salicylic acid ($r = 0.07$; $p = 0.618$), triclocarban ($r = -0.15$; $p = 0.255$), triclosan ($r = 0.21$; $p = 0.125$), acetaminophen ($r = -0.09$; $p = 0.503$) and methylparaben ($r = 0.05$; $p = 0.707$) were also not significantly correlated between the tank water and rooftop debris samples (**Appendix Table A7**).

In addition to the relationships investigated between the indicator organisms, MST and CST markers detected in the tank water and rooftop debris samples, correlations between the rainfall observed for the sampling period and the detection of these markers and indicator organisms in the tank water (**Appendix Table A8)** and rooftop debris (**Appendix Table A9**) samples were also investigated with Pearson's correlations. Significant positive correlations were observed between the rainfall data versus fecal coliforms ($r = 0.89$; $p = 0.018$) detected in the tank water samples (**Appendix Table A8)**. In addition, significant positive correlations were observed for the rainfall data versus; caffeine ($r = 0.83$; $p = 0.043$) and triclocarban ($r = 0.83$; $p = 0.043$) detected in the rooftop debris samples (**Appendix Table A9**). No other significant correlations

between the MST markers, CST markers and the indicator organisms versus the rainfall data were observed.

Furthermore, Pearson's correlations were conducted to determine whether there were any correlations between the rainfall observed during the sampling period and the concentrations of the CST markers detected in the municipal tap water samples. Results then indicated that there were no significant correlations between the rainfall data and the concentrations of caffeine (r = 0.47; p = 0.17), salicylic acid (r = 0.57; p = 0.089), methylparaben (r = 0.25; p = 0.48), triclosan (r = 0.31; p = 0.38), triclocarban (r = -0.0035; p = 0.99) or acetaminophen (r = -0.32; p = 0.37) detected in the municipal tap water samples.

2.4 Discussion

The current study focused on detecting a wide range of MST and CST markers in DRWH systems in order to identify a toolbox of markers that may be applied to monitor the quality of harvested rainwater in future screenings of this water source. In addition, to ascertain whether the detected MST and CST markers can be employed to supplement indicator organism analysis for the monitoring of harvested rainwater, the correlations between these alternative contamination indicators and traditional indicator organisms were evaluated.

Culture based analysis of the traditional indicator organisms indicated that the heterotrophic bacteria and total coliform counts exceeded the recommended drinking water guidelines in all the tank water samples (100%; n = 60) analyzed (DWAF, 1996; SABS, 2005; NHMRC & NRMMC, 2011; WHO, 2011). Similarly, the *E. coli* counts exceeded the recommended drinking water guidelines in 95% (n = 57) of the tank water samples, fecal coliform counts exceeded the recommended guideline in 85% (n = 51) of the tank water samples analyzed, while enterococci were only sporadically detected during the sampling period. *Escherichia coli*, enterococci and fecal coliforms are generally utilized to indicate fecal pollution of a water source by warm-blooded animals, while total coliforms are utilized as an indicator of the general hygienic quality of the water (DWAF, 1996; SABS, 2005; NHMRC & NRMMC, 2011; WHO, 2011). The heterotrophic bacteria count is then utilized to assess the general bacterial load in water sources, where higher numbers may indicate an increased risk of contracting a disease when utilizing the water for domestic purposes (DWAF, 1996; SABS, 2005). Based on the indicator organism analysis of the tank water samples, it is thus hypothesized that the water may be contaminated by fecal material deposited on the rooftops by animals, rodents, and birds, amongst others. Ahmed et al. (2012) found that animal fecal matter on rooftops may significantly contribute to the contamination of the stored tank water as pathogens such as *Campylobacter* spp., *Giardia* spp. and *Salmonella* spp. were detected in harvested rainwater as well as in bird and possum fecal samples collected from the rooftops utilized to capture the rainwater. In addition, the detection of indicator organisms in harvested rainwater from tanks

located in the Kleinmond Housing Scheme site corroborates data obtained by Dobrowsky et al., (2014a) where indicator organisms were also routinely detected in tank water at this sampling site.

Bacteroides HF183, adenovirus, Lachnospiraceae and human mtDNA were identified as the prevalent MST markers in both the rooftop debris and tank water samples by conventional PCR analysis. Quantitative PCR analysis was then conducted in order to quantify these markers, with results indicating that *Bacteroides* HF183, Lachnospiraceae and adenovirus were present in all the tank water (100%; n = 60) and rooftop debris (100%; n = 60) samples analyzed. This was in contrast to the conventional PCR assays where the HF183 marker was only detected in 86.7% and 63.3%, adenovirus was detected in 66.7% (tank water and rooftop debris) and Lachnospiraceae was detected in 55% and 81.7% of the tank water and rooftop debris samples, respectively. For the human mtDNA, the qPCR assays also displayed greater sensitivity to detect this marker as a 90% and 91.7% detection frequency in the tank water and rooftop debris samples, respectively, was observed, compared to the detection of the marker in 35% of the tank water samples and 57% of the rooftop debris samples by conventional PCR analysis. It is hypothesized that the difference in the frequency of detection observed for the conventional PCR assays versus the qPCR assays in the current study, could be due to the dilution of the extracted DNA prior to analysis with the qPCR assays which may have minimized PCR inhibition. This has previously been noted in a study conducted by Cao et al. (2012), where the qPCR analysis of serially diluted DNA was conducted. The results indicated that qPCR inhibition was resolved between 78 – 100% for different qPCR assays with a five-fold dilution of the DNA samples. In addition, it has been reported that qPCR is a more sensitive technique in comparison to conventional PCR, which may also account for the increased detection frequencies observed for the qPCR analyses in the current study (Tuma et al., 1999; Schmittgen et al., 2000; Rengarajan et al., 2002; Boyle et al., 2004; Seurinck et al., 2005). For example, Seurinck et al. (2005) compared conventional PCR with qPCR for the detection of the *Bacteroides* HF183 marker in fresh water samples. Utilizing qPCR analysis, the HF183 marker was detected in 83% of the samples analyzed, while the conventional PCR assays only detected the HF183 marker in 67% of the samples analyzed. Similar results were thus obtained in the current study, with the qPCR analysis proving more sensitive for the detection of the dominant MST markers. However, implementing qPCR assays to routinely screen for these markers in rainwater harvesting systems may become expensive in terms of reagents and consumables. In addition, access to qPCR equipment may be limited and could thus complicate the use of the technology for the routine screening of water samples. Based on the results obtained in the current study it is thus recommended that qPCR assays be utilized to supplement or corroborate conventional PCR analysis for the screening or monitoring of water samples.

The MST markers with the highest frequency of detection in the current study i.e. *Bacteroides* HF183, human mtDNA, adenovirus and Lachnospiraceae, have also previously been successfully applied to identify contamination sources in various water supplies. For example, *Bacteroides* HF183 has been applied to rivers, streams, lakes, ponds, stormwater run-off, harvested rainwater and coastal waters, to identify sources of contamination linked to human activity or wastewater discharge (Sidhu et al., 2013; Kabiri et al., 2016; Kirs et al., 2016; Waso et al., 2016; Wilson et al., 2016). In addition, various adenoviruses such as bovine and human adenovirus have been employed to screen stormwater run-off, harvested rainwater and river water, amongst others, for fecal contamination (Sidhu et al., 2013; Rusinõl et al., 2016; Waso et al., 2016). Moreover, Lachnospiraceae has been applied to marine water sources and human mtDNA has been applied to agricultural run-off and river water to identify the presence of wastewater and human fecal contamination in these water sources (Martellini et al., 2005; Caldwell et al., 2007; Newton et al., 2011; Villemur et al., 2015). In a study conducted by Templar et al. (2016), *Bacteroides* HF183 and Lachnospiraceae were applied to river water (collected from three rivers that drain into Lake Michigan, USA) in order to assess whether wastewater was being discharged into the river system. The results then indicated that sewage was the major source of pollution of the three rivers as *Bacteroides* HF183 and Lachnospiraceae were detected throughout the sampling period in the river water samples. Moreover, in a study conducted by Sidhu et al. (2013) human adenovirus was applied to stormwater run-off in combination with the MST markers *Bacteroides* HF183, *M. smithii nif*H and polyomavirus and the CST markers caffeine, salicylic acid, acetaminophen and acesulfame to ascertain whether wastewater was contaminating stormwater run-off. Based on the results obtained the authors concluded that human adenovirus may be a promising indicator of sewage contamination in captured stormwater and could be useful in future microbial risk assessment studies as an indicator of pathogenic viruses associated with the investigated water source (Sidhu et al., 2013). Based on the MST analysis results obtained in the current study it is thus hypothesized that the tank water at the Kleinmond Housing Scheme site may be contaminated with fecal matter and sewage waste. This result thus corroborates the indicator organism analysis, where counts regularly contravened national and international water quality guidelines.

Triclosan, triclocarban, methylparaben, caffeine and acetaminophen were then the dominant CST markers detected in the tank water and rooftop debris samples analyzed (100%; n = 60). Salicylic acid was also detected in all the rooftop debris samples and 98% (n = 59) of the tank water samples. Triclosan and triclocarban are common antimicrobial compounds associated with toothpaste, hand soaps, dishwashing liquids, liquid soaps, cosmetics, shampoos, deodorants and mouthwash (Cox, 1987; Jones et al., 2000; Yalavarthy et al., 2015). In addition, methylparaben is a preservative associated with various cosmetics and personal care products (Lillo et al., 2016). Moreover, caffeine is a common compound associated with coffee and other

food products regularly consumed by humans and it may also be associated with pharmaceuticals where it is thought to enhance the effect of the medication (Heberer et al., 2002; Buerge et al., 2003). Lastly, salicylic acid and acetaminophen are commonly associated with over-the-counter medications that may be obtained from pharmacies without a prescription. Salicylic acid may also be associated with various personal care products, especially face wash and scrubs utilized for the treatment of acne (Sidhu et al., 2013). These CST markers are thus widely associated with anthropogenic activities, household waste and sewage. During the sampling sessions it was then observed that various household waste and garbage bags were placed around and on top of the DRWH tanks at the sampling site in Kleinmond (**Appendix Fig. A1**)**.** Researchers have then previously noted that anthropogenic activities and waste in close proximity to DRWH systems may significantly influence the quality of the harvested rainwater as was observed in the current study (Mwenge Kahinda et al., 2007).

Additionally, municipal tap water was sporadically analyzed for the presence of the CST markers targeted in the current study. Caffeine, salicylic acid, methylparaben, triclosan, triclocarban and acetaminophen were then all detected in the ten municipal tap water samples collected during sampling three to five (**Fig. 2.5.**). Municipal tap water in the Kleinmond region is treated at the Kleinmond (Palmiet) Water Treatment plant (GPS coordinates: 34°19'56"S 19°1'8"E). The plant mainly treats water from the Palmiet River and to a lesser extent water from boreholes in the Kleinmond area, to produce drinking water. The Palmiet River passes through the Elgin valley which is known to produce large quantities of apples, plums, pears and grapes, often as export produce (Grabouw Tourism Buro, personal communication). The river then also flows through the town of Grabouw (Western Cape, South Africa). Hence, the river flows through an area greatly impacted by agricultural and anthropogenic activities, before it reaches the Kleinmond Water Treatment plant. Rapid gravity filters are then utilized for the treatment of the water at the Kleinmond Water Treatment plant, while the pH of the water is adjusted with lime dosing, coagulation is facilitated with alum dosing, flocculation is achieved with open channel mixing and sedimentation is facilitated in concrete settling tanks. The water is then disinfected with chlorine gas and stabilized with lime addition (Overstrand Municipality, personal communication). Literature has however, reported on the inefficacy of filtration, alum dosing and chlorination to remove chemical compounds, such as the CST markers screened for in the current study, from water sources (Liu & Wong, 2013). It is then recommended that advanced treatments such as ozonation and granular activated carbon adsorption may have to be implemented in order to increase the removal efficiency of these compounds (Boyd et al., 2003; Hagedorn & Weisberg, 2009; Liu & Wong, 2013; Wang & Wang, 2016). Additionally, in comparison to the other sampling occasions (three and four), significantly high concentrations of methylparaben [sampling three (p = 0.0049); sampling four (p = 0.043)] and triclosan [sampling three (p = 0.0044)] were then detected in the municipal tap water during sampling five. The

significant increase in the concentration of these compounds during the fifth sampling occasion may indicate an increase in the pollution of the Palmiet River by anthropogenic activities upstream from the water treatment plant. However, the correlation between an increase in the concentration of CST markers versus increased pollution associated with anthropogenic and agricultural activities in the river system needs to be investigated in future studies.

Comparison of the mean concentrations of the CST markers detected in the municipal tap water samples (MT) with the mean concentrations detected in the tank water samples (T) then indicated no significant differences in the caffeine ($p = 0.99$), salicylic acid ($p = 0.59$), triclosan ($p = 0.061$), triclocarban ($p = 0.21$) or acetaminophen ($p = 0.35$) concentrations between the municipal tap water and the tank water samples. In contrast, the *t*-test revealed a significant difference between the methylparaben concentrations detected in the municipal tap water and the tank water samples ($p = 0.0019$), with the concentrations detected in the municipal tap water samples significantly higher than the concentrations detected in the tank water samples. A number of studies have reported on the presence of chemical compounds associated with personal care products and pharmaceuticals in surface and municipal tap water sources around the world (Boyd et al., 2003; Buerge et al., 2003; Rodil et al., 2011; Qiao et al., 2011; Liu & Wong, 2013; Zhang et al., 2013; Alshakka et al., 2016). For example, Lorraine and Pettigrove (2006) reported on the presence of triclosan at an average concentration of 0.743 µg/L in municipal tap water in Southern California, USA. In addition, Valcárcel et al. (2011) reported on the occurrence of 33 chemical compounds in river water and tap water samples in Madrid, Spain. Carbamazepine and caffeine were both detected in the river water samples at concentrations of 67.7 µg/L and 1.4 µg/L, respectively, while these compounds were detected at ng/L concentrations (up to 75 ng/L) in tap water samples (Valcárcel et al., 2011). The findings in the current study are thus in accordance with literature reporting on the ubiquitous distribution of chemical compounds associated with pharmaceuticals and personal care products in aquatic environments, especially in areas impacted by anthropogenic and agricultural activities. It is however, imperative to note that the residents of the Kleinmond Housing Scheme site are exposed to µg/L levels of these chemical compounds regardless of whether they utilize the tank water or municipal tap water for domestic purposes. In addition, limited information on the long term human health risk associated with exposure to these CST markers at µg/L levels is available. It is thus crucial that future studies focus on elucidating the risk associated with the presence of these compounds, at the concentrations recorded, in water earmarked for domestic purposes. Furthermore, cost-effective treatment methods or intervention strategies for the removal of these compounds from water sources should be investigated and implemented. For example, research has indicated that first-flush diverters improve the physico-chemical quality of harvested rainwater (Gikas & Tsihrintzis, 2012) and may thus reduce the concentration of

chemical contaminants (such as the CST markers) being washed into the tanks from the roof surface.

Correlation analysis between all the indicator organisms (quantified with culture based techniques) versus the detection of *E. coli* and enterococci (qPCR analysis) in the tank water samples was then conducted (**Table 2.4**). Significant correlations were then observed between enterococci (quantified by qPCR) versus total coliforms ($r = 0.473$; $p = 0.000$), enterococci (enumerated with culturing techniques) ($r = 0.401$; $p = 0.002$) and fecal coliforms ($r = 0.437$; $p = 0.001$) (**Fig. 2.6.**), amongst others. The correlations observed between the various indicator organisms are in accordance with previous studies where correlations between these parameters have been noted (Korajkic et al., 2011; McQuaig et al., 2012; Liang et al., 2015). In a study conducted by Liang et al. (2015) significant positive correlations were observed for enterococci (quantified by qPCR) versus enterococci (detected with Enterolert) ($r = 0.67$; $p < 0.01$); *E. coli* (quantified by qPCR) versus enterococci (quantified by qPCR) ($r = 0.84$; $p < 0.01$) and *E. coli* (detected with Colilert) versus enterococci (detected with Enterolert) ($r = 0.84$; $p < 0.01$) in surface water sources in Singapore. In addition, McQuaig et al. (2012) noted significant correlations between fecal coliforms versus total coliforms ($r = 0.88$; $p < 0.05$); enterococci (enumerated with culturing techniques) versus total coliforms ($r = 0.85$; $p < 0.05$) and enterococci (enumerated with culturing techniques) versus fecal coliforms ($r = 0.86$; $p < 0.05$) in beach water samples. The positive correlations observed between the indicator organisms enumerated with the culturing techniques versus those quantified with the qPCR assays in the current study thus indicates that qPCR may also be employed to detect indicator organisms in water sources. In addition, while qPCR analysis does not indicate the number of viable indicator organisms present in a water source, the combination of qPCR assays with the nucleic acid binding dyes ethidium monoazide bromide (EMA) and propidium monoazide (PMA) could be utilized to detect DNA from viable cells (Elizaquível et al., 2011; Truchado et al., 2015; Reyneke et al., 2016; Strauss et al., 2016). Viability qPCR has successfully been utilized by our research group to estimate the proportion of viable *Legionella* spp. and *Pseudomonas* spp. in pasteurized and unpasteurized harvested rainwater (Reyneke et al., 2016; Strauss et al., 2016). The implementation of qPCR and viability qPCR assays to monitor water quality could thus improve monitoring efforts which will be beneficial for the water sector and for health risk assessment.

Moreover, in the current study, the enterococci quantified with the qPCR assays and enterococci enumerated with culturing techniques exhibited a significant positive correlation ($r = 0.401$; $p = 0.002$) in the tank water samples analyzed. Correlations between qPCR assays targeting the 23S rRNA sequence (also targeted in the current study) of enterococci and the culture based detection of enterococci have also previously been observed by Noble et al. (2010). It is also interesting to note that in the current study, enterococci were below the

detection limit in 31 tank water samples using culture based analysis, while enterococci were detected in all the tank water samples ranging from 2.01 gene copies/µL to 5.6×10^3 gene copies/µL utilizing qPCR analysis. Previous research has indicated that the 23S rRNA gene (detected by qPCR analysis in the current study) is a multi-copy gene and as a result the detection of this gene may overestimate the number of enterococci present in a water sample (Noble et al., 2010). However, enterococci gene copies were still detected in all the tank water samples, indicating that enterococci are more prevalent in the tank water samples than initially determined utilizing the culture based analysis.

In contrast, in the current study no significant correlation ($r = 0.0045$; $p = 0.974$) was observed between *E. coli* quantified with the qPCR assays and *E. coli* enumerated with the culturing techniques in the tank water samples. Using culture based analysis *E. coli* were detected in 57 tank water samples and were below the detection limit (< 1 CFU/100 mL) in three tank water samples: TH1; TB4 and TH5. Utilizing qPCR analysis, *uidA* gene copies of *E. coli* were then detected in these samples with 2.7×10^2 gene copies/µL detected in sample TH1, 1.1×10^3 gene copies/µL detected in sample TB4 and 1.9×10^3 gene copies/µL detected in sample TH5. Quantitative PCR assays however, detect DNA from viable but non-culturable (VBNC) and culturable cells, while culture based analysis utilizing growth media only detects culturable cells. This could possibly account for the increased gene copies/µl *E. coli* detected using qPCR analysis (Liang et al., 2015). As previously indicated, viability qPCR could be employed to detect the viable *E. coli* cells in the tank water samples. For example, PMA has previously been employed to successfully monitor for the presence of viable *E. coli* in fresh produce and irrigation water (Elizaquível et al., 2011; Truchado et al., 2015). Future studies could then focus on the utilization of PMA in conjunction with qPCR analysis for the detection of viable *E. coli* and enterococci cells in harvested rainwater. In addition, these assays may allow for the rapid assessment of water quality as results may be obtained within five hours, whereas culturing of *E. coli* from water samples may take up to 24 hours and culturing of enterococci from water samples may take up to three days (Frahm & Obst, 2003).

To assess the effectiveness of the MST and CST markers detected in the current study to supplement indicator organism analysis of harvested rainwater in future screenings, correlations between the indicator organisms and the source tracking markers were also investigated. Significant positive correlations were then recorded between adenovirus and *E. coli* (enumerated with the culturing techniques) ($r = 0.983$; $p = 0.000$), the HF183 marker and *E. coli* (quantified by qPCR) ($r = 0.303$; $p = 0.023$), Lachnospiraceae and heterotrophic bacteria ($r = 0.682$; $p = 0.000$) and human mtDNA and enterococci (enumerated with the culturing techniques) ($r = 0.297$; $p = 0.026$) detected in the tank water samples. In contrast, no significant correlations between the MST markers and the indicator organisms (qPCR analysis) detected in the rooftop debris samples were observed. The correlation observed for *E. coli* (enumerated

with culturing techniques) versus adenovirus in the tank water samples is in accordance with a previous study where correlations between adenovirus and indicator organisms have been noted (McQuaig et al., 2009). McQuaig et al. (2009) noted strong positive correlations between adenovirus and *E. coli* (r = 1.00; p = 0.034), enterococci (r = 1.00; p = 0.009) and fecal coliforms (r = 1.00; p = 0.034) in river water samples. Adenoviruses are considered important indicators of human health risk as they may cause disease such as eye infections, gastroenteritis, pneumonia, meningitis and hepatitis (Van Heerden et al., 2003; Jones et al., 2007). Furthermore, it has been noted that adenovirus strains are host-specific and may be more resilient to environmental stresses such as temperature and pH fluctuations, UV irradiation and disinfection strategies as compared to bacteria and may therefore be promising indicators of host-specific fecal pollution in water sources (McQuaig et al., 2009; Sidhu et al., 2013). Hewitt et al. (2013) then employed human adenovirus to screen for sewage contamination in urban streams, estuaries and river water samples and found these viruses to be abundant at the point where sewage contamination of the water source was probable. Furthermore, Waso et al. (2016) previously detected adenovirus (including human adenovirus) in 42.5% of the tank water and 52.5% of the rooftop debris samples collected at the Kleinmond Housing scheme site.

A significant positive correlation was also observed for the human-specific HF183 marker and *E. coli* (quantified by qPCR) (r = 0.303; p = 0.023). This observation was in accordance with previous studies where significant correlations between different human-specific *Bacteroides* markers and *E. coli* enumerated with culture based techniques or quantified by qPCR have been observed (Layton et al., 2006; Gourmelon et al., 2010; Kapoor et al., 2013; Nshimyimana et al., 2014, Liang et al., 2015; Staley et al., 2016). Utilizing Cluster Analysis, human mtDNA then significantly correlated with enterococci (enumerated with culturing techniques) (r = 0.30; p = 0.026) (**Table 2.4**), but was not significantly related to the HF183 marker and *E. coli* (qPCR analysis) (**Fig. 2.6.**). This result is in contrast with a study conducted by Kapoor et al. (2013) where correlations between *E. coli*, human-specific *Bacteroides* (BacHum) and various host-specific mtDNA markers, including human mtDNA, were investigated. The study found significant positive correlations between the human mtDNA marker and the human-specific *Bacteroides* marker (r = 0.38; p < 0.0001) and between the human mtDNA marker and *E. coli* (quantified by qPCR) (r = 0.11; p < 0.001). However, based on the correlation observed between human mtDNA and enterococci in the current study, it may still be beneficial to include human mtDNA and other mtDNA markers (porcine and bovine) in a “toolbox” of markers for the screening of harvested rainwater in future studies.

Another notable positive correlation was observed between Lachnospiraceae and heterotrophic bacteria (r = 0.682; p = 0.000) detected in the tank water samples. It should be noted that Lachnospiraceae did not correlate with any of the other MST markers or indicator organisms screened for in the tank water samples. This was also in contrast with previous research

conducted where Lachnospiraceae was positively correlated with the human-specific HF183 marker ($r = 0.97$; $p = 0.01$) and *Enterococcus* spp. (quantified by qPCR) ($r = 0.91$; $p < 0.01$) (Newton et al., 2011). Furthermore, Newton et al. (2011) observed a strong relationship between the presence of Lachnospiraceae and adenovirus in the water source with a 154% increase in the likelihood of detecting adenovirus with every ten-fold increase in Lachnospiraceae concentrations. However, based on the correlation observed between the Lachnospiraceae and heterotrophic bacteria, Lachnospiraceae may still be a valuable MST marker to screen for in DRWH systems. Lachnospiraceae is a robust family of obligate anaerobic bacteria which occur in the gut of humans and other animals (Vos et al., 2009). The Lachnospiraceae may thus be utilized to indicate contamination of a water source with fecal anaerobes and may additionally be utilized as an indicator of the microbial load in the water source as a wide range of fecal anaerobes belong to the Lachnospiraceae family.

Correlations between the CST markers and indicator organisms in the tank water and rooftop debris samples were also investigated to determine which CST markers will be valuable markers to include in future screenings of DRWH systems. For the tank water samples, significant positive correlations were observed for caffeine versus; enterococci (quantified by qPCR) ($r = 0.863$; $p = 0.000$); fecal coliforms ($r = 0.447$; $p = 0.001$); total coliforms ($r = 0.483$; $p = 0.000$) and enterococci (enumerated with culturing techniques) ($r = 0.399$; $p = 0.002$). Salicylic acid then also positively correlated with total coliforms ($r = 0.301$; $p = 0.024$) in the tank water samples (**Appendix Table A4**). In a study conducted by Staley et al. (2016) river water samples in Toronto, Canada were investigated for sewage contamination utilizing CST and MST markers. In addition, the correlation of the CST markers to indicator organisms was also investigated in the river water. The study then found significant positive correlations between culturable *E. coli* versus; caffeine ($r = 0.69$; $p < 0.05$), codeine ($r = 0.36$; $p < 0.05$), cotinine ($r = 0.67$; $p < 0.05$) and acetaminophen ($r = 0.41$; $p < 0.05$). It was concluded that the river water was contaminated with sewage discharge and that CST markers in combination with indicator organisms may provide a more comprehensive assessment of water quality (Staley et al., 2016). In addition, caffeine was highlighted as a promising CST marker to monitor water quality for anthropogenic contamination. Based on the correlations between the indicator organisms and caffeine and salicylic acid observed in the current study, these two CST markers may thus be promising markers to screen DRWH systems for contamination originating from anthropogenic activities.

For the rooftop debris samples, correlation analysis between the CST markers versus *E. coli* and enterococci enumerated with the qPCR assays revealed significant correlations for *E. coli* (quantified by qPCR) versus methylparaben ($r = 0.623$; $p = 0.000$) and *E. coli* (quantified by qPCR) versus salicylic acid ($r = 0.273$; $p = 0.042$) (**Appendix Table A6**). Methylparaben has not previously been employed as a CST marker, but research is being conducted on this compound

to investigate its endocrine disrupting characteristics (Lillo et al., 2016). This compound is a common preservative found in thousands of personal care products, is absorbed after application to the skin and may be detected in the blood and urine of the exposed individual within one hour after the use of the product (Lillo et al., 2016). Therefore, methylparaben may be detected in human excrement, wastewater and household waste (Lillo et al., 2016) in high concentrations (µg/L). The current study thus highlights the potential of methylparaben as an anthropogenic CST marker to be utilized to monitor water sources and to monitor DRWH systems for anthropogenic contamination.

Finally, the correlations between the tank water and the rooftop debris samples for each individual MST, CST marker, *E. coli* (quantified by qPCR) and enterococci (quantified by qPCR) were investigated. Of the MST markers, only HF183 exhibited a significant positive correlation between the tank water and rooftop debris samples ($r = 0.50$; $p = 0.000$). The HF183 marker could thus be screened for in tank water samples and may indicate that debris washing into the tank is contributing to the decline of the quality of the stored rainwater. Similar results were obtained in a study by Bradshaw et al. (2016), during which there were no significant differences ($p < 0.05$) observed between the number of HF183 gene copy numbers detected in river water samples and the HF183 gene copy numbers detected in the sediment across three different sites. Bradshaw et al. (2016) then noted that monitoring MST markers in sediment may be just as crucial as monitoring MST levels in a water column, as storm events could lead to the re-introduction of MST markers and potential pathogens into the water source when the sediment is disturbed. None of the indicators (*E. coli* and enterococci quantified with the qPCR assays), the other MST markers or any of the CST markers could be positively correlated between the tank water and rooftop debris samples. However, Lachnospiraceae detected in the tank water samples displayed a significant negative correlation ($r = -0.29$; $p = 0.028$) with the Lachnospiraceae detected in the rooftop debris samples. Limited research on the persistence of Lachnospiraceae in solid matrices versus water sources is available and the persistence mechanism of this group of organisms should thus be investigated in future research. In addition, the overall lack of positive correlations between the tank water and rooftop debris samples for the CST, MST markers and the indicator organisms may be due to differences in persistence of the organisms or markers in the two matrices (water versus debris) which will be influenced by UV (the rooftop debris will be more exposed than the tank water), temperature and pH, amongst other factors.

2.5 Conclusions

In the current study, it was established that the tank water samples analyzed did not adhere to drinking water standards with the indicator organism levels detected generally contravening the recommended national and international drinking water guidelines. *Escherichia coli* and

enterococci were also quantifiable in all the tank water and rooftop debris samples uitlizing qPCR analysis. Thus, based on the levels of indicator organisms detected in the tank water, it is recommended that the tank water not be utilized for consumption. However, the water may be utilized for limited domestic and irrigational activities. In addition, conventional PCR analysis indicated that *Bacteroides* HF183, adenovirus, human mtDNA and Lachnospiraceae were the most readily detected MST markers in the tank water and rooftop debris samples. *Bacteroides* HF183, adenovirus and Lachnospiraceae were then detected in all the tank water and rooftop debris samples, while human mtDNA was detected in 90% and 91.7% of the tank water and rooftop debris samples, respectively, utilizing qPCR analysis. Results obtained in the current study thus corroborated data obtained in previous research were qPCR was generally found to be more sensitive than conventional PCR. Future research should thus employ qPCR analysis to detect and quantify the remaining MST markers (*Bifidobacterium*, *Enterococcus esp*, bovine mtDNA, porcine mtDNA and enterovirus) detected utilizing conventional PCR assays in the current study. Furthermore, the CST markers; caffeine, salicylic acid, acetaminophen, triclosan, triclocarban and methylparaben were all detected at µg/L levels in the tank water and rooftop debris samples. Based on the indicator organism analysis as well as the detection of the prevalent MST and CST markers, it was thus evident that the tank water samples in the Kleinmond Housing Scheme site are contaminated with fecal matter and that anthropogenic activity in the vicinity of the tanks significantly influences the tank water quality.

Based on the correlations observed in the current study between the MST markers and indicator organisms, *Bacteroides* HF183 and human mtDNA may be employed to monitor for the presence of human associated fecal microorganisms and possibly human pathogens in DRWH systems. In addition, Lachnospiraceae may be employed to monitor for the presence of fecal anaerobes and the microbial load (as a wide range of strictly anaerobic organisms are associated with the Lachnospiraceae family) in DRWH systems, while adenovirus may be employed as a marker of fecal contamination and as an indication of the potential health risk associated with utilizing the harvested rainwater. For the correlation analysis of CST markers versus the indicator organisms, salicylic acid, caffeine and methylparaben exhibited the greatest potential as supplementary contamination indicators and may be employed to indicate contamination of the DRWH systems by anthropogenic activities and household waste.

Moreover, based on the wide spectrum of human associated MST markers detected in the current study, it is recommended that wildlife animal MST markers targeting for example bird species and MST markers specific for domesticated animals should be designed and screened for in DRWH tanks. This is crucial as various bird species (including pigeons and chickens) and domestic animals (dogs and cats) were observed at the sampling site throughout the sampling period. Correlations between these markers and standard indicator organisms should then also

be investigated in order to accurately identify all possible sources of fecal contamination that contribute to the decline of the quality of the stored rainwater.

Based on the detection of the CST markers (triclosan, triclocarban, methylparaben, caffeine and acetaminophen) in the tank water, rooftop debris and municipal tap water samples collected in this study, it was also evident that these compounds are widespread in the environment. It would thus be beneficial to elucidate the effect of exposure to these compounds, specifically the compounds with endocrine disrupting characteristics such as triclosan, triclocarban and methylparaben, at µg/L concentrations on human and animal health. Additionally, water treatment methods for the effective removal of these compounds from water sources earmarked for domestic purposes, should be investigated and implemented.

Furthermore, concurrence analysis between the MST and CST markers present in DRWH systems and pathogens indigenous to stored harvested rainwater such as *Pseudomonas* spp., *Klebsiella* spp., and *Legionella* spp. amongst others, should be investigated in order to elucidate which of these markers could be utilized as surrogates for these microbial pathogens. It is also recommended that qPCR analysis should be coupled to nucleic acid binding dyes such as EMA or PMA, for the development of viability qPCR assays to detect only the viable indicator organisms and pathogens in harvested rainwater. Accurately determining the viable population in this water source will then aid in quantitative microbial risk assessment in future studies to determine if there is a health risk associated with the use of harvested rainwater for daily domestic purposes and for consumption.

Aknowledgements

The authors would like to acknowledge the following individuals and institutions for their contributions to this project:

- The financial assistance of the **National Research Foundation (NRF)** (Grant Number: 90320) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NRF.
- The **South African Weather Services** for providing the rainfall data for the sampling period.
- Prof Gideon Wolfaardt, Mr. Edward Archer and Mr. Niel Olivier for their assistance in obtaining standards for triclosan, triclocarban, carbamazepine and methylparaben and setting up the methods for the HPLC/MS/MS analysis.

2.6 References

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.M., 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 (3), 724-732.

Ahmed, W., Gardner, T., Toze, S., 2011. Microbiological quality of roof-harvested rainwater and health risks: A review. Journal of Environmental Quality 40 (1), 13-21.

Ahmed, W., Hamilton, K., Gyawali, P., Toze, S., Haas, C., 2016. Evidence of avian and possum fecal contamination in rainwater tanks as determined by microbial source tracking approaches. Applied and Environmental Microbiology DOI:10.1128/AEM.00892-16.

Ahmed, W., Hodgers, L., Sidhu, J.P., Toze, S., 2012. Fecal indicators and zoonotic pathogens in household drinking water taps fed from rainwater tanks in Southeast Queensland, Australia. Applied and Environmental Microbiology 78 (1), 219-226.

Ahmed, W., Huygens, F., Goonetilleke, A., Gardner, T., 2008b. Real-time PCR detection of pathogenic microorganisms in roof-harvested rainwater in Southeast Queensland, Australia. Applied and Environmental Microbiology 74 (17), 5490-5496.

Ahmed, W., Stewart, J., Powell, D., Gardner, T., 2008a. Evaluation of the host-specificity and prevalence of enterococci surface protein (*esp*) marker in sewage and its application for sourcing human fecal pollution. Journal of Environmental Quality 37 (4), 1583-1588.

Ahmed, W., Vieritz, A., Goonetilleke, A., Gardner, T., 2010a. Health risk from the use of roof-harvested rainwater in Southeast Queensland, Australia, as potable or non-potable water, determined using quantitative microbial risk assessment. Applied and Environmental Microbiology 76 (22), 7382-7391.

Ahmed, W., Yusuf, R., Hasan, I., Goonetilleke, A., Gardner, T., 2010b. Quantitative PCR assay of sewage-associated *Bacteroides* markers to assess sewage pollution in an urban lake in dhaka, bangladesh. Canadian Journal of Microbiology 56 (10), 838-845.

Alshakka, M., Ibrahim, M.I.M., Hassali, M.A., Palaian, S., Aljadhey, H., 2016. Hazards of pharmaceuticals in water as new area in eco-pharmacovigilance research. Journal of Pharmacy Practice and Community Medicine 2 (1), 3-8.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, Z., Miller, W., Lipman, D.J, 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research 25 (17), 3389-3402.

Anderson, K.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 (6), 3041-3048.

Boyd, G.R., Reemtsma, H., Grimm, D.A., Mitra, S., 2003. Pharmaceuticals and personal care products (PPCPs) in surface and treated waters of Louisiana, USA and Ontario, Canada. Science of the Total Environment 311 (1), 135-149.

Boyle, D.G., Boyle, D.B., Olsen, V., Morgan, J.A.T., Hyatt, A., 2004. Rapid quantitative detection of chytridiomycosis (*Batrachochytrium dendrobatidis*) in amphibian samples using real-time Taqman PCR assay. Diseases of Aquatic Organisms 60 (2), 141-148.

Bradshaw, J.K., Snyder, B.J., Oladeinde, A., Spidle, D., Berrang, M.E., Meinersmann, R.J., Oakley, B., Sidle, R.C., Sullivan, K., Molina, M., 2016. Characterizing relationships among fecal indicator bacteria, microbial source tracking markers, and associated waterborne pathogen occurrence in stream water and sediments in a mixed land use watershed. Water Research 101, 498-509.

Brownell, M., Harwood, V.J., Kurz, R., McQuaig, S.M., Lukasik, J., Scott, T., 2007. Confirmation of putative stormwater impact on water quality at a Florida beach by microbial source tracking methods and structure of indicator organism populations. Water Research 41 (16), 3747-3757.

Buerge, I.J., Poiger, T., Müller, M.D., Buser, H., 2003. Caffeine, an anthropogenic marker for wastewater contamination of surface waters. Environmental Science and Technology 37 (4), 691-700.

Byappanahalli, M.N., Fujioka, R., 2004. Indigenous soil bacteria and low moisture may limit but allow fecal bacteria to multiply and become a minor population in tropical soils. Water Science and Technology 50 (1), 27-32.

Byappanahalli, M.N., Whitman, R.L., Shively, D.A., Sadowsky, M.J., Ishii, S., 2006b. Population structure, persistence, and seasonality of autochthonous *Escherichia coli* in temperate, coastal forest soil from a Great Lakes watershed. Environmental Microbiology 8 (3), 504-513.

Byappanahalli, M.N., Whitman, R.L., Shively, D.A., Ting, W.E., Tseng, C.C., Nevers, M.B., 2006a. Seasonal persistence and population characteristics of *Escherichia coli* and enterococci in deep backshore sand of two freshwater beaches. Journal of Water and Health 4 (3), 313-320.

Caldwell, J.M., Raley, M.E., Levine, J.F., 2007. Mitochondrial multiplex real-time PCR as a source tracking method in fecal-contaminated effluents. Environmental Science and Technology 41 (9), 3277-3283.

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. Journal of Applied Microbiology 113 (1), 66-75.

Crabtree, K.D., Ruskin, R.H., Shaw, S.B., Rose, J.B., 1996. The detection of *Cryptosporidium* oocysts and *Giardia* cysts in cistern water in the US Virgin Islands. Water Research 30 (1), 208-216.

Council for Scientific and Industrial Research (CSIR), 2011. Sustainable Implementation, Monitoring and Evaluation. Available: http://www.csir.co.za/publications/pdfs/Sustainable_implementation_monitoring_evaluation.pdf [2016, July 28].

Cox, A.R., 1987. Efficacy of the antimicrobial agent triclosan in topical deodorant products: Recent developments *in vivo*. Journal of the Society for Cosmetic Chemists 38, 223-231.

Despins, C., Farahbakhsh, K., Leidl, C., 2009. Assessment of rainwater quality from rainwater harvesting systems in Ontario, Canada. Journal of Water Supply: Research and Technology - AQUA 58 (2), 117-134.

Dobrowsky, P.H., Mannel, D., De Kwaadsteniet, M., Prozesky, H., Khan, W., Cloete, T., 2014a. Quality assessment and primary uses of harvested rainwater in Kleinmond, South Africa. Water SA 40 (3), 401-406.

Dobrowsky, P.H., De Kwaadsteniet, M., Cloete, T.E., Khan, W., 2014b. Distribution of indigenous bacterial pathogens and potential pathogens associated with roof-harvested rainwater. Applied and Environmental Microbiology 80 (7), 2307-2316.

Dobrowsky, P.H., Van Deventer, A., De Kwaadsteniet, M., Ndlovu, T., Khan, S., Cloete, T.E., Khan, W., 2014c. Prevalence of virulence genes associated with pathogenic *Escherichia coli* strains isolated from domestically harvested rainwater during low- and high-rainfall periods. Applied and Environmental Microbiology 80 (5), 1633-1638.

Department of Water Affairs and Forestry (DWAF), 1996. South African Water Quality Guidelines 2nd Edition, Volume 1: Domestic Water Use. CSIR Environmental Services, Pretoria.

Elizaquível, P., Sánchez, G., Aznar, R., 2011. Quantitative detection of viable foodborne *E. coli* O157: H7, *Listeria monocytogenes* and *Salmonella* in fresh-cut vegetables combining propidium monoazide and real-time PCR. Food Control 25 (2), 704-708.

Field, K.G., Samadpour, M., 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Research 41 (16), 3517-3538.

Frahm, E., Obst, U., 2003. Application of the fluorogenic probe technique (TaqMan PCR) to the detection of *Enterococcus* spp. and *Escherichia coli* in water samples. Journal of Microbiological Methods 52 (1), 123-131.

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. Journal of Applied Microbiology 85 (S1), 83S-89S.

Gikas, G.D., Tsihrintzis, V.A., 2012. Assessment of water quality of first-flush roof runoff and harvested rainwater. Journal of Hydrology 466, 115-126.

Gourmelon, M., Caprais, M., Mieszkin, S., Marti, R., Wery, N., Jardé, E., Derrien, M., Jadas-Hécart, A., Communal, P., Jaffrezic, A., 2010. Development of microbial and chemical MST tools to identify the origin of the fecal pollution in bathing and shellfish harvesting waters in France. Water Research 44 (16), 4812-4824.

Hagedorn, C., Weisberg, S.B., 2009. Chemical-based fecal source tracking methods: current status and guidelines for evaluation. Reviews in Environmental Science and Biotechnology 8 (3), 275-287.

Handia, L., 2005. Comparative study of rainwater quality in urban Zambia. Journal of Water Supply: Research and Technology - AQUA 54 (1), 55-64.

Harwood, V.J., Staley, C., Badgley, B.D., 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), 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 (6), 3163-3170.

Heberer, T., Reddersen, K., Mechlinski, A., 2002. From municipal sewage to drinking water: fate and removal of pharmaceutical residues in the aquatic environment in urban areas. Water Science and Technology 46 (3), 81-88.

Heim, A., Ebnet, C., Harste, G., Pring-Åkerblom, P., 2003. Rapid and quantitative detection of human adenovirus DNA by real-time PCR. Journal of Medical Virology 70 (2), 228-239.

Hewitt, J., Greening, G.E., Leonard, M., Lewis, G.D., 2013. Evaluation of human adenovirus and human polyomavirus as indicators of human sewage contamination in the aquatic environment. Water Research 47 (17), 6750-6761.

Ishii, S., Ksoll, W.B., Hicks, R.E., Sadowsky, M.J., 2006. Presence and growth of naturalized *Escherichia coli* in temperate soils from Lake Superior watersheds. Applied and Environmental Microbiology 72 (1), 612-621.

Jackson, V.A., Paulse, A., Bester, A., Neethling, J., Khan, S., Khan, W., 2009. Bioremediation of metal contamination in the Plankenburg River, Western Cape, South Africa. International Biodeterioration and Biodegradation 63 (5), 559-568.

Jones, M.S., Harrach, B., Ganac, R.D., Gozum, M.M., Dela Cruz, W.P., Riedel, B., Pan, C., Delwart, E.L., Schnurr, D.P, 2007. New adenovirus species found in a patient presenting with gastroenteritis. Journal of Virology 81 (11), 5978-5984.

Jones, R.D., Jampani, H.B., Newman, J.L., Lee, A.S., 2000. Triclosan: A review of effectiveness and safety in health care settings. American Journal of Infection Control 28 (2), 184-196.

Jones-Lepp, T., 2006. Chemical markers of human waste contamination: analysis of urobilin and pharmaceuticals in source waters. Journal of Environmental Monitoring 8 (4), 472-478.

Kabiri, L., Alum, A., Rock, C., McLain, J.E., Abbaszadegan, M., 2016. A toolbox strategy using *Bacteroides* genetic markers to differentiate human from non-human sources of fecal contamination in natural water. Science of the Total Environment 572, 897-905.

Kapoor, V., Smith, C., Santo Domingo, J.W., Lu, T., Wendell, D., 2013. Correlative assessment of fecal indicators using human mitochondrial DNA as a direct marker. Environmental Science and Technology 47 (18), 10485-10493.

Kirs, M., Caffaro-Filho, R.A., Wong, M., Harwood, V.J., Moravcik, P., Fujioka, R.S., 2016. Human-associated *Bacteroides* spp. and human polyomaviruses as microbial source tracking markers in Hawaii. Applied and Environmental Microbiology 82 (22), 6757-6767.

Kobayashi, A., Sano, D., Hatori, J., Ishii, S., Okabe, S., 2013. Chicken-and duck-associated *Bacteroides–Prevotella* genetic markers for detecting fecal contamination in environmental water. Applied Microbiology and Biotechnology 97 (16), 7427-7437.

Korajkic, A., Brownell, M., Harwood, V.J., 2011. Investigation of human sewage pollution and pathogen analysis at Florida Gulf coast beaches. Journal of Applied Microbiology 110 (1), 174-183.

Kuan, M.M., 1997. Detection and rapid differentiation of human enteroviruses following genomic amplification. Journal of Clinical Microbiology 35 (10), 2598-2601.

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 (6), 4214-4224.

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 and Environmental Microbiology 81 (3), 850-860.

Lillo, M., Nichols, C., Perry, C., Runke, S., Krutilina, R., Seagroves, T.N., Miranda-Carboni, G.A., Krum, S.A. 2016. Methylparaben stimulates tumor initiating cells in ER breast cancer models. Journal of Applied Toxicology DOI 10.1002/jat.3374.

Liu, J., Wong, M., 2013. Pharmaceuticals and personal care products (PPCPs): a review on environmental contamination in China. Environment International 59, 208-224.

Loraine, G.A., Pettigrove, M.E., 2006. Seasonal variations in concentrations of pharmaceuticals and personal care products in drinking water and reclaimed wastewater in Southern California. Environmental Science and Technology 40 (3), 687-695.

Martellini, A., Payment, P., Villemur, R., 2005. Use of eukaryotic mitochondrial DNA to differentiate human, bovine, porcine and ovine sources in fecally contaminated surface water. Water Research 39 (4), 541-548.

McQuaig, S.M., Griffith, J., Harwood, V.J., 2012. Association of fecal indicator bacteria with human viruses and microbial source tracking markers at coastal beaches impacted by nonpoint source pollution. Applied and Environmental Microbiology 78 (18), 6423-6432.

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 (11), 3379-3388.

Muscillo, M., Pourshaban, M., Iaconelli, M., Fontana, S., Di Grazia, A., Manzara, S., Fadda, G., Santangelo, R., La Rosa, G., 2008. Detection and quantification of human adenoviruses in surface waters by nested PCR, TaqMan real-time PCR and cell culture assays. Water, Air and Soil Pollution 191 (1), 83-93.

Mwenge Kahinda, J., Taigbenu, A.E., Boroto, J.R., 2007. Domestic rainwater harvesting to improve water supply in rural South Africa. Physics and Chemistry of the Earth Parts A/B/C 32 (15), 1050-1057.

National Health and Medical Research Council (NHMRC), Natural Resource Management Ministerial Council (NRMMC), 2011. Australian Drinking Water Guidelines Paper 6 National Water Quality Management Strategy. National Health and Medical Research Council, National Resource Management Ministerial Council Commonwealth of Australia, Canberra.

Newton, R.J., Vandewalle, J.L., Borchardt, M.A., Gorelick, M.H., McLellan, S.L., 2011. Lachnospiraceae and Bacteroidales alternative fecal indicators reveal chronic human sewage contamination in an urban harbor. Applied and Environmental Microbiology 77 (19), 6972-6981.

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 (22), 7437-7443.

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

Okuma, T.A., Hellberg, R.S., 2014. Identification of meat species in pet foods using a real-time polymerase chain reaction (PCR) assay. Food Control 50, 9-17.

Olapade, O.A., Depas, M.M., Jensen, E.T., McLellan, S.L., 2006. Microbial communities and fecal indicator bacteria associated with *Cladophora* mats on beach sites along Lake Michigan shores. Applied and Environmental Microbiology 72 (3), 1932-1938.

Qiao, T., Yu, Z., Zhang, X., Au, D.W., 2011. Occurrence and fate of pharmaceuticals and personal care products in drinking water in southern China. Journal of Environmental Monitoring 13 (11), 3097-3103.

Rengarajan, K., Cristol, S.M., Mehta, M., Nickerson, J.M., 2002. Quantifying DNA concentrations using fluorometry: a comparison of fluorophores. Molecular Vision 8, 416-421.

Reyneke, B., Dobrowsky, P.H., Ndlovu, T., Khan, S., Khan, W., 2016. EMA-qPCR to monitor the efficiency of a closed-coupled solar pasteurization system in reducing *Legionella* contamination of roof-harvested rainwater. Science of the Total Environment 553, 662-670.

Rodil, R., Quintana, J.B., Concha-Graña, E., López-Mahía, P., Muniategui-Lorenzo, S., Prada-Rodríguez, D., 2011. Emerging pollutants in sewage, surface and drinking water in Galicia (NW Spain). Chemosphere 86 (10), 1040-1049.

Rusiñol, M., Moriarty, E., Lin, S., Bofill-Mas, S., Gilpin, B., 2016. Human-, ovine-, and bovine-specific viral source tracking tools to discriminate between the major fecal sources in agricultural waters. Food and Environmental Virology 8 (1), 34-45.

Saayman, M.J., Tobin, M., Khan, W., Khan, S., 2012. The establishment of a routine monitoring technique for detecting the most prevalent pathogenic viruses in river water, Western Cape, South Africa. Masters Thesis, Cape Peninsula University of Technology, Western Cape, South Africa.

Sauer, E.P., VandeWalle, J.L., Bootsma, M.J., McLellan, S.L., 2011. Detection of the human specific *Bacteroides* genetic marker provides evidence of widespread sewage contamination of stormwater in the urban environment. Water Research 45 (14), 4081-4091.

Schmittgen, T.D., Zakrajsek, B.A., Mills, A.G., Gorn, V., Singer, M.J., Reed, M.W., 2000. Quantitative reverse transcription–polymerase chain reaction to study mRNA decay: comparison of endpoint and real-time methods. Analytical Biochemistry 285 (2), 194-204.

Sercu, B., Van De Werfhorst, Laurie C, Murray, J.L., Holden, P.A., 2011. Terrestrial sources homogenize bacterial water quality during rainfall in two urbanized watersheds in Santa Barbara, CA. Microbial Ecology 62 (3), 574-583.

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

Sidhu, J., 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, 488-496.

Simmons, G., Jury, S., Thornley, C., Harte, D., Mohiuddin, J., Taylor, M., 2008. A Legionnaires' disease outbreak: a water blaster and roof-collected rainwater systems. Water Research 42 (6), 1449-1458.

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 (1), 230-237.

South African Bureau of Standards (SABS), 2005. South African National Standards (SANS) 241: Drinking Water Quality Management Guide for Water Services Authorities. Annexure 1. South African Bureau of Standards ISBN 0-626-17752-9.

Staley, Z.R., Grabuski, J., Sverko, E., Edge, T.A., 2016. Comparison of microbial and chemical source tracking markers to identify fecal contamination sources in the Humber River and associated stormwater outfalls. Applied and Environmental Microbiology 82 (21), 6357-6366.

Strauss, A., Dobrowsky, P.H., Ndlovu, T., Reyneke, B., Khan, W., 2016. Comparative analysis of solar pasteurization versus solar disinfection for the treatment of harvested rainwater. BMC Microbiology. 16 (289), 1-16.

Templar, H.A., Dila, D.K., Bootsma, M.J., Corsi, S.R., McLellan, S.L., 2016. Quantification of human-associated fecal indicators reveal sewage from urban watersheds as a source of pollution to Lake Michigan. Water Research 100, 556-567.

Truchado, P., Gil, M.I., Kostic, T., Allende, A., 2015. Optimization and validation of a PMA qPCR method for *Escherichia coli* quantification in primary production. Food Control 62, 150-156.

Tuma, R.S., Beaudet, M.P., Jin, X., Jones, L.J., Cheung, C., Yue, S., Singer, V.L., 1999. Characterization of SYBR gold nucleic acid gel stain: A dye optimized for use with 300-nm ultraviolet transilluminators. Analytical Biochemistry 268 (2), 278-288.

Ufnar, J., Wang, S.Y., Christiansen, J., Yampara-Iquise, H., Carson, C., Ellender, R., 2006. Detection of the *nif*H gene of *Methanobrevibacter smithii*: a potential tool to identify sewage pollution in recreational waters. Journal of Applied Microbiology 101 (1), 44-52.

United States Environmental Protection Agency (US EPA), 2007. Method 1694: Pharmaceuticals and personal care products in water, soil, sediment, and biosolids by HPLC/MS/MS. US Environmental Protection Agency, Washington, DC, United States of America.

United States Environmental Protection Agency (US EPA), 2008. Analytical methods approved for drinking water compliance monitoring under the total coliform rule. US Environmental Protection Agency, Washington, DC, United States of America.

Valcárcel, Y., Alonso, S.G., Rodríguez-Gil, J., Gil, A., Catalá, M., 2011. Detection of pharmaceutically active compounds in the rivers and tap water of the Madrid region (Spain) and potential ecotoxicological risk. Chemosphere 84 (10), 1336-1348.

Van Heerden, J., Ehlers, M.M., Van Zyl, W.B., Grabow, W.O. 2003. Incidence of adenoviruses in raw and treated water. Water Research 37 (15), 3704-3708.

Verrinder, G., Keleher, H., 2001. Domestic drinking water in rural areas: are water tanks on farms a health hazard?. Environmental Health 1 (3), 51.

Villemur, R., Imbeau, M., Vuong, M.N., Masson, L., Payment, P., 2015. An environmental survey of surface waters using mitochondrial DNA from human, bovine and porcine origin as fecal source tracking markers. Water Research 69, 143-153.

Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K., Whitman, W.B., 2009. Bergey's manual of systematic bacteriology: Volume 3: The firmicutes.

Wang, J., Wang, S., 2016. Removal of pharmaceuticals and personal care products (PPCPs) from wastewater: A review. Journal of Environmental Management 182, 620-640.

Waso, M., Ndlovu, T., Dobrowsky, P.H., Khan, S., Khan, W., 2016. Presence of microbial and chemical source tracking markers in roof-harvested rainwater and catchment systems for the detection of fecal contamination. Environmental Science and Pollution Research 23 (17), 16987-17001.

Whitman, R.L., Byers, S.E., Shively, D.A., Ferguson, D.M., Byappanahalli, M., 2005. Occurrence and growth characteristics of *Escherichia coli* and enterococci within the accumulated fluid of the northern pitcher plant (*Sarracenia purpurea L*.). Canadian Journal of Microbiology 51 (12), 1027-1037.

Whitman, R.L., Shively, D.A., Pawlik, H., Nevers, M.B., Byappanahalli, M.N., 2003. Occurrence of *Escherichia coli* and enterococci in *Cladophora* (Chlorophyta) in nearshore water and beach sand of Lake Michigan. Applied and Environmental Microbiology 69 (8), 4714-4719.

Wilson, J.L., Schumacher, J.G., Burken, J.G., 2016. Persistence and microbial source tracking of *Escherichia coli* at a swimming beach at lake of the Ozarks State Park, Missouri. JAWRA Journal of the American Water Resources Association 58 (2), 508-522.

Wolf, S., Hewitt, J., Greening, G.E., 2010. Viral multiplex quantitative PCR assays for tracking sources of fecal contamination. Applied and Environmental Microbiology 76 (5), 1388-1394.

World Health Organization (WHO), 2011. Guidelines for drinking-water quality, Fourth edition. World Health Organization, Geneva, Switzerland. ISBN 978 92 4 154815 1.

Yalavarthy, D.P., Kumar, K.R., Rohini, P., Md, A.R., 2015. A review on occurrence, fate and toxicity of triclosan. World Journal of Pharmacy and Pharmaceutical Sciences 4 (7), 336-369.

Zhang, H., Du, Z., Ji, Y., Mei, M., 2013. Simultaneous trace determination of acidic non-steroidal anti-inflammatory drugs in purified water, municipal tap water, juice, soda and energy drink by hollow fiber-based liquid-phase microextraction and ultra-high pressure liquid chromatography coupled to tandem mass spectrometry. Talanta 109, 177-184.

Chapter 3:

(Chapter 3 is compiled in the format of the Science of the Total Environment journal and US spelling is employed)

Development and Small-Scale Validation of Novel Avian-Associated Mitochondrial DNA Source Tracking Markers for the Detection of Avian Fecal Contamination in Rainwater Catchment Systems

Waso, M.[1], Khan, S.[2] and Khan, W.[1*]

[1] *Department of Microbiology, Faculty of Science, Stellenbosch University, Private Bag X1, Stellenbosch, 7602, South Africa.*

[2] *Faculty of Health and Applied Sciences, Namibia University of Science and Technology, 13 Storch Street, Private Bag 13388, Windhoek, 9000, Namibia.*

Abstract

Avian fecal matter may negatively influence the quality of environmental waters and by extension harvested rainwater, as it harbors high concentrations of fecal indicator bacteria (FIB) and may also contain pathogens. The current study was aimed at designing and validating (on a small-scale) novel avian mitochondrial DNA (mtDNA) source tracking markers for the detection of avian fecal matter in rainwater catchment systems. Three primer sets were designed to target the NADH dehydrogenase subunit 5 gene of avian mtDNA. The three primer sets [AVF1 and AVR (designated AV1); AVF2 and AVR (designated AV2); and ND5F and ND5R (designated ND5)] used as markers were then validated by screening 38 non-host (non-avian) and 37 host (avian) fecal samples. The host-sensitivity of the assays were determined; AV1 displayed a host-sensitivity of 1.00, AV2 displayed a host-sensitivity of 0.892 and ND5 displayed a host-sensitivity of 0.622. The host-specificity of each assay was however, equal to 0.316, 0.0526 and 0.237 for AV1, AV2 and ND5, respectively. Tank water samples and rooftop debris were then screened for the prevalence of the three markers. Overall the AV1 marker was detected in 85% and 90% of the tank water and rooftop debris samples, respectively. The AV2 marker was detected in 50% and 28% and the ND5 marker was detected in 73% and 42% of the tank water and rooftop debris samples, respectively. Bayes' theorem was then applied to calculate the conditional probability of each marker detecting true avian contamination in the tank water and rooftop debris. For the AV1 marker there was an 89.2% and a 92.9% probability that the marker detected true avian fecal contamination in the tank water and rooftop debris samples, respectively. The probability values for the AV2 and ND5 markers were lower at 48.5% and 69.1% for the tank water samples and 27.1% and 36.8% for the rooftop debris samples, respectively. The AV1 marker thus exhibits the greatest potential as an avian mtDNA marker for the detection of avian fecal contamination in rainwater harvesting systems. However, based on the low host-specificity of this primer set future research will involve investigating mismatch amplification mutation assays to develop host-specific primers for the detection of avian mtDNA in environmental samples.

Keywords: rainwater harvesting; microbial source tracking; avian mitochondrial DNA; avian fecal pollution; Bayes' theorem

3.1 Introduction

Domestic rainwater harvesting (DRWH) refers to the collection and storage of fresh rainwater and has been earmarked as an additional water source to increase water supplies directly to households (Gould & Nissen-Peterson, 1999; Mwenge Kahinda & Taigbenu, 2011). Rainwater is harvested worldwide and in some countries, such as Australia, harvested rainwater is frequently utilized as the primary fresh water source for households in regions where water is scarce (Ahmed et al., 2010). In South Africa, it is estimated that approximately 69 746 DRWH tanks capture rainwater as a primary source of fresh water, and these tanks are primarily utilized in rural areas in the Eastern Cape and Kwazulu-Natal (Malema et al., 2016). Furthermore, rainwater harvesting tanks are being implemented in all nine provinces of South Africa, in order to decrease the number of households without adequate water supplies (Mwenge Kahinda et al., 2010; Mwenge Kahinda & Taigbenu, 2011).

Despite the benefits of harvesting rainwater, numerous studies have indicated that the quality of harvested rainwater is a major concern (Ahmed et al., 2008; 2010; 2011; 2012a; De Kwaadsteniet et al., 2013; Dobrowsky et al., 2014a; 2014b; 2014c), especially in regions where DRWH tanks are utilized as the primary source of fresh water and the water is subsequently utilized for potable purposes. The chemical quality of harvested rainwater is generally within the respective drinking water guidelines, however a number of pathogenic microorganisms such as virulent *Escherichia coli* (*E. coli*) strains, *Cryptosporidium* spp., *Legionella* spp. and *Salmonella* spp., amongst others have been associated with harvested rainwater, indicating that the microbiological quality of harvested rainwater is compromised (Ahmed et al., 2008; 2010; 2011 2012a; De Kwaadsteniet et al., 2013; Dobrowsky et al., 2014a; 2014b; 2014c). In addition, indicator organisms which include total coliforms, fecal coliforms, *E. coli* and enterococci, have been detected in stored rainwater (Crabtree et al., 1996; Verrinder & Keleher, 2001; Handia, 2005; Ahmed et al., 2008; Despins et al., 2009; Ahmed et al., 2010; 2011; 2012a; Dobrowsky et al., 2014c).

Avian species have been identified as one of the primary sources of fecal pollution of DRWH tanks as they can easily access rooftops utilized for the collection of rainwater (Mwenge Kahinda et al., 2007; De Kwaadsteniet et al., 2013). Avian fecal matter has also been reported to harbor high concentrations of fecal indicator bacteria (FIB) such as *E. coli* and enterococci (Alderisio & DeLuca, 1999; Fogarty et al., 2003; Ge et al., 2010). In addition, virulent *E. coli* (Wallace et al., 1997; Ahmed et al., 2012b), *Salmonella* spp. (Fallacara et al., 2004; Kinzelman et al., 2008), *Campylobacter* spp. (Fallacara et al., 2004; Kinzelman et al., 2008), *Giardia lamblia* (Kuhn et al., 2002), *Cryptosporidium parvum* (Kuhn et al., 2002) and antibiotic resistant genes (Middleton & Ambrose, 2005; Čížek et al., 2007; Simões et al., 2010; Chidamba & Korsten, 2015), have all been associated with avian fecal matter. This implies that the

microorganisms that are present in the fecal matter of avian species may be washed into the DRWH tanks during a rain event, thereby contaminating the harvested rainwater. In addition, birds travel great distances in the environment, display various feeding habits and may therefore be exposed to diverse microbial species (Ahmed et al., 2015). Subsequently, avian species may act as important vectors of microorganisms and more importantly may contribute to the spread of pathogens in the environment (Ahmed et al., 2015). Environmental waters contaminated by avian fecal matter thus pose a significant health risk to humans, whether it is through consumption or recreational activities such as swimming.

Microbial source tracking (MST) is the process whereby molecular markers specific to a microorganism associated with a host or markers specific to the selected host's DNA are utilized to screen environmental waters for fecal pollution originating from that specific host (Caldwell et al., 2007; Field & Samadpour, 2007; Harwood et al., 2014; Ahmed et al., 2015). Microbial source tracking may thus be employed to identify fecal contamination by avian species in environmental waters and the development of MST markers specific to bird fecal matter is therefore required. Green et al. (2012) reported on the development of two Quantitative Real-Time Polymerase Chain Reaction (qPCR) assays for the detection of gull, Canada goose, duck and chicken fecal-associated genetic markers. The MST markers were developed based on unique sequences obtained from gull fecal samples after performing microplate subtractive hybridization. These unique sequences then shared similarity with *Fusobacterium* spp. (GFB), *Catellicoccus marimammalium* (*C. marimammalium*) and *Helicobacter* spp. and subsequently three MST markers were developed. Based on conventional PCR results, the markers for *C. marimammalium* (GFC) and *Helicobacter* spp. (GFD) were developed into qPCR assays. The newly developed GFC and GFD markers were then screened for in host and non-host groups, to determine their avian sensitivity and specificity. Both markers (GFC and GFD) displayed high specificity values (0.98 and 1.00), but low sensitivity values (0.17 and 0.57). In a follow up study by Ahmed et al. (2015) the GFD marker was further validated based on the high specificity value reported for this marker by Green et al. (2012). The host-specificity and -sensitivity of the GFD marker was evaluated by screening a wide range of host and non-host fecal samples and subsequently environmental waters from Florida (United States of America) and Brisbane (Australia) were also screened for the GFD marker. The GFD marker again exhibited high avian specificity (0.96) but low sensitivity (0.52). The detection of the GFC and GFD markers in the host fecal samples in both aforementioned studies were however highly variable and the variability observed was attributed to varying feeding habits of birds based on seasonal changes and bird migrations (Green et al., 2012; Ahmed et al., 2015). This variability consequently influenced the sensitivity of the developed assays.

Based on cross-reactivity frequently observed for MST markers targeting microorganisms associated with an animal or human host and the variable detection of these MST markers in

host fecal samples, mitochondrial DNA (mtDNA) has been proposed as an alternative target for the design of MST markers. Mitochondrial DNA is host-specific and is not influenced by geographical changes, diet or seasonality and occur in high copy numbers in eukaryotic cells. Therefore, mtDNA may be amplified in a robust manner comparable to that of a 16S rRNA PCR assay (Caldwell et al., 2007). Caldwell et al. (2007), Caldwell and Levine (2009) and Villemur et al. (2015) subsequently reported on targeting specific hosts' mtDNA for the direct screening of fecal pollution sources in environmental waters. Hence, mtDNA markers specific for the detection of pig, cow, dog, Canadian goose and human fecal contamination were developed. The advantages of this approach are; mtDNA sequences of animals and humans are readily available on databases; mtDNA is present in fecal matter of all animals and humans and is highly species specific and may therefore be utilized to discriminate between different fecal sources by targeting host DNA directly.

Limited literature is however available on MST markers for the detection of fecal contamination by a range of avian species with most studies focusing on markers for a specific avian group: gulls (Lu et al., 2011; Ryu et al., 2012), poultry (Weidhaas et al., 2010), ducks (Sun et al., 2016) and Canadian geese (Fremaux et al., 2010) and only one study has reported on targeting mtDNA for avian fecal contamination originating from Canadian geese (Caldwell & Levine, 2009), which only targets this group of birds and not a wide range of avian species. In addition, no host-specificity and -sensitivity validation is reported for this avian marker (Caldwell & Levine, 2009). The aim of this study was thus to design a novel avian associated mtDNA marker to detect avian fecal contamination in rainwater harvesting systems in Kleinmond, Western Cape. This site was selected as indicator organisms, pathogens, MST markers and chemical source tracking (CST) markers have been detected in harvested rainwater from this site (Dobrowsky et al., 2014a; 2014b; 2014c; Waso et al., 2016). Furthermore, based on results presented by Caldwell et al. (2007) the NADH dehydrogenase subunit 5 gene was targeted for the design of the avian mtDNA MST markers. This study is a proof of concept study and therefore provides data on the usefulness of targeting mtDNA for the design of MST markers specific to a wide range of avian species. The aim was achieved by: i) identifying the dominant avian species in the Kleinmond area and avian species associated with anthropogenic activities, which could contaminate the rainwater harvesting systems, ii) obtaining the mtDNA NADH dehydrogenase subunit 5 sequences of the identified avian species from the GenBank database, iii) designing primers targeting the NADH dehydrogenase subunit 5 gene and optimizing the conventional PCR assays, iv) validating the newly designed primers on a small scale by screening a variety of host and non-host fecal samples, v) screening rooftop debris and harvested rainwater from the Kleinmond Housing Scheme site for the presence of the newly developed avian source tracking markers and vi) applying Bayesian statistics to determine the

conditional probability that the avian markers detected in the harvested rainwater and rooftop debris was as a result of true avian fecal pollution.

3.2 Materials and Methods

3.2.1 Primer Design and PCR Assay Optimization

Primers for the detection of avian fecal matter were designed to target indigenous and related avian species of the Kleinmond area, Western Cape, and avian species associated with anthropogenic activities, which may contribute to the fecal contamination of rainwater harvesting systems in this area. Primers were designed for the amplification of the mitochondrial gene: NADH dehydrogenase subunit 5. The NADH dehydrogenase subunit 5 nucleotide sequences for: chicken spp., seagull spp., dove spp., pigeon spp., falcon spp., guinea fowl spp. and penguin spp. were retrieved from GenBank (http://www3.ncbi.nlm.nih.gov) (**Table 3.1**).

Table 3.1 Avian species and the accession numbers of the nucleotide sequences obtained from GenBank, utilized to design avian mtDNA source tracking markers.

Avian Type	Avian Species	Accession Numbers
Doves and Pigeons	*Columba livia*	GU908131
	Geotrygon violacea	HM640213.1
	Hemiphaga novaeseelandiae	EU725864
	Leptotila verreauxi	HM640214
	Streptopelia chinensis	KP273832.1
	Zenaida auriculata	HM640211
Seagulls	*Chroicocephalus ridibundus*	KM577662.1
	Ichthyaetus relictus	KC760146.1
	Larus brunnicephalus	JX155863.1
	Larus crassirostris	KM507782.1
	Larus dominicanus	AY293619.1
	Saundersilarus saundersi	JQ071443.1
Penguin	*Spheniscus demersus*	KC914350.1
Falcon	*Falco peregrinus*	NC_000878
Chicken	*Gallus gallus gallus*	NC_007236.1
	Gallus gallus bankiva	NC_007237
	Gallus gallus spadiceus	NC_007235.1
	Gallus lafayetii	NC_007239.1
	Gallus sonneratii	NC_007240.1
	Gallus varius	NC_007238.1
Guinea Fowl	*Numida meleagris*	NC_006382.1

The sequences were aligned using the ClustalW software (version 2.0.10) (Larkin et al., 2007). Primers (**Table 3.2**) were then designed based on these multiple sequence alignments, targeting the same conserved regions in the NADH dehydrogenase subunit 5 gene as described by Caldwell et al. (2007). The first two degenerate primer sets (AV1 and AV2) shared the same reverse primer (AVR) with different forward primers: either AVF1 (primer set AV1) or AVF2 (primer set AV2) (**Table 3.2**). The third primer set (ND5) was designed to increase avian specificity and therefore did not include degenerate base pairs (**Table 3.2**).

Table 3.2 Avian mtDNA primers designed to target the NADH dehydrogenase subunit 5 gene.

Primer	Nucleotide Sequence (5'-3')	Tm (°C)	Location in Target (bp)	Amplicon Size
AVF1	GCC AAC ACA GCY GCM CTC CAG GC	67.3	497-518	190
AVF2	GCC AAY ACA GCY GCM CTM CAR GC	64.0	497-518	
AVR	GCC RAA TTG RGC DGA TTT TCC TG	57.8	666-687	
ND5 Fwd	CCA ATA CAG CCG CCC TCC AAG	66.5	497-517	201
ND5 Rev	GCC GAA TTG GGC TGA TTT TCC TG	64.6	676-698	

The conventional PCR assays were optimized in terms of the annealing temperature and $MgCl_2$ concentration. Optimization of the PCR assays was performed on DNA extracted from chicken meat in order to optimize the primers on DNA from a known avian species. Thirty grams (30 g) of fresh chicken meat (obtained from a local supermarket) was homogenized in 60 mL sterile water using a Seward Stomacher® 400 at 230 rpm for 1 min (Okuma & Hellburg, 2014). The DNA was extracted from the homogenized meat with the QIAamp® DNA Mini Kit (Qiagen, Hilden, Germany) as per manufacturer's instructions (Okuma & Hellburg, 2014).

The resulting optimized conventional PCR assays were performed in a total volume of 25 µL and consisted of 1X Green GoTaq® Flexi Buffer (Promega Corp, Madison, USA), 4 mM $MgCl_2$, 0.2 mM of each dNTP, 0.3 µM of each primer and 1U GoTaq® DNA Polymerase (Promega Corp, Madison, USA). For each assay 2.5 µL of DNA was utilized as PCR template, sterile milliQ was included as a negative control and gull fecal DNA served as a positive control. The cycling parameters consisted of an initial denaturation at 94 °C for 2 min, 40 cycles: 94 °C for 10 s, 57 °C (AV primer sets) or 60 °C (ND5 primer set) for 10 s, 72 °C for 30 s and a final extension at 72 °C for 5 min.

The PCR products were electrophoresed on an agarose gel (1.5%) stained with ethidium bromide (0.5 µg/mL) in 1X tris borate ethylenediaminetetraacetic acid (TBE) buffer, for 1 h 20 min at 80 V. The products were then visualized using the Vilber Lourmat gel documentation system (Vilber Lourmat, Collégien, France) to confirm the presence of the desired amplicon.

Selected PCR products were cleaned and concentrated using the Zymo Research DNA Clean & Concentrator 5 kit™ (Zymo Research, Irvine, USA) as per manufacturer's instruction. The purified PCR products were sequenced at the Central Analytical Facility (CAF) at Stellenbosch University utilizing the BigDye Terminator v3.1 Sequencing Kit (Applied Biosystems, Foster City, USA). The sequencing data was analyzed using the online Basic Local Alignment Search Tool (BLAST) program of the National Centre for Biotechnology Information (NCBI) (Altschul et al., 1997). The DNA sequences that showed > 97% similarity (< 3% diversity) to organisms on the database were recorded.

3.2.2 Avian and Non-Avian Fecal Sample Collection and DNA Extractions

To validate and assess the performance of the newly developed AV and ND5 primer pairs, fecal samples from host and non-host sources were collected. Host (avian species) fecal samples included: pigeon (n = 9), seagull (n = 4), goose (n = 11), chicken (n = 11), macaw (n = 2), and hadeda (ibis) (n = 1); and non-host (non-avian species) fecal samples included: human (n = 8), dog (n = 6), cat (n = 2), pig (n = 15) and cow (n = 6). In addition, wastewater was included as a non-host sample. Pigeon, goose, macaw, hadeda, dog, cat and two pig fecal samples were collected from private residences in Stellenbosch and Cape Town (Northern Suburbs). Seagull fecal samples were collected from the Gordon's Bay Harbor (GPS co-ordinates: 34°09'07.2"S, 18°51'24.8"E). Eleven pig fecal samples were collected from a local piggery (GPS co-ordinates: 34°01'11.0"S, 18°49'12.5"E), chicken fecal samples were collected from the Mariendahl Experimental Farm (GPS co-ordinates: 33°49'53.6"S, 18°47'51.0"E) and cow fecal samples were collected from the Welgevallen Experimental farm (GPS co-ordinates: 33°56'40.3"S, 18°51'41.4"E). Human fecal samples were obtained from a local pathologist and wastewater was collected in a sterile Schott bottle from the influent point of the Stellenbosch Wastewater Treatment Plant (GPS co-ordinates: 33°59'21.13"S, 18°47'47.75"E).

All the animal fecal samples were collected in sterile falcon tubes with sterilized spatulas. DNA was then extracted from all the fecal samples utilizing the QIAamp® DNA Stool Mini Kit (Qiagen, Hilden, Germany) as per manufacturer's instructions. The PCR assays were then conducted as outlined in Section 3.2.1.

3.2.3 Harvested Rainwater and Rooftop Debris Sample Collection and DNA Extractions

All harvested rainwater samples (henceforth referred to as tank water) and rooftop debris samples were collected from ten DRWH systems connected to ten houses located in the Kleinmond Housing Scheme site in Kleinmond (Western Cape) (GPS co-ordinates: 34°20.11'81"S 19°00.59'74"E), which is a semi-urban coastal region. Each DRWH tank has a capacity of 2000 L and no first flush diverters are installed which eliminate debris from washing into the tanks from the roof surface after the first rainfall.

Sampling was conducted one to four days after a rain event from August 2015 to June 2016, during low and high rainfall periods. The total monthly rainfall data for Kleinmond was obtained from the South African Weather Services (Pretoria, South Africa). In total, six sampling sessions were conducted and while ten tank water and ten rooftop debris samples were collected per sampling occasion, a total of 60 tank water and 60 rooftop debris samples were collected overall. Tank water samples were collected in 5 L sterile polypropylene bottles which were rinsed with distilled water, sterilized with 70% ethanol and dried overnight in an oven at 60 °C. Rooftop debris samples (~20 g) were collected in sterile falcon tubes using sterilized spatulas, from the gutter system connected to each sampled rainwater tank. After collection, all samples were stored at 4 °C until further processing.

One liter (1 L) tank water aliquot, collected from each 5 L tank water sample, was concentrated as previously described by Waso et al. (2016) and was subsequently utilized for DNA extraction with the Zymo Research Soil Microbe DNA MiniPrep™ kit (Zymo Research, Irvine, USA) as per manufacturer's instructions.

For the DNA extractions from the rooftop debris samples, 10 g of each debris sample was sonicated for 10 min in 5 mL sterile distilled water in a Branson 5510 sonication bath (Bransonic® Ultrasonic Cleaner) (Jackson et al., 2009). After sonication, 2 mL of the supernatant was transferred into a sterile Eppendorf tube and then centrifuged at 15 000 x *g* for 15 min. After centrifugation, the supernatant was discarded. This procedure was repeated three times. The resulting pellets were then suspended in 200 µL tris ethylenediaminetetraacetic acid (TE) buffer (pH 8). The DNA was extracted from the resuspended pellets utilizing the Zymo Research Soil Microbe DNA MiniPrep™ kit as per manufacturer's instructions.

The PCR assays were then conducted on the DNA extracted from the tank water and rooftop debris as outlined in Section 3.2.1.

3.2.4 Statistical Analysis

Statistical analysis was conducted using Microsoft Office 2016 Excel software to determine the performance characteristics of the newly designed avian mitochondrial MST markers. The host-specificity and host-sensitivity of each marker was calculated as follows: host-sensitivity = TP/(TP + FN) and host-specificity = TN/(TN + FP) where TP is the number of true positive samples, FN is the number of false negative samples, TN is the number of true negative samples and FP is the number of false positive samples. Fecal samples were considered positive for an avian mitochondrial marker if a PCR product could be visualized on an agarose gel and negative if no PCR product was observed.

The positive predictive value (PPV) and negative predictive value (NPV) of each primer set was calculated as follows: PPV = TP/(TP + FP) and NPV = TN/(TN + FN) where TP is the number of true positive samples, FP is the number of false positive samples, TN is the number of true negative samples and FN is the number of false negative samples. The PPV describes the ratio of true positive samples that tested positive as compared to the total number of samples that tested positive for each respective marker; whereas the NPV describes the ratio of true negative samples that tested negative as compared to the total number of samples that tested negative for each respective marker.

Bayes' theorem was then applied to determine the conditional probability that the detection of the AV1, AV2 and ND5 markers in tank water and rooftop debris samples originated from avian feces (thus true avian fecal pollution) as opposed to fecal contamination from non-host groups. Bayes' theorem was applied as described by Kildare et al. (2007) with the following equation (Eq. 1):

$$P(H\backslash T)=\frac{P(T\backslash H)P(H)}{P(T\backslash H)P(H)+P(T\backslash H')P(H')} \quad (1)$$

Therefore, Eq. 1 estimates the probability [P(H\T)] that there is avian fecal contamination in an analyzed sample if the sample tests positive with the AV1, AV2 or ND5 assays; where P(T\H) is the probability of a positive signal with an assay (AV1, AV2 or ND5) in a fecal sample that is of avian origin, P(T\H') is the probability of a positive signal with an assay in a fecal sample that is of non-avian origin, P(H) is the background probability of detecting the individual avian markers in harvested rainwater or rooftop debris and P(H') is the background probability of the individual avian markers being absent from harvested rainwater or rooftop debris. It is assumed that the background probability [P(H)] may be determined by conducting environmental sampling and P(H') will then be equal to 1 - P(H) (Kildare et al., 2007).

In addition, Spearman Rank Order Correlation analysis was performed with the statistical software package Statistica™ 64 Version 13 (2016) to determine if there were any significant correlations between the detection of the AV1, AV2 and ND5 markers in the tank water samples and the rainfall observed during the sampling period.

3.3 Results

3.3.1 Physico-Chemical Parameters

Overall 60 tank water samples were collected from ten DRWH tanks, while 60 rooftop debris samples were collected from the gutter systems connected to each sampled rainwater storage tank at the Kleinmond Housing Scheme site (Western Cape, South Africa). Samples were collected on six occasions after a rain event and the total rainfall recorded per month was

40.4 mm in August 2015 (sampling one), 50.2 mm in September 2015 (sampling two), 13.4 mm in October 2015 (sampling three), 39.4 mm in March 2016 (sampling four), 26.6 mm in April 2016 (sampling five) and 88.6 mm in June 2016 (sampling six) (South African Weather Services, Pretoria, South Africa). Therefore, for August 2015, September 2015, March 2016 and June 2016 higher rainfall was observed, in comparison to October 2015 and April 2016, where lower rainfall was observed.

3.3.2 Host-Sensitivity and -Specificity of the AV and ND5 Markers

To determine the host-sensitivity (proportion of target host samples identified as positive) of all three primer pairs, undiluted DNA from avian fecal samples were screened for the respective markers. All (100%) individual avian fecal samples (n = 37; pigeon, seagull, chicken, macaw, hadeda and geese) were PCR positive for the AV1 (AVF1 and AVR primers) marker, 33 (89.2%) were PCR positive for the AV2 (AVF2 and AVR primers) marker and 23 (62.2%) were PCR positive for the ND5 marker (**Table 3.3**). Therefore, the host-sensitivity of the AV1 marker was 1.00, for the AV2 marker the sensitivity was calculated as 0.892 and the host-sensitivity of the ND5 marker was calculated as 0.622.

Table 3.3 Number of avian and non-avian fecal samples positive for the AV1, AV2 and ND5 markers

Host groups		**No. of samples tested**	**No. of PCR positive samples for each marker**		
			AVF1	**AVF2**	**ND5**
Non-Avian/Non-Specific	Cat	2	0	2	2
	Dog	6	2	6	6
	Cow	6	2	6	0
	Sewage	1	1	1	0
	Human	8	8	8	8
	Pig	15	13	13	13
	Total	**38**	**26**	**36**	**29**
Avian/Specific	Goose	10	10	8	8
	Pigeon	9	9	8	8
	Seagull	4	4	4	4
	Macaw	2	2	2	2
	Hadeda	1	1	0	1
	Chicken	11	11	11	0
	Total	**37**	**37**	**33**	**23**

To determine the host-specificity (proportion of non-target hosts that produce negative results), non-avian fecal samples were also screened for the three avian MST markers. Among the 38 individual non-avian fecal samples (dog, cat, pig, cow, human, wastewater), 26 (68.4%) were PCR positive for the AV1 marker, 36 (94.7%) were PCR positive for the AV2 marker and 29 (76.3%) were PCR positive for the ND5 marker (**Table 3.3**). Therefore, the host-specificity for

AV1 was equal to 0.316, for AV2 the host-specificity was 0.0526 and the host-specificity for the ND5 marker was 0.237.

Furthermore, the PPV and NPV of each primer set was calculated to determine the ratio of true positive samples as compared to the total number of positive samples and the true negative samples as compared to the total number of negative samples. These values describe the probability of obtaining a positive or negative result with the respective markers. The PPV for the AV1 marker was 58.7%, for the AV2 marker the PPV was 47.8% and for the ND5 marker the PPV was 44.2%. The NPV for the AV1 marker was 100%, for the AV2 marker it was 33.3% and for the ND5 marker it was 39.1%.

3.3.3 Detection of the AV1, AV2 and ND5 Markers in Tank Water and Rooftop Debris Samples

To determine the prevalence of the AV1, AV2 and ND5 markers in harvested rainwater and rooftop debris at the Kleinmond Housing Scheme site and to subsequently calculate the background probability [P(H)] of detecting the avian markers in the environmental samples from this site, DNA was extracted from 60 tank water samples and 60 rooftop debris samples which were collected from ten DRWH systems. These samples were screened for the AV1, AV2 and ND5 markers utilizing PCR as described in Section 3.2.1. For the 60 tank water samples, the AV1 marker was detected in 51 samples (85%), the AV2 marker was detected in 30 samples (50%) and the ND5 marker was detected in 44 samples (73%) (**Fig. 3.1.**).

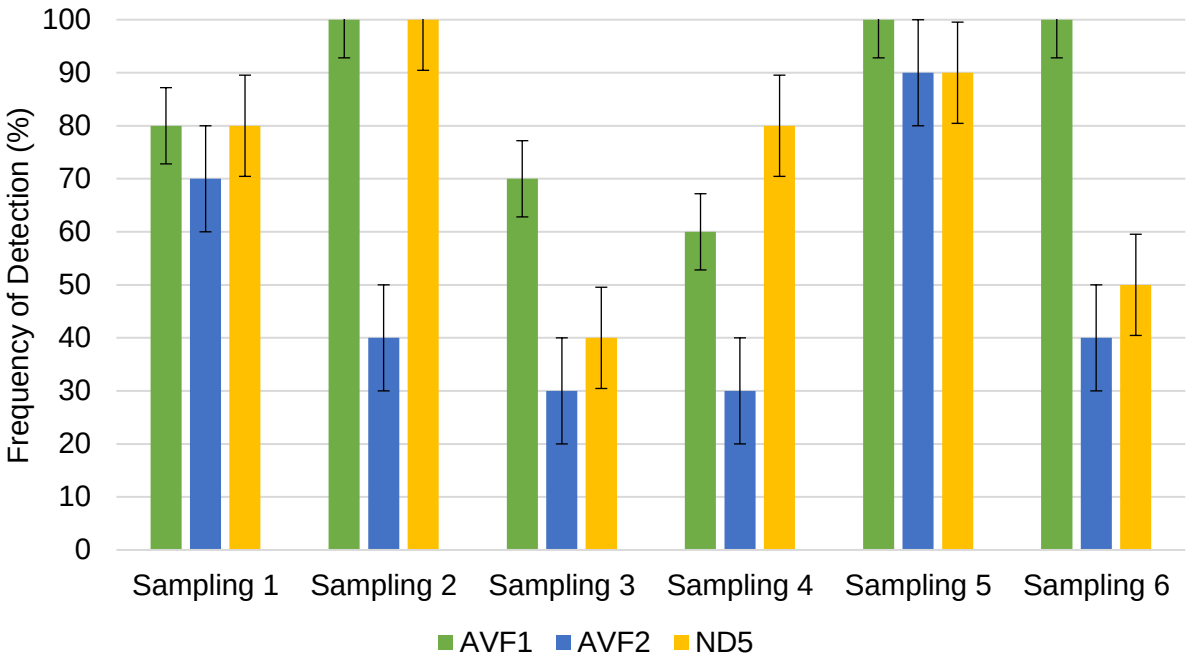


Fig. 3.1. Frequency of detection of the AV1, AV2 and ND5 mtDNA MST markers in tank water samples (n = 10) collected during six sampling sessions with the standard error indicated with error bars.

For the 60 rooftop debris samples, 54 samples (90%) tested positive for the AV1 marker, 17 samples (28%) tested positive for the AV2 marker and 25 samples (42%) tested positive for the ND5 marker (**Fig. 3.2.**).

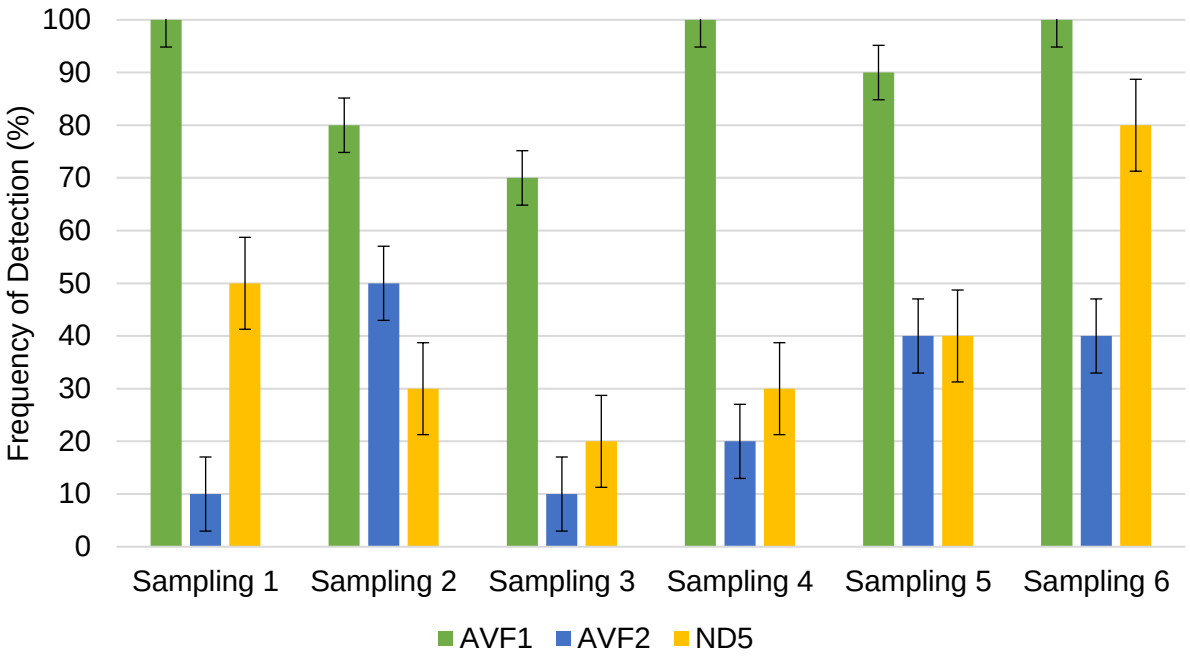


Fig. 3.2. Frequency of detection of the AV1, AV2 and ND5 mtDNA MST markers in rooftop debris samples (n = 10) collected during six sampling sessions with the standard error indicated with error bars.

Representative PCR products obtained from the tank water samples were sequenced in order to confirm the detection of avian mtDNA in the environmental samples. The majority of the sequences from the PCR products obtained with all three markers (AV1, AV2 and ND5) shared sequence similarity with *Columbia livia* (pigeon) mtDNA (KP319029.1) (100% similarity; n = 6). Polymerase chain reaction products obtained from the AV1 and AV2 markers (n = 4) also showed sequence similarity (97% - 100%) to another avian species: *Calidris melanotos* ND5 NADH dehydrogenase subunit 5 gene (JQ251367.1).

3.3.4 Application of Bayes' Theorem to Estimate the Conditional Probability of Accurately Detecting Avian Fecal Contamination in Tank Water and Rooftop Debris Samples

Bayes' theorem was utilized to estimate the conditional probability of accurately detecting avian fecal contamination in tank water and rooftop debris samples collected from the Kleinmond Housing Scheme site with the AV1, AV2 and ND5 markers, respectively, as cross-reactivity with non-host fecal samples was observed for all three markers.

The background probability of detecting the individual avian markers in harvested rainwater or rooftop debris [P(H)] was estimated by screening tank water samples and rooftop debris samples for the individual markers (Kildare et al., 2007). The P(H) value for the AV1 marker was calculated as 0.85 for the rainwater samples and 0.90 for the rooftop debris samples, as the AV1 marker was detected in 85% and 90% of the tank water and debris samples, respectively (**Table 3.4**). The same approach was employed to determine the P(H) values for AV2 and ND5 and the values was thus determined to be 0.50 and 0.28 for AV2 for the tank water and debris samples, respectively and 0.73 and 0.42 for ND5 for the tank water and debris samples, respectively. The P(H') [P(H') = 1-P(H)] values, which describes the probability of the avian markers being absent from the tank water and rooftop debris samples, were then determined and the values were equal to 0.15 and 0.10 for the AV1 marker, 0.50 and 0.72 for the AV2 marker and 0.27 and 0.58 for the ND5 marker, for the tank water and rooftop debris samples, respectively (**Table 3.4**). P(T\H) for each assay was then equal to the host-sensitivity as determined for each marker by screening host and non-host fecal samples (AV1 = 1.00; AV2 = 0.892; ND5 = 0.622) and the P(T\H') for each assay was determined as the probability of a positive signal in a non-avian fecal sample (AV1 = 0.684; AV2 = 0.947; ND5 = 0.763) (**Table 3.4**). For the AV1 marker, the P(H) values of 0.85 and 0.90 for the tank water samples and rooftop debris samples, respectively, corresponded to an 89.2% and 92.9% probability that the detection of the AV1 marker was as a result of true avian fecal pollution of the tank water and rooftop debris, respectively (**Table 3.4**). For the AV2 marker, the P(H) values corresponded to a 48.5% and 27.1% probability that the AV2 marker was detected in the tank water and rooftop debris as a result of avian fecal pollution (**Table 3.4**). Lastly, for the ND5 marker, there was a 69.1% and 36.8% probability that the detection of the marker was as a result of true avian fecal pollution of the tank water and rooftop debris, respectively (**Table 3.4**).

Table 3.4 Sensitivity, specificity, positive and negative predictive values and the conditional probability for the three avian mtDNA markers: AV1, AV2 and ND5 for the detection of true avian fecal contamination in tank water and rooftop debris samples

Marker	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	P(T\H)	P(T\H')	P(H)		P(H')		Conditional Probability[a]	
							Tank Water	Rooftop Debris	Tank Water	Rooftop Debris	Tank Water (%)	Rooftop Debris (%)
AV1	1.00	0.316	0.587	1.00	1.00	0.684	0.850	0.900	0.150	0.100	89.2	92.9
AV2	0.892	0.0526	0.478	0.333	0.892	0.947	0.500	0.283	0.500	0.717	48.5	27.1
ND5	0.622	0.237	0.442	0.391	0.622	0.763	0.733	0.417	0.267	0.583	69.1	36.8

[a] Probability analysis based on Bayes' theorem

3.4 Discussion

Based on previous research, pathogens and indicator organisms have been associated with harvested rainwater in numerous countries such as New Zealand, Nigeria, Australia, US Virgin Islands, Canada and Zambia (De Kwaadsteniet at al., 2013). In addition, it was shown that pathogens, indicator organisms and source tracking markers are present in harvested rainwater collected from the Kleinmond Housing Scheme site in South Africa (Dobrowsky et al., 2014a; 2014b; 2014c; Waso et al., 2016). It is hypothesized that avian species may play a significant role in the contamination of the harvested rainwater and rainwater catchment systems and therefore three novel avian mtDNA MST markers were designed to detect avian fecal contamination in the tank water and rooftop debris samples. To validate the newly designed avian markers, the performance characteristics were evaluated by testing a range of host and non-host fecal samples and the prevalence of the three markers in harvested rainwater and rooftop debris samples was determined.

Host-specificity and -sensitivity are regarded as important performance characteristics when validating newly developed source tracking markers. Markers that are non-specific and are scarce in host fecal matter, may lead to false positive or negative detection of fecal contamination in environmental waters (Stoeckel & Harwood, 2007; Ahmed et al., 2015). The United States Environmental Protection Agency (US EPA) recommends a host-specificity of more than 80% or a value of 0.80 for newly developed MST markers (maximum value of 1.00 or 100%) (US EPA, 2005). Therefore, it is important to assess these performance characteristics to determine if a marker will accurately detect fecal contamination in environmental waters and thus the higher the host-specificity and -sensitivity values, the better the probability of detecting true fecal pollution. In the current study, small-scale validation of the designed markers was conducted as a preliminary determination of the value of targeting mtDNA for the development of MST markers specific for the detection of avian fecal contamination in harvested rainwater. Overall, 38 non-avian fecal samples and 37 avian fecal samples were screened for the designed AV1, AV2 and ND5 markers. The AV1 marker showed the highest host-specificity value of all three markers, with a value of 0.316. However, when compared to the recommended value of 0.80 for host-specificity stipulated by the US EPA (2005), the host-specificity values of the developed markers [AV1 (0.316); AV2 (0.0526); ND5 (0.237)] were well below this guideline. Green et al. (2012) reported avian specificity values of 0.98, 0.99 and 1.00 for the GFC, GFB and GFD avian associated MST markers (targeting bacteria associated with avian fecal matter, not avian mtDNA), respectively. In a follow-up study by Ahmed et al. (2015), host-specificity values for the GFD marker was reported as 0.94 and 1.00 for fecal samples collected in Brisbane and Florida, respectively. These values are higher than the guideline of 0.80 as stipulated by the US EPA (2005) and thus the markers are highly host-specific.

Comparatively, the host-sensitivity of the AV1 marker was 1.00, which implies that the marker was 100% sensitive for the detection of avian fecal contamination and corresponds to the detection of the AV1 marker in all host (avian) fecal samples screened for in this study. The sensitivity values of the AV2 and ND5 markers were lower, with values of 0.892 and 0.622 recorded, respectively. In comparison, Green et al. (2012) reported avian sensitivity values of 0.08, 0.17 and 0.57 for the GFB, GFC and GFD markers, respectively. In addition, Ahmed et al. (2015) reported host-sensitivity (host-prevalence) values of 0.58 and 0.30 for the GFD marker for fecal samples obtained in Brisbane and Florida, respectively. Therefore, the host-sensitivity values obtained in the current study are markedly higher than values reported in literature. This may be due to the fact that the presence of mtDNA in fecal matter is not influenced by the diet of the animal or change in the geographical region due to migrations and seasonal changes. Therefore, the detection of avian mtDNA in avian fecal samples is less variable, and thus markers targeting mtDNA in avian species may lead to the design of more host-sensitive MST markers.

It is, however, important to note that the host-specificity in the current study was notably lower than the recommended guideline of 0.80 stipulated by the US EPA (2005). The low specificity values obtained in the current study may be due to the small pool of host and non-host fecal samples ($n = 75$) screened for and may be more accurately determined by screening a larger number of host and non-host fecal samples ($n > 100$), however no guideline for the number of samples to screen when validating a new MST marker has been stipulated (Harwood et al., 2014). Therefore, it has been suggested that the larger the pool of host and non-host fecal samples screened to validate a new MST marker, the more accurate the marker's sensitivity and specificity can be determined (Harwood et al., 2014). Furthermore, the degenerate bases in the AV1 and AV2 primers could also have influenced the specificity of the markers however, the specificity value (0.237) observed for the non-degenerate primer pair (ND5) was still lower than the value observed for the AV1 marker. The ND5 marker was also designed to target more than one avian species. Thus, based on the fact that the non-degeneracy did not improve the specificity of the primer set, single avian species may need to be targeted with a single primer pair, rather than targeting all avian species with one primer pair. Designing primers that target single avian species may also lead to the generation of more host-specific MST markers. However, a wide variety of avian species may be associated with a specific site and this approach may lead to the exclusion of avian species that may still play a significant role in the contamination of the investigated water source. A possible alternative could then include combining the primers designed in the current study with qPCR and Taqman® chemistries to increase the specificity of the markers as Taqman qPCR has been shown to increase PCR specificity (Boyle et al., 2004).

The relatively high host-sensitivity values obtained in the current study, prompted the investigation of the probability of these markers to then detect avian fecal contamination in environmental waters, even though the markers showed low host-specificity values. All three markers were then screened for in harvested rainwater and rooftop debris samples collected from the Kleinmond Housing Scheme site, where a variety of bird species have been observed perched on the rooftops utilized for the collection of the rainwater (Waso et al., 2016). Overall, the AV1 marker was detected in 85%, the AV2 marker was detected in 50% and the ND5 marker was detected in 73% of the tank water samples. In addition, the detection of the AV1, AV2 and ND5 markers in the tank water samples were compared with the rainfall observed for the sampling period. No significant correlations ($p < 0.05$) were observed between the average rainfall data and the detection of the AV1 ($p = 0.516$), AV2 ($p = 0.177$) and ND5 ($p = 0.232$) markers in the tank water samples. For the screening of the rooftop debris samples, the AV1 marker was detected in 90%, AV2 was detected in 28% and the ND5 marker was detected in 42% of the rooftop debris samples. It was interesting to note that the AV1 marker, which exhibited the highest host-specificity and -sensitivity values of the three designed markers, was also the most prevalent marker detected in the tank water and rooftop debris samples. The AV1 marker was thus highly prevalent and therefore may indicate chronic avian pollution of the rainwater harvesting systems at this site.

In order to determine the probability that the detected markers in the environmental samples, were detected as a result of true avian pollution, Bayes' theorem was applied as stipulated by Kildare et al. (2007). It was therefore assumed that the prevalence of the avian markers in the tank water and rooftop debris samples were indicative of the background avian fecal pollution at this site and were thus equal to the P(H) values for each marker. This corresponded to a probability of 89.2% and 92.9% that the AV1 marker was detected as a result of true avian pollution of the tank water samples and rooftop debris samples, respectively. For the AV2 and ND5 marker the probability percentages were lower at 48.5% and 69.1% for the tank water samples and 27.1% and 36.8% for the rooftop debris samples, respectively. Ahmed et al. (2015) reported a 98% probability that the GFD marker detected in river water was because of true avian fecal pollution. The values reported in the current study for the AV1 mtDNA marker is therefore comparable to the values reported in literature for the GFD avian associated marker. In addition, sequencing of representative PCR products were conducted to confirm that the primer pairs were detecting avian mtDNA in the tank water samples. All of the sequenced products were identified as avian mtDNA. The products obtained with the AV1, AV2 and ND5 markers shared sequence similarity with *Columbia livia* (pigeon) mtDNA (KP319029.1) (n = 6). In addition, the AV1 and AV2 markers also shared sequence similarity with *Calidris melanotos* (Pectoral Sandpiper) NADH dehydrogenase subunit 5 gene (JQ251367.1) (n = 4). The detection of the mtDNA of pigeons at the sampling site is not surprising as pigeons are often associated

with urbanised environments, frequently build their nests in gutters of houses (Chidamba & Korsten, 2015) and has been observed at the sampling site during sample collection. In addition, the presence of Pectoral Sandpiper mtDNA at the sampling site is not unexpected as these bird species are often found in areas with abundant water sources such as rivers, lakes, dams and streams and are often also found in the coastal regions of South Africa, such as Kleinmond (Roberts et al., 2005). These two avian species could therefore be targeted specifically in future studies for the development of specific avian mtDNA markers as these two avian species have now been identified as specific avian fecal contamination sources at the Kleinmond Housing Scheme site.

3.5 Conclusions

This study was a preliminary proof of concept study to indicate the potential use of mtDNA for the design of novel MST markers specifically targeted at detecting avian fecal contamination in harvested rainwater. Based on the conditional probability values obtained for the AV1 marker, this marker exhibited the greatest potential to indicate the presence of true avian fecal contamination in rainwater harvesting systems. However, the AV1 marker still exhibited low avian specificity and therefore it is recommended that future studies focus on increasing the host-specificity of the AV1 marker by combining the primers with Taqman® chemistry and designing a probe. Alternatively, future studies could focus on designing more host-specific avian mtDNA MST markers by targeting single avian species [for example *Columbia livia* (pigeon) and *Calidris melanotos* (Pectoral Sandpiper)] with separate primer pairs, rather than targeting a wide range of avian species with a single primer set. In addition, mismatch amplification mutation assays could be investigated in future studies to design primers specific to avian species by targeting SNP's in the conserved NADH dehydrogenase subunit 5 gene. These SNP's could then be identified between different avian species and different mammalian hosts such as human and pigs, to design primers highly specific to avian species. The primer sets could then be utilized in combination with qPCR chemistry to design a multiplex qPCR assay to detect all avian fecal contamination in DRWH systems.

Despite the low host-specificity of the avian mtDNA markers designed in this study, the sequencing results indicated that all three primer sets detected avian mtDNA in the tank water samples. Thus, avian fecal matter is likely to be a source of microbial contamination of the tanks. In addition, based on the presence of the avian mtDNA in the tank water samples, mtDNA may be utilized to design host-specific MST markers which may be employed in future screening of DRWH systems.

Acknowledgements

The authors would like to acknowledge the following individuals and institutions for their contribution to this project:

- The financial assistance of the **National Research Foundation (NRF)** (Grant Number: 90320) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NRF.
- The **South African Weather Services** for providing the rainfall data for the sampling period.
- All the individuals who assisted in obtaining the fecal samples needed for this study are hereby acknowledged and thanked: Prof W.H. van Zyl, Dr. S. Deane, Mrs. T. Jansen, Mrs. P.H. Stevens, Mr. B. Reyneke, Mr. W.H. van Kerwel, Miss T. Conradie, Mrs. W. Wentzel, Mrs. T. van der Merwe, the Snyman Family, Mr. H. Cronje and the workers at Sweetwell Piggery.

3.6 References

Ahmed, W., Gardner, T., Toze, S., 2011. Microbiological quality of roof-harvested rainwater and health risks: a review. Journal of Environmental Quality 40 (1), 13-21.

Ahmed, W., Harwood, V.J., Nguyen, K., Young, S., Hamilton, K., Toze, S., 2015. Utility of *Helicobacter* spp. associated GFD markers for detecting avian fecal pollution in natural waters of two continents. Water Research 88, 613-622.

Ahmed, W., Hodgers, L., Sidhu, J. P., Toze, S., 2012a. Fecal indicators and zoonotic pathogens in household drinking water taps fed from rainwater tanks in Southeast Queensland, Australia. Applied and Environmental Microbiology 78 (1), 219-226.

Ahmed, W., Huygens, F., Goonetilleke, A., Gardner, T., 2008. Real-time PCR detection of pathogenic microorganisms in roof-harvested rainwater in Southeast Queensland, Australia. Applied and Environmental Microbiology 74 (17), 5490-5496.

Ahmed, W., Sidhu, J.P., Toze, S., 2012b. An attempt to identify the likely sources of *Escherichia coli* harboring toxin genes in rainwater tanks. Environmental Science and Technology 46 (9), 5193-5197.

Ahmed, W., Vieritz, A., Goonetilleke, A., Gardner, T., 2010. Health risk from the use of roof-harvested rainwater in Southeast Queensland, Australia, as potable or non-potable water,

determined using quantitative microbial risk assessment. Applied and Environmental Microbiology 76 (22), 7382-7391.

Alderisio, K.A., DeLuca, N., 1999. Seasonal enumeration of fecal coliform bacteria from the feces of ring-billed gulls (*Larus delawarensis*) and Canada geese (*Branta canadensis*). Applied and Environmental Microbiology 65 (12), 5628-5630.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1997. Basic local alignment search tool. Journal of Molecular Biology 215 (3), 403-410.

Boyle, D.G., Boyle, D.B., Olsen, V., Morgan, J.A.T., Hyatt, A., 2004. Rapid quantitative detection of chytridiomycosis (*Batrachochytrium dendrobatidis*) in amphibian samples using real-time taqman PCR assay. Diseases of Aquatic Organisms 60 (2), 141-148.

Caldwell, J.M., Levine, J.F., 2009. Domestic wastewater influent profiling using mitochondrial real-time PCR for source tracking animal contamination. Journal of Microbiological Methods 77 (1), 17-22.

Caldwell, J.M., Raley, M.E., Levine, J.F., 2007. Mitochondrial multiplex real-time PCR as a source tracking method in fecal-contaminated effluents. Environmental Science and Technology 41 (9), 3277-3283.

Chidamba, L., Korsten, L., 2015. Antibiotic resistance in *Escherichia coli* isolates from roof-harvested rainwater tanks and urban pigeon faeces as the likely source of contamination. Environmental Monitoring and Assessment 187 (7), 1-15.

Čížek, A., Dolejska, M., Karpíšková, R., Dědičová, D., Literák, I., 2007. Wild black-headed gulls (*Larus ridibundus*) as an environmental reservoir of *Salmonella* strains resistant to antimicrobial drugs. European Journal of Wildlife Research 53 (1), 55-60.

Crabtree, K.D., Ruskin, R.H., Shaw, S.B., Rose, J.B., 1996. The detection of *Cryptosporidium* oocysts and *Giardia* cysts in cistern water in the US Virgin Islands. Water Research 30 (1), 208-216.

De Kwaadsteniet, M., Dobrowsky, P.H., Van Deventer, A., Khan, W., Cloete, T., 2013. Domestic rainwater harvesting: microbial and chemical water quality and point-of-use treatment systems. Water, Air and Soil Pollution 224 (7), 1-19.

Despins, C., Farahbakhsh, K., Leidl, C., 2009. Assessment of rainwater quality from rainwater harvesting systems in Ontario, Canada. Journal of Water Supply: Research and Technology - AQUA 58 (2), 117.

Dobrowsky, P.H., De Kwaadsteniet, M., Cloete, T.E., Khan, W., 2014b. Distribution of indigenous bacterial pathogens and potential pathogens associated with roof-harvested rainwater. Applied and Environmental Microbiology 80 (7), 2307-2316.

Dobrowsky, P.H., Mannel, D., De Kwaadsteniet, M., Prozesky, H., Khan, W., Cloete, T., 2014c. Quality assessment and primary uses of harvested rainwater in Kleinmond, South Africa. Water SA 40 (3), 401-406.

Dobrowsky, P.H., Van Deventer, A., De Kwaadsteniet, M., Ndlovu, T., Khan, S., Cloete, T.E., Khan, W., 2014a. Prevalence of virulence genes associated with pathogenic *Escherichia coli* strains isolated from domestically harvested rainwater during low- and high-rainfall periods. Applied and Environmental Microbiology 80 (5), 1633-1638.

Fallacara, D.M., Monahan, C.M., Morishita, T.Y., Bremer, C.A., Wack, R.F., 2004. Survey of parasites and bacterial pathogens from free-living waterfowl in zoological settings. Avian Diseases 48 (4), 759-767.

Field, K.G., Samadpour, M., 2007. Fecal source tracking, the indicator paradigm, and managing water quality. Water Research 41 (16), 3517-3538.

Fogarty, L., Haack, S., Wolcott, M., Whitman, R., 2003. Abundance and characteristics of the recreational water quality indicator bacteria *Escherichia coli* and enterococci in gull feces. Journal of Applied Microbiology 94 (5), 865-878.

Fremaux, B., Boa, T., Yost, C.K., 2010. Quantitative real-time PCR assays for sensitive detection of Canada goose-specific fecal pollution in water sources. Applied and Environmental Microbiology 76 (14), 4886-4889.

Ge, Z., Nevers, M.B., Schwab, D.J., Whitman, R.L., 2010. Coastal loading and transport of *Escherichia coli* at an embayed beach in Lake Michigan. Environmental Science and Technology 44 (17), 6731-6737.

Gould, J., Nissen-Petersen, E., 1999. Rainwater catchment systems for domestic supply. London: Intermediate Technology Publications.

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 (2), 503-510.

Handia, L., 2005. Comparative study of rainwater quality in urban Zambia. Journal of Water Supply: Research and Technology - AQUA 54, 55-64.

Harwood, V.J., Staley, C., Badgley, B.D., 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), 1-40.

Jackson, V.A., Paulse, A., Bester, A., Neethling, J., Khan, S., Khan, W., 2009. Bioremediation of metal contamination in the Plankenburg River, Western Cape, South Africa. International Biodeterioration and Biodegradation 63 (5), 559-568.

Kildare, B.J., Leutenegger, C.M., McSwain, B.S., Bambic, D.G., Rajal, V.B., 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 (16), 3701-3715.

Kinzelman, J., McLellan, S.L., Amick, A., Preedit, J., Scopel, C.O., Olapade, O., Gradus, S., Singh, A., Sedmak, G., 2008. Identification of human enteric pathogens in gull feces at Southwestern Lake Michigan bathing beaches. Canadian Journal of Microbiology 54 (12), 1006-1015.

Kuhn, R.C., Rock, C.M., Oshima, K.H., 2002. Occurrence of *Cryptosporidium* and *Giardia* in wild ducks along the Rio Grande River valley in southern New Mexico. Applied and Environmental Microbiology 68 (1), 161-165.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., Thompson, J.D., Gibson, T.J., Higgins, D.G., 2007. Clustal W and Clustal X version 2.0. Bioinformatics 23 (21), 2947-2948.

Lu, J., Ryu, H., Hill, S., Schoen, M., Ashbolt, N., Edge, T.A., Santo Domingo, J., 2011. Distribution and potential significance of a gull fecal marker in urban coastal and riverine areas of southern Ontario, Canada. Water Research 45 (13), 3960-3968.

Malema, S., Abia, L.K.A., Mwenge Kahinda, J., Ubomba-Jaswa, E., 2016. Gaining a better understanding of the factors that influence the quality of harvested rainwater in South Africa – a review. WISA 2016 Biennial Conference. 16 -19 May 2016. Durban, South Africa.

Middleton, J., Ambrose, A., 2005. Enumeration and antibiotic resistance patterns of fecal indicator organisms isolated from migratory Canada geese (*Branta canadensis*). Journal of Wildlife Diseases 41 (2), 334-341.

Mwenge Kahinda, J., Taigbenu, A.E., 2011. Rainwater harvesting in South Africa: Challenges and opportunities. Physics and Chemistry of the Earth, Parts A/B/C 36 (14), 968-976.

Mwenge Kahinda, J., Taigbenu, A.E., Boroto, J.R., 2007. Domestic rainwater harvesting to improve water supply in rural South Africa. Physics and Chemistry of the Earth, Parts A/B/C 32 (15), 1050-1057.

Mwenge Kahinda, J., Taigbenu, A.E., Boroto, J.R., 2010. Domestic rainwater harvesting as an adaptation measure to climate change in South Africa. Physics and Chemistry of the Earth, Parts A/B/C 35 (13), 742-751.

Okuma, T.A., Hellberg, R.S., 2014. Identification of meat species in pet foods using a real-time polymerase chain reaction (PCR) assay. Food Control 50, 9-17.

Roberts, A., Hockey, P.A.R, Dean, W.R.J., Ryan, P., 2005. Roberts' birds of Southern Africa 7th Edition. Trustees of the J. Voelcker Bird Book Fund, Cape Town, South Africa.

Ryu, H., Griffith, J.F., Khan, I.U., Hill, S., Edge, T.A., Toledo-Hernandez, C., Gonzalez-Nieves, J., Santo Domingo, J., 2012. Comparison of gull feces-specific assays targeting the 16S rRNA genes of *Catellicoccus marimammalium* and *Streptococcus* spp. Applied and Environmental Microbiology 78 (6), 1909-1916.

Simões, R.R., Poirel, L., Da Costa, P.M., Nordmann, P., 2010. Seagulls and beaches as reservoirs for multidrug-resistant *Escherichia coli*. Emerging Infectious Diseases 16 (1), 110-112.

Stoeckel, D.M., Harwood, V.J., 2007. Performance, design, and analysis in microbial source tracking studies. Applied and Environmental Microbiology 73 (8), 2405-2415.

Sun, D., Duan, C., Shang, Y., Ma, Y., Tan, L., Zhai, J., Gao, X., Guo, J., Wang, G., 2016. Application of *Faecalibacterium* 16S rDNA genetic marker for accurate identification of duck faeces. Environmental Science and Pollution Research 23 (8), 7639-7647.

United States Environmental Protection Agency (US EPA), 2005. Microbial source tracking guide document EPA/600/R-05/064. US Environmental Protection Agency, Washington, DC.

Verrinder, G., Keleher, H., 2001. Domestic drinking water in rural areas: are water tanks on farms a health hazard?. Environmental Health 1 (3), 51.

Villemur, R., Imbeau, M., Vuong, M.N., Masson, L., Payment, P., 2015. An environmental survey of surface waters using mitochondrial DNA from human, bovine and porcine origin as fecal source tracking markers. Water Research 69, 143-153.

Wallace, J., Cheasty, T., Jones, K., 1997. Isolation of Verocytotoxin-producing *Escherichia coli* O157 from wild birds. Journal of Applied Microbiology 82 (3), 399-404.

Waso, M., Ndlovu, T., Dobrowsky, P.H., Khan, S., Khan, W., 2016. Presence of microbial and chemical source tracking markers in roof-harvested rainwater and catchment systems for the detection of fecal contamination. Environmental Science and Pollution Research 23 (17), 16987-17001.

Weidhaas, J.L., Macbeth, T.W., Olsen, R.L., Sadowsky, M.J., Norat, D., Harwood, V.J., 2010. Identification of a *Brevibacterium* marker gene specific to poultry litter and development of a quantitative PCR assay. Journal of Applied Microbiology 109 (1), 334-347.

Chapter 4:

General Conclusions and Recommendations

(UK spelling is employed)

General Conclusions and Recommendations

Domestic rainwater harvesting (DRWH) is currently being utilised worldwide as an alternative fresh water source, however numerous studies have indicated that the microbial quality of harvested rainwater does not adhere to drinking water guidelines as indicator organisms and various pathogens have been detected in stored rainwater (Crabtree et al. 1996; Verrinder & Keleher, 2001; Handia, 2005; Field & Samadpour, 2007; Ahmed et al. 2008; Simmons et al. 2008; Despins et al. 2009; Ahmed et al. 2010a; 2011; 2012a). In addition, studies conducted by our research group have indicated that indicator organisms and pathogens have been detected in harvested rainwater in South Africa (Dobrowsky 2014a; 2014b; 2014c). However, while the suitability of a water source for its intended purpose is usually determined by monitoring for indicator organisms, there is growing evidence that indicator organisms may establish populations and persist in water sources (Leclerc et al. 2001; Anderson et al. 2005; Tallon et al. 2005; Field & Samadpour, 2007; Griffith et al. 2009; Harwood et al. 2014). The detection of indicator organisms is therefore not always an indication of a recent contamination event and the risk associated with the use of the water source cannot be accurately assessed (Field & Samadpour, 2007; Harwood et al. 2014). In addition, faecal indicator bacteria (a subgroup of the indicator organisms employed to specifically screen for faecal contamination) are ubiquitously distributed in host species (warm-blooded animals) and the detection of these bacteria in a water source does not provide information regarding the specific sources of faecal contamination. Moreover, a number of studies have indicated that indicator organism analysis generally does not positively correlate with the presence of pathogens in contaminated water sources (Lund, 1996; Bonadonna et al. 2002; Lemarchand & Lebaron, 2003; Anderson et al. 2005; Harwood et al. 2005; 2014).

While it has been suggested that pathogens could be screened for directly, as a measure of the health risk associated with the use of a contaminated water source, a wide range of assays will need to be utilised to screen for all possible pathogens associated with a specific water source, which may become technically complicated and expensive (Field & Samadpour, 2007). In addition, pathogens may be difficult to culture and are often unevenly distributed in the water source, influencing the accurate detection of these disease-causing organisms (Field & Samadpour, 2007; Harwood et al. 2014). Taking all of these pitfalls and challenges associated with indicator organism analysis into consideration, it is clear that supplementary or alternative indicators for monitoring water quality are required. In this regard, microbial (MST) and chemical source tracking (CST) markers have been investigated and are being applied to various water sources in order to accurately identify the source of contamination (Jones-Lepp, 2006; Field & Samadpour, 2007; Muscillo et al. 2008; Ahmed et al. 2010b; Kobayashi et al. 2013; Sidhu et al. 2013; Harwood et al. 2014).

The primary aim of the current study was thus to identify a toolbox of MST and CST markers present in DRWH systems, which may be utilised to augment or supplement indicator organism analysis in future screenings of harvested rainwater (**Chapter two**). Identifying source tracking markers that correlate well with indicator organisms may improve their predictive capability to indicate faecal contamination of a water source (Harwood et al. 2014; Liang et al. 2015). Correlations between the primary source tracking markers detected and traditional indicator organisms were thus also analysed. Subsequently, 60 tank water samples and 60 rooftop debris samples were collected from ten houses located in the Kleinmond Housing Scheme site (Kleinmond, Western Cape, South Africa). The tank water samples were then screened for *Escherichia coli* (*E. coli*), enterococci, faecal coliforms, total coliforms and heterotrophic bacteria utilising traditional culture based methods. In addition, all the tank water and rooftop debris samples were screened for *Bacteroides* HF183, Lachnospiraceae, *Bifidobacterium*, *Methanobrevibacter* spp., *Methanobrevibacter smithii (nif*H gene), *Enterococcus* (*esp* gene), adenovirus, polyomavirus, enterovirus, human, bovine and porcine mitochondrial DNA (mtDNA) utilising conventional PCR assays. The most frequently detected MST markers, *E. coli* and enterococci were then quantified in all the tank water and rooftop debris samples utilising quantitative real-time PCR (qPCR) assays. Furthermore, all the tank water and rooftop debris samples were screened for the CST markers caffeine, salicylic acid, acetaminophen, carbamazepine, triclosan, triclocarban and methylparaben utilising high-performance liquid chromatography tandem mass spectrometry (HPLC/MS/MS) analysis.

The culture based analysis of the indicator organisms in the tank water samples indicated that the heterotrophic bacteria and total coliform counts exceeded the recommended drinking water guidelines [Department of Water Affairs and Forestry (DWAF), 1996; South African Bureau of Standards (SABS), 2005] in all the tank water samples (100%) analysed, *E. coli* counts exceeded the recommended drinking water guidelines in 95% of the tank water samples analysed, while the faecal coliforms exceeded the recommended guidelines in 85% of the tank water samples analysed [DWAF, 1996; SABS, 2005; National Health and Medical Research Council (NHMRC) & Natural Resource Management Ministerial Council (NRMMC), 2011; World Health Organization (WHO), 2011]. Enterococci sporadically exceeded the recommended drinking water guidelines (32%) (NHMRC & NRMMC, 2011). *Escherichia coli*, enterococci and faecal coliforms are utilised to indicate faecal pollution in a water source, while heterotrophic bacteria are utilised as an indicator of the microbial load in a water source and total coliforms are utilised to indicate the general hygienic quality of the water source (DWAF, 1996; SABS, 2005; NHMRC & NRMMC, 2011; WHO, 2011). These results are in accordance with a study by Dobrowsky et al. (2014b) during which, indicator organisms were also detected in the tank water collected from the Kleinmond site. Therefore, based on the indicator organism analysis, it is hypothesised that the tank water is contaminated by faecal matter. Water contaminated by faecal matter may then present a significant health risk to the individuals consuming and using

the water on a regular basis as pathogens such as *Salmonella enterica*, *Listeria monocytogenes*, *Pseudomonas aeruginosa* and enteric viruses, amongst others, may be present in the contaminated water (Liang et al. 2015; Bradshaw et al. 2016).

Based on the conventional PCR results, *Bacteroides* HF183 [tank water, 86.7%; rooftop debris samples, 63.3%] adenovirus [tank water, 66.7%; rooftop debris, 66.7%], Lachnospiraceae [tank water, 55%; rooftop debris, 81.7%] and human mtDNA [tank water, 35%; rooftop debris, 57%] were the MST markers most frequently detected in the tank water and rooftop debris samples. These MST markers were subsequently quantified by qPCR analysis. The HF183 marker, Lachnospiraceae and adenovirus were then quantifiable in all the tank water (100%) and rooftop debris (100%) samples analysed, while the human mtDNA marker was quantifiable in 90% and 91.7% of the tank water and rooftop debris samples, respectively. The HF183 marker, Lachnospiraceae and human mtDNA have previously been applied to water sources such as rivers, harbours and stormwater run-off to monitor for the presence of human faecal matter, faecal anaerobes and wastewater contamination, while adenovirus has been employed to indicate faecal contamination from various human and animal sources in rivers and stormwater run-off (McQuaig et al. 2009; Newton et al. 2011; McQuaig et al. 2012; Sidhu et al. 2013; Harwood et al. 2014; Bradshaw et al. 2016; Rusinõl et al. 2016; Staley et al. 2016). The gene copies detected in the current study for *Bacteroides* HF183 in the tank water (mean concentration of 5.1×10^3 gene copies/µL) and rooftop debris (mean concentration of 4.7×10^3 gene copies/µL) samples were generally lower than copies of this marker detected in surface water in a previous study (10^4 - 10^7 gene copies/µL detected per 100 mL of water) (Ahmed et al. 2010b). In contrast, the mean Lachnospiraceae (tank water, 3.0×10^4 gene copies/µL; rooftop debris, 6.9×10^3 gene copies/µL) and adenovirus (tank water, 3.2×10^2 gene copies/µL; rooftop debris, 6.4×10^3 gene copies/µL) gene copies detected in the current study in the tank water and rooftop debris samples were comparable to the gene copies of Lachnospiraceae previously detected in marine water (Newton et al. 2011) and adenovirus gene copies previously detected in river water samples (Wolf et al. 2010). The human mtDNA gene copies detected in the current study in the tank water (mean concentration of 1.1×10^6 gene copies/µL) and rooftop debris (mean concentration of 3.0×10^5 gene copies/µL) samples were however, generally higher than the concentrations of human mtDNA gene copies previously detected in surface water sources in Canada (mean concentration of 10^3 - 10^4 gene copies/µL per 100 mL of surface water) (Villemur et al. 2015).

The detection of the MST markers in the tank water and rooftop debris samples indicate that the DRWH systems are contaminated by faecal matter, thus corroborating the indicator organism analysis of the tank water samples. It should however be noted that direct contamination of the DRWH tanks by sewage or human faecal matter is unlikely. Therefore, it is hypothesised that the source of the human associated markers may be animals such as birds or rodents that may

easily gain access to the rooftops utilised to harvest the rainwater. Avian species are receiving increased interest as vectors of pathogens in the environment (Ahmed et al. 2015; 2016; Chidamba & Korsten, 2015), while rodents have been associated with the transmission of pathogenic microorganisms to humans via the faecal-oral route and ectoparasites (such as ticks and fleas) (Zheng et al. 2016). Rodents have also recently been associated with the spread of adenovirus strains in the environment (Zheng et al. 2016). These animals are known to live in close proximity to humans and may therefore be exposed to human waste. Subsequently birds and rodents may act as vectors of pathogens and by extension MST markers in the environment. Regardless of the origin of the human associated MST markers detected in the DRWH systems in the current study (direct or indirect via an animal vector), the presence of these markers may indicate the presence of human pathogens in the tank water and rooftop debris. Pathogens such as *Pseudomonas* spp., *Legionella* spp. and *Klebsiella* spp., amongst others, have previously been detected by members of our research group in tank water samples collected from the Kleinmond Housing Scheme site (Dobrowsky et al. 2014a). Thus, based on the fact that a number of pathogenic microorganisms have previously been detected in the tank water at the Kleinmond sampling site (Dobrowsky et al. 2014a), correlations between these organisms and source tracking markers should be investigated. Bradshaw et al. (2016) applied the MST markers for general, ruminant and human *Bacteroides* to river water samples. Correlations between these markers and *Listeria*, *Campylobacter*, *Salmonella* and the specific virulence gene encoding for the Shiga toxin (stx_2) were then investigated. The results indicated that the ruminant MST marker correlated with the presence of the Shiga toxin gene and *Campylobacter*, while the human MST marker correlated with the presence of *Listeria*. The presence of the Shiga toxin gene and *Campylobacter* could subsequently be attributed to agricultural activities along the river where cattle pastures were observed and the presence *Listeria* could be attributed to sewage contamination of the river.

In addition, *E. coli* and enterococci were quantifiable in all the tank water (100%) and rooftop debris (100%) samples by qPCR analysis. The qPCR assays detected the indicator organisms in all the tank water samples, whereas with the culture based methods, *E. coli* was only detected in 57 (95%) of the tank water samples and enterococci was only detected in 19 (32%) tank water samples. The qPCR assays may thus be valuable tools to be employed for the monitoring of water quality as they may be more sensitive to detect indicator organisms in water samples as compared to traditional culture based detection. In addition, based on the results obtained in the current study, the qPCR technique was also more sensitive than conventional PCR for the detection of the MST markers in the harvested rainwater and rooftop debris samples. However, it should be noted that while qPCR may be more sensitive, molecular techniques such as qPCR may detect DNA from viable cells, viable but non-culturable (VBNC) cells and DNA from dead cells in a sample, while culture based analysis only detects viable and culturable cells. It may thus be beneficial to combine qPCR with nucleic acid binding dyes such

as propidium monoazide (PMA) and ethidium monoazide bromide (EMA) for the development of viability qPCR assays in order to detect only the viable cells in a sample. Viability qPCR has successfully been utilised by our research group to estimate the proportion of viable *Legionella* spp. and *Pseudomonas* spp. in pasteurized and unpasteurized harvested rainwater (Reyneke et al. 2016; Strauss et al. 2016). Determining the viable portion of indicator organisms and pathogens in a water source may be utilised to determine the health risks associated with the use of the water source, with increased accuracy.

Based on the HPLC/MS/MS analysis results, caffeine, acetaminophen, triclosan, triclocarban, methylparaben and salicylic acid were detected in all the tank water (except salicylic acid which was detected in 98% of the tank water samples) and rooftop debris samples at µg/L concentrations. The detection of caffeine, salicylic acid and acetaminophen at µg/L concentrations in the tank water and rooftop debris samples are in accordance with a previous study conducted by our research group where these compounds were detected at similar concentrations in the tank water and rooftop debris collected at the Kleinmond Housing scheme site (Waso et al. 2016). In addition, caffeine, acetaminophen and triclosan have also been detected at µg/L concentrations in river water samples from ten sites across America (Glassmeyer et al. 2005). While, methylparaben was detected at µg/L concentrations in river water samples in China (Peng et al. 2008).

Caffeine, acetaminophen, triclosan, triclocarban, methylparaben and salicylic acid are widely associated with anthropogenic activities, household waste and sewage. During the sampling sessions it was observed that various household waste and garbage bags were placed around and on top of the DRWH tanks at the sampling site in Kleinmond. Researchers have previously noted that anthropogenic activities and waste in close proximity to DRWH systems may significantly influence the quality of the harvested rainwater as was observed in the current study (Mwenge Kahinda et al. 2007). In addition, these CST markers were detected in all of the municipal tap water samples (n = 10) collected intermittently during the sampling period, at µg/L concentrations. These results thus indicate that these anthropogenic associated CST markers are ubiquitously distributed in the environment and that the residents of the Kleinmond Housing Scheme site are exposed to these compounds regardless of whether they utilise the tank water or the municipal tap water for domestic activities. Future studies should focus on elucidating the health risk associated with chronic exposure to these compounds at the concentrations (µg/L) recorded. This is especially important for compounds such as triclosan, triclocarban and methylparaben which exhibit endocrine disrupting characteristics. Epidemiological and animal studies link exposure to these compounds to neurobehavioural and neurodevelopmental changes, Parkinson's disease, asthma, reproductive effects and infertility, metabolic syndrome (including the development of obesity), bone and immune disorders and cancer in humans, amongst other conditions (Schug et al. 2016).

Correlation analysis between the MST markers, CST markers and the indicator organisms (enumerated with culturing techniques and quantified by qPCR) then revealed significant positive correlations for: adenovirus versus *E. coli* (enumerated with culturing techniques) ($p = 0.000$), the HF183 marker versus *E. coli* (quantified by qPCR) ($p = 0.023$), Lachnospiraceae versus heterotrophic bacteria ($p = 0.000$) and human mtDNA versus enterococci (enumerated with the culturing techniques) ($p = 0.026$) detected in the tank water samples. In addition, significant positive correlations were observed for caffeine versus; enterococci (quantified by qPCR) ($p = 0.000$); faecal coliforms ($p = 0.001$); total coliforms ($p = 0.000$) and enterococci (enumerated with culturing techniques) ($p = 0.002$), while salicylic acid positively correlated with total coliforms ($p = 0.024$) in the tank water samples. For the rooftop debris samples, significant correlations were observed for *E. coli* (quantified by qPCR) versus methylparaben ($p = 0.000$) and salicylic acid ($p = 0.042$), respectively. Based on these correlations between the various MST markers, CST markers and indicator organisms, it is recommended that the HF183 marker, Lachnospiraceae, human mtDNA, adenovirus, caffeine, salicylic acid and methylparaben should be included in a toolbox of markers for the screening of DRWH systems to supplement indicator organism analysis. Furthermore, as the qPCR assays employed to quantify the HF183 marker, Lachnospiraceae, human mtDNA and adenovirus proved to be more sensitive than the conventional PCR assays, it is recommended that qPCR analysis of the remaining MST markers (*Bifidobacterium*, *Enterococcus esp*, enterovirus, bovine mtDNA and porcine mtDNA) in the tank water and rooftop debris samples be conducted followed by correlation analysis of these MST markers with indicator organisms to determine which of these markers could also be incorporated into the DRWH source tracking toolbox.

The objective of **Chapter three** was to design and validate (on a small-scale) novel avian associated mtDNA markers to detect avian faecal contamination in rainwater harvesting systems in Kleinmond, Western Cape. Avian faecal matter has been reported to harbour high concentrations of faecal indicator bacteria (FIB) such as *E. coli* and enterococci (Alderisio & DeLuca, 1999; Fogarty et al. 2003; Ge et al. 2010). In addition, virulent *E. coli* (Wallace et al. 1997; Ahmed et al. 2012b; Chidamba & Korsten, 2015), *Salmonella* spp. (Fallacara et al. 2004; Kinzelman et al. 2008), *Campylobacter* spp. (Fallacara et al. 2004; Kinzelman et al. 2008), *Giardia lamblia* (Kuhn et al. 2002) and *Cryptosporidium parvum* (Kuhn et al. 2002) have all been associated with avian faecal matter. It is also known that birds travel great distances in the environment, display various feeding habits and may therefore be exposed to diverse microbial species (Ahmed et al. 2015). Subsequently, avian species may act as vectors of microorganisms and more importantly may contribute to the spread of pathogens in the environment (Ahmed et al. 2015). Therefore, there may be a significant health risk associated with exposure to water contaminated by avian faecal matter.

Primers for the detection of avian faecal matter were designed to target indigenous and related avian species of the Kleinmond area, which may contribute to the faecal contamination of rainwater harvesting systems investigated. Primers were designed for the amplification of the mitochondrial gene NADH dehydrogenase subunit 5, as mtDNA is species specific, abundant in eukaryotic cells and in faecal matter. In total, three primer sets were designed namely: AVF1 and AVR (designated AV1); AVF2 and AVR (designated AV2); and ND5F and ND5R (designated ND5). To validate the designed MST markers, 38 faecal samples from non-avian sources and 37 avian faecal samples from avian groups were screened for the markers. The avian host-sensitivity and host-specificity were then calculated. The host-sensitivity for the AV1, AV2 and ND5 markers was then equal to 1.00, 0.892 and 0.622, respectively. While the host-specificity was equal to 0.316, 0.0526 and 0.237 for the AV1, AV2 and ND5 markers, respectively. The tank water and rooftop debris samples collected in the Kleinmond Housing Scheme site (Chapter two) were then screened for the three avian markers in order to determine the background probability of avian pollution in the rainwater harvesting systems. Bayes' theorem was then applied to calculate the conditional probability that the avian mtDNA markers detected in the tank water and rooftop debris samples were as a result of true avian faecal contamination. For the 60 tank water samples, 51 (85%), 30 (50%) and 44 samples (73%) tested positive for the AV1, AV2 and ND5 markers, respectively. For the 60 rooftop debris samples, 54 (90%), 17 (28%) and 25 samples (42%) tested positive for the AV1, AV2 and ND5 markers, respectively. These detection frequencies then corresponded to an 89.2% and 92.9% probability that the detection of the AV1 marker was as a result of true avian faecal pollution of the tank water and rooftop debris, respectively. For the AV2 marker, the values corresponded to a 48.5% and 27.1% probability that the AV2 marker was detected in the tank water and rooftop debris as a result of avian faecal pollution, respectively. Lastly, for the ND5 marker, there was a 69.1% and 36.8% probability that the detection of the marker was as a result of true avian faecal pollution of the tank water and rooftop debris, respectively. The AV1 marker thus exhibited the greatest potential as an avian MST marker, despite the low host-specificity observed.

Avian mtDNA may thus be a valuable target for the development of avian source tracking markers, especially as a high host-sensitivity was obtained for the AV1 marker in the current study. Future research should however focus on increasing the host-specificity of the AV1 marker (which exhibited the highest host-sensitivity in this study), which may be achieved by combining the designed primers with qPCR and Taqman™ chemistries by designing a Taqman™ probe specific to avian mtDNA. In addition, instead of targeting a wide range of avian species with a single primer set (for example targeting pigeons, doves, seagulls and chickens with one primer set), specificity of avian mtDNA markers may be increased by targeting individual avian families or species with specific primer pairs for each family or species (thus

design primers specific to chickens, pigeons, doves and seagulls, respectively). Alternatively, new avian mtDNA primers could be designed by identifying SNP's and using mismatch amplification mutation assays (MAMA's) to increase specificity of the primers by targeting base pair mismatches between avian mtDNA from different avian species and mtDNA from mammalian faecal sources such as humans and pigs (for which a high percentage of cross-reactivity was observed in this study). Mismatch amplification mutation assays have previously been employed by Cebula et al. (1995) to distinguish between wild-type *E. coli*, Shiga-like toxin (SLT) producing *E. coli* of the non-O157:H7 serotype and various *E. coli* O157:H7 isolates. Cebula et al. (1995) then utilised three primer sets targeting the genes encoding for SLT-1 and SLT-2 and the *uidA* gene of *E. coli* respectively, as enterohemorrhagic *E. coli* (EHEC) strains may be distinguished using the SLT genes. In addition, *E. coli* O157:H7 contained a base pair mismatch in the *uidA* gene (a G instead of a T at position 92) and this base pair mismatch could be targeted to specifically identify *E. coli* O157:H7 in the samples analysed. Based on the results obtained, the MAMA primer sets could thus be utilised to characterise EHEC strains based on the amplification of the SLT genes and could specifically detect *E. coli* O157:H7 based on the presence of a product corresponding to the specific *uidA* gene. A similar assay to distinguish between enterotoxigenic *E. coli* strains has been developed, during which MAMA were combined with qPCR chemistry for the development of MAMA-qPCR (Sabui et al. 2012). Mismatch amplification mutation assays may thus be a valuable tool to increase avian mtDNA marker specificity in future studies. These assays could also be exploited in future studies to then design an avian mtDNA multiplex qPCR to detect all possible avian species in the Kleinmond area by targeting distinct base pair mismatches between different avian species.

The results in this study illustrated that avian species are contributing to the microbial contamination of the harvested rainwater as all three avian mtDNA MST markers were detected in the rooftop debris and tank water samples. Representative PCR products were sequenced to confirm the detection of avian mtDNA, and the results indicated that *Columbia livia* (pigeon) and *Calidris melanotos* (Pectoral Sandpiper) mtDNA was detected in the DRWH systems. Therefore, based on the confirmation of the presence of avian mtDNA (and thus avian faecal matter) in the DRWH systems, there may be a significant health risk associated with the use of the harvested rainwater as pathogens have been associated with avian faecal matter. For example, pigeons have been shown to act as vectors of antibiotic resistant and virulent *E. coli* in the environment (Chidamba & Korsten, 2015). It is therefore recommended that cost-effective treatment options such as solar pasteurization and solar disinfection be implemented to remove pathogenic organisms from harvested rainwater prior to utilising the rainwater (Dobrowsky et al. 2015; Reyneke et al. 2016; Strauss et al. 2016).

The current study thus highlighted that source tracking markers are present in rainwater harvesting systems, source tracking markers exhibit relationships with indicator organisms in

rainwater harvesting systems and avian species are a source of faecal contamination of harvested rainwater. Future research should focus on developing source tracking markers that may accurately detect other wild life and domestic animal faecal contamination (for example markers specific to dogs, cats, monkeys and rats) in order to characterise all the contamination sources possibly polluting the rainwater harvesting systems. Identifying the dominant contamination sources impacting environmental waters, allows for the design and implementation of effective prevention and remediation strategies to preserve or restore the quality of the water source. In addition, characterising all of the contamination sources impacting on the quality of roof-harvested rainwater and determining the viability of the microbial contaminants in harvested rainwater will thus aid in risk identification, risk characterisation and risk estimation in future studies. This information could then be utilised for quantitative microbial risk assessment, which should be conducted to estimate the health risk associated with the use of harvested rainwater for consumption and other domestic purposes such as cooking, cleaning and bathing.

References

Ahmed, W., Gardner, T. & Toze, S. 2011. Microbiological quality of roof-harvested rainwater and health risks: A review. *Journal of Environmental Quality.* 40(1):13-21.

Ahmed, W., Hamilton, K., Gyawali, P., Toze, S. & Haas, C. 2016. Evidence of avian and possum fecal contamination in rainwater tanks as determined by microbial source tracking approaches. *Applied and Environmental Microbiology.* doi:10.1128/AEM.00892-16

Ahmed, W., Harwood, V.J., Nguyen, K., Young, S., Hamilton, K. & Toze, S. 2015. Utility of *Helicobacter* spp. associated GFD markers for detecting avian faecal pollution in natural waters of two continents. *Water Research*. 88:613-622.

Ahmed, W., Hodgers, L., Sidhu, J.P. & Toze, S. 2012a. Faecal indicators and zoonotic pathogens in household drinking water taps fed from rainwater tanks in Southeast Queensland, Australia. *Applied and Environmental Microbiology.* 78(1):219-226.

Ahmed, W., Huygens, F., Goonetilleke, A. & Gardner, T. 2008. Real-time PCR detection of pathogenic microorganisms in roof-harvested rainwater in Southeast Queensland, Australia. *Applied and Environmental Microbiology.* 74(17):5490-5496.

Ahmed, W., Sidhu, J.P & Toze, S. 2012b. An attempt to identify the likely sources of *Escherichia coli* harboring toxin genes in rainwater tanks. *Environmental Science and Technology*. 46(9):5193-5197.

Ahmed, W., Vieritz, A., Goonetilleke, A. & Gardner, T. 2010a. Health risk from the use of roof-harvested rainwater in Southeast Queensland, Australia, as potable or non-potable water, determined using quantitative microbial risk assessment. *Applied and Environmental Microbiology.* 76(22):7382-7391.

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

Alderisio, K.A. & DeLuca, N. 1999. Seasonal enumeration of fecal coliform bacteria from the feces of ring-billed gulls (*Larus delawarensis*) and Canada geese (*Branta canadensis*). *Applied and Environmental Microbiology*. 65(12):5628-5630.

Anderson, K.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(6):3041-3048.

Bradshaw, J.K., Snyder, B.J., Oladeinde, A., Spidle, D., Berrang, M.E., Meinersmann, R.J., Oakley, B., Sidle, R.C., Sullivan, K. & Molina, M. 2016. Characterizing relationships among fecal indicator bacteria, microbial source tracking markers, and associated waterborne pathogen occurrence in stream water and sediments in a mixed land use watershed. *Water Research*. 101, 498-509.

Bonadonna, L., Briancesco, R., Ottaviani, M. & Veschetti, E. 2002. Occurrence of *Cryptosporidium* oocysts in sewage effluents and correlation with microbial, chemical and physical water variables. *Environmental Monitoring and Assessment.* 75(3):241-252.

Cebula, T.A., Payne, W.L. & Feng, P. 1995. Simultaneous identification of strains of *Escherichia coli* serotype O157:H7 and their shiga-like toxin type by mismatch amplification mutation assay-multiplex PCR. *Journal of Clinical Microbiology*. 33(1), 248-250.

Chidamba, L. & Korsten, L. 2015. Antibiotic resistance in *Escherichia coli* isolates from roof-harvested rainwater tanks and urban pigeon faeces as the likely source of contamination. *Environmental Monitoring and Assessment*. 187(7):1-15.

Crabtree, K.D., Ruskin, R.H., Shaw, S.B. & Rose, J.B. 1996. The detection of *Cryptosporidium* oocysts and *Giardia* cysts in cistern water in the US Virgin Islands. *Water Research.* 30(1):208-216.

Despins, C., Farahbakhsh, K. & Leidl, C. 2009. Assessment of rainwater quality from rainwater harvesting systems in Ontario, Canada. *Aqua.* 58(2):117.

Dobrowsky, P.H., Carstens, M., De Villiers, J., Cloete, T. & Khan, W. 2015. Efficiency of a closed-coupled solar pasteurization system in treating roof harvested rainwater. *Science of the Total Environment.* 536:206-214.

Dobrowsky, P.H., De Kwaadsteniet, M., Cloete, T.E. & Khan, W. 2014a. Distribution of indigenous bacterial pathogens and potential pathogens associated with roof-harvested rainwater. *Applied and Environmental Microbiology.* 80(7):2307-2316.

Dobrowsky, P.H., Mannel, D., De Kwaadsteniet, M., Prozesky, H., Khan, W. & Cloete, T.E. 2014b. Quality assessment and primary uses of harvested rainwater in Kleinmond, South Africa. *Water SA.* 40(3):401-406.

Dobrowsky, P.H., Van Deventer, A., De Kwaadsteniet, M., Ndlovu, T., Khan, S., Cloete, T.E. & Khan, W. 2014c. Prevalence of virulence genes associated with pathogenic *Escherichia coli* strains isolated from domestically harvested rainwater during low- and high-rainfall periods. *Applied and Environmental Microbiology.* 80(5):1633-1638.

Department of Water Affairs and Forestry (DWAF). 1996. South African water quality guidelines 2nd Edition, Volume 1: Domestic water use. CSIR Environmental Services, Pretoria.

Fallacara, D.M., Monahan, C.M., Morishita, T.Y., Bremer, C.A. & Wack, R.F. 2004. Survey of parasites and bacterial pathogens from free-living waterfowl in zoological settings. *Avian Diseases*. 48(4):759-767.

Field, K.G. & Samadpour, M. 2007. Fecal source tracking, the indicator paradigm, and managing water quality. *Water Research.* 41(16):3517-3538.

Fogarty, L., Haack, S., Wolcott, M. & Whitman, R. 2003. Abundance and characteristics of the recreational water quality indicator bacteria *Escherichia coli* and enterococci in gull feces. *Journal of Applied Microbiology*. 94 (5), 865-878.

Ge, Z., Nevers, M.B., Schwab, D.J. & Whitman, R.L. 2010. Coastal loading and transport of *Escherichia coli* at an embayed beach in Lake Michigan. *Environmental Science and Technology*. 44(17):6731-6737.

Glassmeyer, S.T., Furlong, E.T., Kolpin, D.W., Cahill, J.D., Zaugg, S.D., Werner, S.L., Meyer, M.T. & Kryak, D.D. 2005. Transport of chemical and microbial compounds from known wastewater discharges: Potential for use as indicators of human faecal contamination. *Environmental Science & Technology.* 39(14):5157-5169.

Griffith, J.F., Cao, Y., McGee, C.D. & Weisberg, S.B. 2009. Evaluation of rapid methods and novel indicators for assessing microbiological beach water quality. *Water Research*. 43(19):4900-4907.

Handia, L. 2005. Comparative study of rainwater quality in urban zambia. *Aqua.* 54:55-64.

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(6):3163-3170.

Harwood, V.J., Staley, C., Badgley, B.D., Borges, K. & Korajkic, A. 2014. Microbial source tracking markers for detection of faecal contamination in environmental waters: relationships between pathogens and human health outcomes. *FEMS Microbiology Reviews.* 38(1):1-40.

Jones-Lepp, T. 2006. Chemical markers of human waste contamination: Analysis of urobilin and pharmaceuticals in source waters. *Journal of Environmental Monitoring.* 8(4):472-478.

Kinzelman, J., McLellan, S.L., Amick, A., Preedit, J., Scopel, C.O., Olapade, O., Gradus, S., Singh, A. & Sedmak, G. 2008. Identification of human enteric pathogens in gull feces at Southwestern Lake Michigan bathing beaches. *Canadian Journal of Microbiology*. 54(12):1006-1015.

Kobayashi, A., Sano, D., Hatori, J., Ishii, S. & Okabe, S. 2013. Chicken- and duck-associated *Bacteroides–Prevotella* genetic markers for detecting faecal contamination in environmental water. *Applied Microbiology and Biotechnology.* 97(16):7427-7437.

Kuhn, R.C., Rock, C. M. & Oshima, K.H. 2002. Occurrence of *Cryptosporidium* and *Giardia* in wild ducks along the Rio Grande River valley in southern New Mexico. *Applied and Environmental Microbiology*. 68(1):161-165.

Leclerc, H., Mossel, D., Edberg, S. & Struijk, C. 2001. Advances in the bacteriology of the coliform group: their suitability as markers of microbial water safety. *Annual Reviews in Microbiology*. 55(1):201-234.

Lemarchand, K. & Lebaron, P. 2003. Occurrence of *Salmonella* spp. and *Cryptosporidium* spp. in a French coastal watershed: relationship with faecal indicators. *FEMS Microbiology Letters.* 218(1):203-209.

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 and Environmental Microbiology*. 81(3):850-860.

Lund, V. 1996. Evaluation of *E. coli* as an indicator for the presence of *Campylobacter jejuni* and *Yersinia enterocolitica* in chlorinated and untreated oligotrophic lake water. *Water Research.* 30(6):1528-1534.

McQuaig, S.M., Griffith, J. & Harwood, V.J. 2012. Association of fecal indicator bacteria with human viruses and microbial source tracking markers at coastal beaches impacted by nonpoint source pollution. *Applied and Environmental Microbiology*. 78(18):6423-6432.

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(11), 3379-3388.

Muscillo, M., Pourshaban, M., Iaconelli, M., Fontana, S., Di Grazia, A., Manzara, S., Fadda, G., Santangelo, R. & La Rosa, G. 2008. Detection and quantification of human adenoviruses in surface waters by nested PCR, TaqMan real-time PCR and cell culture assays. *Water, Air, & Soil Pollution.* 191(1-4):83-93.

Mwenge Kahinda, J., Taigbenu, A.E. & Boroto, J.R. 2007. Domestic rainwater harvesting to improve water supply in rural South Africa. *Physics and Chemistry of the Earth, Parts A/B/C.* 32(15):1050-1057.

National Health and Medical Research Council (NHMRC) & Natural Resource Management Ministerial Council (NRMMC). 2011. Australian Drinking Water Guidelines Paper 6 National Water Quality Management Strategy. National Health and Medical Research Council, National Resource Management Ministerial Council Commonwealth of Australia, Canberra.

Newton, R.J., Vandewalle, J.L., Borchardt, M.A., Gorelick, M.H. & McLellan, S.L. 2011. Lachnospiraceae and Bacteroidales alternative fecal indicators reveal chronic human sewage contamination in an urban harbor. *Applied and Environmental Microbiology*. 77(19):6972-6981.

Peng, X., Yu, Y., Tang, C., Tan, J., Huang, Q. & Wang, Z. 2008. Occurrence of steroid estrogens, endocrine-disrupting phenols, and acid pharmaceutical residues in urban riverine water of the pearl river delta, South China. *Science of the Total Environment*. 397 (1):158-166.

Reyneke, B., Dobrowsky, P.H., Ndlovu, T., Khan, S. & Khan, W. 2016. EMA-qPCR to monitor the efficiency of a closed-coupled solar pasteurization system in reducing *Legionella* contamination of roof-harvested rainwater. *Science of the Total Environment*. 553:662-670.

Rusiñol, M., Moriarty, E., Lin, S., Bofill-Mas, S. & Gilpin, B. 2016. Human-, ovine-, and bovine-specific viral source tracking tools to discriminate between the major fecal sources in agricultural waters. *Food and Environmental Virology*. 8(1):34-45.

Sabui, S., Dutta, S., Debnath, A., Ghosh, A., Hamabata, T., Rajendran, K., Ramamurthy, T., Nataro, J.P., Sur, D., Levine, M.M. & Chatterjee, N.S. 2012. Real-time PCR-based mismatch amplification mutation assay for specific detection of CS6-expressing allelic variants of enterotoxigenic *Escherichia coli* and its application in assessing diarrheal cases and asymptomatic controls. *Journal of Clinical Microbiology*. 50(4):1308-1312.

Schug, T.T., Johnson, A.F., Birnbaum, L.S., Colborn, T., Guillette Jr, L.J., Crews, D.P., Collins, T., Soto, A.M., vom Saal, F.S. & McLachlan, J.A. 2016. Minireview: Endocrine disruptors: Past lessons and future directions. *Molecular Endocrinology.* 30(8):833-847.

Sidhu, J.P., 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.* 463:488-496.

Simmons, G., Jury, S., Thornley, C., Harte, D., Mohiuddin, J. & Taylor, M. 2008. A Legionnaires' Disease outbreak: a water blaster and roof-collected rainwater systems. *Water Research.*42(6):1449-1458.

South African Bureau of Standards (SABS). 2005. South African National Standards (SANS) 241: Drinking water quality management guide for water services authorities. Annexure 1. South African Bureau of Standards. ISBN 0-626-17752-9.

Staley, Z.R., Grabuski, J., Sverko, E. & Edge, T.A. 2016. Comparison of microbial and chemical source tracking markers to identify fecal contamination sources in the Humber River and associated stormwater outfalls. *Applied and Environmental Microbiology*. 82(21):6357-6366.

Strauss, A., Dobrowsky, P.H., Ndlovu, T., Reyneke, B. & Khan, W. 2016. Comparative analysis of solar pasteurization versus solar disinfection for the treatment of harvested rainwater. *BMC Microbiology*. 16(289):1-16.

Tallon, P., Magajna, B., Lofranco, C. & Leung, K.T. 2005. Microbial indicators of faecal contamination in water: a current perspective. *Water, Air and Soil Pollution*. 166(1-4):139-166.

Verrinder, G. & Keleher, H. 2001. Domestic drinking water in rural areas: Are water tanks on farms a health hazard. *Environmental Health.* 1(3):51.

Villemur, R., Imbeau, M., Vuong, M.N., Masson, L. & Payment, P. 2015. An environmental survey of surface waters using mitochondrial DNA from human, bovine and porcine origin as faecal source tracking markers. *Water Research.* 69:143-153.

Wallace, J., Cheasty, T. & Jones, K. 1997. Isolation of Verocytotoxin-producing *Escherichia coli* O157 from wild birds. *Journal of Applied Microbiology*. 82(3):399-404.

Waso, M., Ndlovu, T., Dobrowsky, P.H., Khan, S. & Khan, W. 2016. Presence of microbial and chemical source tracking markers in roof-harvested rainwater and catchment systems for the detection of fecal contamination. *Environmental Science and Pollution Research*. 23(17):16987–17001.

Wolf, S., Hewitt, J. & Greening, G.E. 2010. Viral multiplex quantitative PCR assays for racking sources of faecal contamination. *Applied and Environmental Microbiology.* 76(5):1388-1394.

World Health Organization (WHO). 2011. Guidelines for drinking-water quality, Fourth edition. World Health Organization, Geneva, Switzerland. ISBN 978 92 4 154815 1.

Zheng, X., Qiu, M., Ke, X., Guan, W., Li, J., Huo, S., Chen, S., Zhong, X., Zhou, W. & Xiong, Y. 2016. Detection of novel adenoviruses in fecal specimens from rodents and shrews in southern china. *Virus Genes.* 52(3):417-421.

Appendix

Fig. A1 Images illustrating the extent of human activity around the DRWH tanks sampled in this study where household waste is frequently stored on top of the storage tanks (**A** and **B**), household waste is prevalent around the rainwater storage tanks (**C**) and some of the sampled rainwater storage tanks do not have lids or the lids are broken (**D**).

Table A1 The qPCR performance characteristics (± standard deviation) of the respective qPCR assays (n = 8) utilized to quantify *Bacteroides* HF183, adenovirus, human mtDNA, Lachnospiraceae, *Escherichia coli* and enterococci in all the tank water and rooftop debris samples.

qPCR Assay	Amplification efficiency (*E*)	Slope	y-intercept	Correlation coefficient (r^2)
Bacteroides HF183	99.4 ± 0.05	-3.3583 ± 0.14	36.915 ± 0.50	0.98 ± 0.014
Adenovirus	94.8 ± 0.16	-3.67435 ± 0.55	36.835 ± 4.64	0.99 ± 0.013
Human mtDNA	92.8 ± 0.02	-3.73055 ± 0.07	38.260 ± 0.88	0.98 ± 0.014
Lachnospiraceae	93.3 ± 0.16	-3.75713 ± 0.51	40.360 ± 2.99	0.99 ± 0.019
Escherichia coli	96.9 ± 0.10	-3.50578 ± 0.28	37.572 ± 3.16	0.95 ± 0.005
Enterococci	98.3 ± 0.05	-3.41653 ± 0.13	36.108 ± 1.26	0.98 ± 0.009

Table A2 Concentrations in µg/L of the CST markers detected in the tank water samples per tank collected during sampling one to six.

Sampling	Tank	Caffeine (µg/L)	Salicylic acid (µg/L)	Methyl-paraben (µg/L)	Triclosan (µg/L)	Triclo-carban (µg/L)	Acetamino-phen (µg/L)
1	A	1.60	14.30	5.80	1.40	1.10	9.60
	B	4.20	17.00	2.80	0.90	0.80	9.30
	C	1.40	11.30	1.70	0.70	1.00	9.40
	D	1.00	14.60	11.64	0.80	0.70	9.60
	E	1.40	13.70	3.30	0.90	1.10	9.80
	F	2.38	12.70	1.80	0.50	0.90	9.10
	G	1.00	10.80	1.80	0.50	1.00	8.90
	H	1.20	15.80	3.00	0.60	0.90	8.70
	I	2.65	15.70	1.50	0.70	1.10	8.00
	J	1.00	14.60	2.70	0.60	0.90	8.80
2	A	2.20	12.60	4.90	2.20	1.10	8.80
	B	2.30	18.70	3.60	0.80	1.10	8.90
	C	1.30	11.40	3.50	0.80	1.00	9.20
	D	1.70	13.40	4.20	1.00	1.00	9.80
	E	4.10	20.00	4.20	0.80	1.00	8.50
	F	1.90	22.60	4.50	1.10	1.20	9.20
	G	1.70	16.70	4.60	1.00	1.30	8.70
	H	1.90	19.20	17.40	0.50	1.20	9.20
	I	1.70	14.50	3.20	1.00	1.10	9.50
	J	1.80	15.00	6.30	1.70	1.40	9.60
3	A	1.90	11.50	2.40	0.90	1.30	9.30
	B	1.00	13.50	1.40	0.70	1.10	9.70
	C	0.90	13.80	2.00	0.90	1.30	9.60
	D	3.20	12.20	3.70	1.10	1.30	9.60
	E	1.60	11.80	2.70	1.00	1.30	9.00
	F	1.70	14.50	2.40	2.30	1.40	9.60
	G	1.00	11.90	1.50	1.20	1.20	8.50
	H	1.50	10.10	2.50	0.70	1.20	7.90
	I	1.60	10.90	2.60	0.90	1.30	8.10
	J	1.80	11.40	1.80	0.90	1.20	8.10
4	A	2.80	12.10	3.50	16.00	1.10	8.70
	B	5.00	10.30	4.50	10.10	1.20	8.50
	C	2.00	11.80	2.50	9.60	1.00	8.00
	D	2.10	11.80	2.30	7.30	1.00	9.80
	E	5.00	10.20	0.50	6.30	0.40	7.50
	F	2.80	11.70	2.90	4.90	0.40	6.00
	G	1.90	11.50	1.00	5.30	0.40	6.80
	H	2.00	8.00	1.10	6.80	0.40	6.00
	I	1.50	12.10	1.40	3.80	0.40	6.70
	J	1.90	9.30	1.40	3.40	0.30	6.90

Table A2 (Continued) Concentrations in µg/L of the CST markers detected in the tank water samples per tank collected during sampling one to six.

Sampling	Tank	Caffeine (µg/L)	Salicylic acid (µg/L)	Methyl-paraben (µg/L)	Triclosan (µg/L)	Triclo-carban (µg/L)	Acetamino-phen (µg/L)
5	A	4.00	17.50	2.50	26.10	4.90	6.90
	B	8.00	16.70	0.20	12.60	0.70	10.40
	C	3.50	14.50	3.30	11.60	0.90	10.60
	D	3.00	13.10	1.80	7.40	0.90	10.10
	E	2.80	11.90	1.70	7.50	0.80	9.90
	F	2.80	13.30	1.70	4.50	0.60	10.40
	G	3.90	BDL[a]	1.40	18.30	0.80	9.60
	H	2.70	10.50	3.10	4.80	0.80	8.50
	I	2.70	13.30	2.10	3.90	0.70	9.40
	J	2.70	13.60	2.20	3.00	0.70	10.10
6	A	3.50	12.30	2.70	14.60	0.70	10.30
	B	6.70	10.90	2.50	12.50	0.70	9.70
	C	2.40	10.20	3.40	9.80	0.70	9.80
	D	2.90	14.70	1.50	11.20	0.70	4.00
	E	3.20	9.90	0.20	9.50	0.70	8.90
	F	2.70	10.40	1.60	7.30	0.70	9.00
	G	2.90	11.80	1.40	7.90	0.70	7.90
	H	3.50	10.10	2.30	12.10	0.60	9.60
	I	2.30	13.80	2.60	4.80	0.70	10.20
	J	2.10	2.90	2.80	0.40	0.50	1.00

[a]BDL – below detection limit (< 0.1 µg/L)

Table A3 Concentrations in µg/L of the CST markers detected in the rooftop debris samples per gutter system connected to a sampled rainwater harvesting tank, collected during sampling one to six.

Sampling	Gutter connected to tank	Caffeine (µg/L)	Salicylic acid (µg/L)	Methyl-paraben (µg/L)	Triclosan (µg/L)	Triclo-carban (µg/L)	Acetamino-phen (µg/L)
1	A	2.00	12.90	3.00	2.40	1.40	9.30
	B	1.50	10.30	1.80	0.90	1.10	8.60
	C	1.30	11.50	1.50	0.90	0.90	6.80
	D	1.60	11.60	2.90	1.00	0.90	9.60
	E	2.20	11.90	1.90	1.00	2.00	9.60
	F	2.20	13.40	1.40	1.00	1.50	9.20
	G	2.40	11.30	2.50	1.30	1.30	9.30
	H	1.30	9.50	2.30	0.90	1.20	8.90
	I	1.60	10.70	1.20	1.60	1.10	8.70
	J	1.40	11.10	1.80	0.90	1.10	7.60
2	A	1.80	12.10	7.30	1.50	1.10	6.60
	B	1.40	11.00	11.90	7.20	1.30	10.00
	C	2.60	15.50	19.50	1.80	1.60	9.90
	D	2.10	11.70	16.20	1.80	1.20	9.70
	E	3.00	10.80	8.10	0.80	1.10	10.00
	F	2.90	12.40	16.40	15.00	1.70	9.10
	G	2.30	9.90	13.00	2.00	1.00	9.30
	H	1.00	10.50	11.20	1.80	1.30	7.30
	I	1.60	16.40	10.20	1.70	0.80	5.70
	J	1.20	15.70	14.70	1.00	1.10	9.70
3	A	1.30	12.50	1.40	0.80	0.90	9.60
	B	1.60	10.20	1.30	1.90	1.50	9.60
	C	1.70	9.60	1.30	1.20	1.40	9.80
	D	1.90	9.60	1.40	1.00	1.00	9.60
	E	2.70	10.80	2.00	1.30	1.20	9.20
	F	1.80	7.50	2.00	0.70	0.80	8.80
	G	1.70	9.70	2.30	1.00	1.40	8.60
	H	1.50	11.60	1.40	0.60	0.90	8.90
	I	1.40	9.90	1.00	0.80	0.90	9.00
	J	1.90	13.40	2.00	1.30	1.80	9.30
4	A	1.70	10.80	1.40	1.00	1.10	9.00
	B	1.50	10.60	0.90	1.60	1.50	9.20
	C	2.00	8.70	1.40	1.10	1.00	9.40
	D	1.40	9.90	1.60	0.90	1.00	8.50
	E	1.50	10.00	2.00	1.00	1.00	8.20
	F	1.50	7.20	1.90	0.80	0.90	8.30
	G	1.90	10.40	1.90	1.30	1.60	9.30
	H	1.60	9.50	1.70	0.70	0.80	8.90
	I	1.30	10.60	1.20	0.70	1.00	8.50
	J	3.40	15.50	2.40	1.40	1.90	10.90

Table A3 (Continued) Concentrations in µg/L of the CST markers detected in the rooftop debris samples per gutter system connected to a sampled rainwater harvesting tank, collected during sampling one to six.

Sampling	Gutter connected to tank	Caffeine (µg/L)	Salicylic acid (µg/L)	Methyl-paraben (µg/L)	Triclosan (µg/L)	Triclo-carban (µg/L)	Acetamino-phen (µg/L)
5	A	1.80	10.40	1.60	6.30	1.20	8.80
	B	1.40	10.40	1.50	2.40	1.00	8.60
	C	1.70	10.20	1.50	5.50	1.40	7.50
	D	1.50	11.80	2.10	3.40	1.90	9.10
	E	1.50	6.00	0.40	0.30	1.40	5.20
	F	2.90	10.90	1.90	2.60	2.20	8.70
	G	1.30	10.00	1.20	1.50	0.80	7.50
	H	2.20	8.60	1.80	3.90	1.20	7.70
	I	1.70	10.30	1.50	3.50	1.30	8.30
	J	2.00	9.00	1.90	3.70	1.90	8.20
6	A	2.20	9.70	2.30	5.90	1.40	6.30
	B	2.00	10.00	1.40	2.40	1.20	8.30
	C	1.70	9.10	3.40	6.00	1.50	8.20
	D	1.90	10.50	1.60	3.50	2.00	8.00
	E	2.40	12.60	2.00	5.40	3.30	8.70
	F	2.70	10.80	2.30	1.90	1.90	7.60
	G	1.50	9.30	0.80	1.70	0.90	7.40
	H	2.10	10.70	1.20	4.10	1.40	7.10
	I	1.80	10.20	1.30	3.40	1.40	6.60
	J	1.90	10.90	1.80	3.90	2.20	8.70

Table A4 Summary of the correlations observed between the indicator organisms and CST markers detected in the tank water samples with the significant correlations ($p < 0.05$) indicated in red.

	E. coli (qPCR)	Entero-cocci (qPCR)	*E. coli* (culturing)	Total coliforms	Entero-cocci (culturing)	Fecal coliforms	Hetero-trophic bacteria	Caffeine	Salicylic acid	Aceta-minophen	Triclo-carban	Triclo-san
E. coli (qPCR)	1	-0.004	0.0045	-0.0487	0.074	0.154	-0.14	0.0338	-0.295	-0.224	-0.291	0.257
	p = ---	*p* = 0.977	*p* = 0.974	*p* = 0.722	*p* = 0.588	*p* = 0.259	*p* = 0.303	*p* = 0.805	*p* = 0.028	*p* = 0.098	*p* = 0.029	*p* = 0.056
Entero-cocci (qPCR)	-0.004	1	0.0386	0.473	0.401	0.437	0.137	0.863	0.262	0.0538	0.0098	-0.126
	p = 0.977	*p* = ---	*p* = 0.778	*p* = 0.000	*p* = 0.002	*p* = 0.001	*p* = 0.313	*p* = 0.000	*p* = 0.051	*p* = 0.694	*p* = 0.943	*p* = 0.353
E. coli (culturing)	0.0045	0.0386	1	0.223	0.242	0.216	-0.0377	-0.0538	0.0546	0.0724	0.0204	-0.0826
	p = 0.974	*p* = 0.778	*p* = ---	*p* = 0.098	*p* = 0.073	*p* = 0.109	*p* = 0.783	*p* = 0.693	*p* = 0.690	*p* = 0.596	*p* = 0.881	*p* = 0.545
Total coliforms	-0.0487	0.473	0.223	1	0.609	0.64	0.0856	0.483	0.301	0.0775	-0.0971	-0.0065
	p = 0.722	*p* = 0.000	*p* = 0.098	*p* = ---	*p* = 0.000	*p* = 0.000	*p* = 0.531	*p* = 0.000	*p* = 0.024	*p* = 0.570	*p* = 0.477	*p* = 0.962
Entero-cocci (culturing)	0.074	0.401	0.242	0.609	1	0.589	0.0164	0.399	0.244	-0.0928	-0.120	0.028
	p = 0.588	*p* = 0.002	*p* = 0.073	*p* = 0.000	*p* = ---	*p* = 0.000	*p* = 0.904	*p* = 0.002	*p* = 0.071	*p* = 0.496	*p* = 0.378	*p* = 0.838
Fecal coliforms	0.154	0.437	0.216	0.64	0.589	1	0.0436	0.447	0.171	-0.267	-0.0433	0.125
	p = 0.259	*p* = 0.001	*p* = 0.109	*p* = 0.000	*p* = 0.000	*p* = ---	*p* = 0.750	*p* = 0.001	*p* = 0.208	*p* = 0.047	*p* = 0.752	*p* = 0.358
Hetero-trophic bacteria	-0.14	0.137	-0.0377	0.0856	0.0164	0.0436	1	0.0885	0.0763	0.107	0.182	-0.359
	p = 0.303	*p* = 0.313	*p* = 0.783	*p* = 0.531	*p* = 0.904	*p* = 0.750	*p* = ---	*p* = 0.517	*p* = 0.576	*p* = 0.432	*p* = 0.179	*p* = 0.007
Methyl-paraben	-0.0787	0.0248	-0.0176	-0.0032	-0.0739	-0.0407	-0.0664	-0.0765	0.133	0.100	-0.0403	-0.123
	p = 0.564	*p* = 0.856	*p* = 0.898	*p* = 0.981	*p* = 0.588	*p* = 0.766	*p* = 0.627	*p* = 0.575	*p* = 0.330	*p* = 0.463	*p* = 0.768	*p* = 0.367
Caffeine	0.0338	0.863	-0.0538	0.483	0.399	0.447	0.0885	1	0.184	0.0178	0.0308	0.124
	p = 0.805	*p* = 0.000	*p* = 0.693	*p* = 0.000	*p* = 0.002	*p* = 0.001	*p* = 0.517	*p* = ---	*p* = 0.174	*p* = 0.896	*p* = 0.822	*p* = 0.361
Salicylic acid	-0.295	0.262	0.0546	0.301	0.244	0.171	0.0763	0.184	1	0.316	0.339	-0.242
	p = 0.028	*p* = 0.051	*p* = 0.690	*p* = 0.024	*p* = 0.071	*p* = 0.208	*p* = 0.576	*p* = 0.174	*p* = ---	*p* = 0.018	*p* = 0.011	*p* = 0.072
Acetamino-phen	-0.224	0.0538	0.0724	0.0775	-0.0928	-0.267	0.107	0.0178	0.316	1	0.0746	-0.0804
	p = 0.098	*p* = 0.694	*p* = 0.596	*p* = 0.570	*p* = 0.496	*p* = 0.047	*p* = 0.432	*p* = 0.896	*p* = 0.018	*p* = ---	*p* = 0.585	*p* = 0.556
Triclo-carban	-0.291	0.0098	0.0204	-0.0971	-0.120	-0.0433	0.182	0.0308	0.339	0.0746	1	0.352
	p = 0.029	*p* = 0.943	*p* = 0.881	*p* = 0.477	*p* = 0.378	*p* = 0.752	*p* = 0.179	*p* = 0.822	*p* = 0.011	*p* = 0.585	*p* = ---	*p* = 0.008
Triclosan	0.257	-0.126	-0.0826	-0.0065	0.028	0.125	-0.359	0.124	-0.242	-0.0804	0.352	1
	p = 0.056	*p* = 0.353	*p* = 0.545	*p* = 0.962	*p* = 0.838	*p* = 0.358	*p* = 0.007	*p* = 0.361	*p* = 0.072	*p* = 0.556	*p* = 0.008	*p* = ---

Table A5 Summary of the correlations observed between the MST and CST markers detected in the tank water samples with the significant correlations ($p < 0.05$) indicated in red.

	HF183	Adenovirus	Human mtDNA	Lachno-spiraceae	Methyl-paraben	Caffeine	Salicylic acid	Acetamino-phen	Triclocarban	Triclosan
HF183	1	-0.0927	0.0875	0.0233	-0.084	-0.0017	-0.175	-0.166	-0.249	0.282
	p = ---	*p* = 0.497	*p* = 0.522	*p* = 0.865	*p* = 0.538	*p* = 0.990	*p* = 0.198	*p* = 0.222	*p* = 0.065	*p* = 0.036
Adenovirus	-0.0927	1	-0.0536	0.0588	-0.0196	-0.0573	0.0408	0.0867	0.0222	-0.106
	p = 0.497	*p* = ---	*p* = 0.695	*p* = 0.667	*p* = 0.886	*p* = 0.675	*p* = 0.765	*p* = 0.525	*p* = 0.871	*p* = 0.435
Human mtDNA	0.0875	-0.0536	1	-0.0335	-0.0523	-0.0027	-0.0226	-0.0086	-0.140	0.0339
	p = 0.522	*p* = 0.695	*p* = ---	*p* = 0.806	*p* = 0.702	*p* = 0.984	*p* = 0.869	*p* = 0.950	*p* = 0.304	*p* = 0.804
Lachnospiraceae	0.0233	0.0588	-0.0335	1	-0.0739	-0.157	-0.0698	0.0862	0.171	-0.200
	p = 0.865	*p* = 0.667	*p* = 0.806	*p* = ---	*p* = 0.588	*p* = 0.249	*p* = 0.609	*p* = 0.528	*p* = 0.209	*p* = 0.139
Methylparaben	-0.084	-0.0196	-0.0523	-0.0739	1	-0.0765	0.133	0.100	-0.0403	-0.123
	p = 0.538	*p* = 0.886	*p* = 0.702	*p* = 0.588	*p* = ---	*p* = 0.575	*p* = 0.330	*p* = 0.463	*p* = 0.768	*p* = 0.367
Caffeine	-0.0017	-0.0573	-0.0027	-0.157	-0.0765	1	0.184	0.0178	0.0308	0.124
	p = 0.990	*p* = 0.675	*p* = 0.984	*p* = 0.249	*p* = 0.575	*p* = ---	*p* = 0.174	*p* = 0.896	*p* = 0.822	*p* = 0.361
Salicylic acid	-0.175	0.0408	-0.0226	-0.0698	0.133	0.184	1	0.316	0.339	-0.242
	p = 0.198	*p* = 0.765	*p* = 0.869	*p* = 0.609	*p* = 0.330	*p* = 0.174	*p* = ---	*p* = 0.018	*p* = 0.011	*p* = 0.072
Acetaminophen	-0.166	0.0867	-0.0086	0.0862	0.100	0.0178	0.316	1	0.0746	-0.0804
	p = 0.222	*p* = 0.525	*p* = 0.950	*p* = 0.528	*p* = 0.463	*p* = 0.896	*p* = 0.018	*p* = ---	*p* = 0.585	*p* = 0.556
Triclocarban	-0.249	0.0222	-0.140	0.171	-0.0403	0.0308	0.339	0.0746	1	0.352
	p = 0.065	*p* = 0.871	*p* = 0.304	*p* = 0.209	*p* = 0.768	*p* = 0.822	*p* = 0.011	*p* = 0.585	*p* = ---	*p* = 0.008
Triclosan	0.282	-0.106	0.0339	-0.200	-0.123	0.124	-0.242	-0.0804	0.352	1
	p = 0.036	*p* = 0.435	*p* = 0.804	*p* = 0.139	*p* = 0.367	*p* = 0.361	*p* = 0.072	*p* = 0.556	*p* = 0.008	*p* = ---

Table A6 Summary of the correlations observed between the indicator organisms, MST and CST markers detected in the rooftop debris samples with the significant correlations ($p < 0.05$) indicated in red.

	HF183	Adeno-virus	Human mtDNA	Lachno-spiraceae	*E. coli* (qPCR)	Entero-cocci (qPCR)	Methyl-paraben	Caffeine	Salicylic acid	Aceta-minophen	Triclo-carban	Triclosan
HF183	1.00	0.40	-0.116	0.214	-0.210	-0.0509	-0.294	0.0913	-0.260	0.0327	-0.0053	-0.137
	p = ---	*p* = 0.002	*p* = 0.394	*p* = 0.113	*p* = 0.121	*p* = 0.710	*p* = 0.028	*p* = 0.503	*p* = 0.053	*p* = 0.811	*p* = 0.969	*p* = 0.313
Adenovirus	0.40	1	-0.0728	0.318	-0.121	-0.124	-0.210	0.129	-0.0452	-0.0943	0.121	0.160
	p = 0.002	*p* = ---	*p* = 0.594	*p* = 0.017	*p* = 0.375	*p* = 0.362	*p* = 0.121	*p* = 0.344	*p* = 0.741	*p* = 0.489	*p* = 0.376	*p* = 0.239
Human mtDNA	-0.12	-0.0728	1	-0.0952	0.104	-0.0495	0.121	-0.173	0.0389	-0.0395	-0.0792	0.151
	p = 0.394	*p* = 0.594	*p* = ---	*p* = 0.485	*p* = 0.447	*p* = 0.717	*p* = 0.375	*p* = 0.202	*p* = 0.776	*p* = 0.772	*p* = 0.562	*p* = 0.268
Lachno-spiraceae	0.21	0.318	-0.0952	1	-0.159	-0.0676	-0.231	0.167	-0.199	-0.293	0.508	0.299
	p = 0.113	*p* = 0.017	*p* = 0.485	*p* = ---	*p* = 0.242	*p* = 0.620	*p* = 0.087	*p* = 0.219	*p* = 0.142	*p* = 0.028	*p* = 0.000	*p* = 0.025
E. coli (qPCR)	-0.210	-0.121	0.104	-0.159	1	-0.0116	0.623	0.144	0.273	0.220	0.0132	0.004
	p = 0.121	*p* = 0.375	*p* = 0.447	*p* = 0.242	*p* = ---	*p* = 0.933	*p* = 0.000	*p* = 0.291	*p* = 0.042	*p* = 0.104	*p* = 0.923	*p* = 0.977
Entero-cocci (qPCR)	-0.0509	-0.124	-0.0495	-0.0676	-0.0116	1	-0.0925	-0.100	-0.0409	0.0091	0.178	-0.0163
	p = 0.710	*p* = 0.362	*p* = 0.717	*p* = 0.620	*p* = 0.933	*p* = ---	*p* = 0.498	*p* = 0.463	*p* = 0.765	*p* = 0.947	*p* = 0.190	*p* = 0.905
Methyl-paraben	-0.294	-0.210	0.121	-0.231	0.623	-0.0925	1	0.196	0.470	0.1823	-0.026	0.324
	p = 0.028	*p* = 0.121	*p* = 0.375	*p* = 0.087	*p* = 0.000	*p* = 0.498	*p* = ---	*p* = 0.148	*p* = 0.000	*p* = 0.179	*p* = 0.849	*p* = 0.015
Caffeine	0.0913	0.129	-0.173	0.167	0.144	-0.100	0.196	1	0.267	0.358	0.486	0.26
	p = 0.503	*p* = 0.344	*p* = 0.202	*p* = 0.219	*p* = 0.291	*p* = 0.463	*p* = 0.148	*p* = ---	*p* = 0.047	*p* = 0.007	*p* = 0.000	*p* = 0.051
Salicylic acid	-0.260	-0.0452	0.0389	-0.199	0.273	-0.0409	0.470	0.267	1	0.278	0.220	0.0684
	p = 0.053	*p* = 0.741	*p* = 0.776	*p* = 0.142	*p* = 0.042	*p* = 0.765	*p* = 0.000	*p* = 0.047	*p* = ---	*p* = 0.038	*p* = 0.103	*p* = 0.617
Acetamino-phen	0.0327	-0.0943	-0.0395	-0.293	0.220	0.0091	0.182	0.358	0.278	1	0.134	0.0099
	p = 0.811	*p* = 0.489	*p* = 0.772	*p* = 0.028	*p* = 0.104	*p* = 0.947	*p* = 0.179	*p* = 0.007	*p* = 0.038	*p* = ---	*p* = 0.325	*p* = 0.942
Triclo-carban	-0.0053	0.121	-0.0792	0.508	0.0132	0.178	-0.026	0.486	0.220	0.134	1	0.381
	p = 0.969	*p* = 0.376	*p* = 0.562	*p* = 0.000	*p* = 0.923	*p* = 0.190	*p* = 0.849	*p* = 0.000	*p* = 0.103	*p* = 0.325	*p* = ---	*p* = 0.004
Triclosan	-0.137	0.160	0.151	0.299	0.004	-0.0163	0.324	0.262	0.0684	0.0099	0.381	1
	p = 0.313	*p* = 0.239	*p* = 0.268	*p* = 0.025	*p* = 0.977	*p* = 0.905	*p* = 0.015	*p* = 0.051	*p* = 0.617	*p* = 0.942	*p* = 0.004	*p* = ---

Table A7 Summary of the correlations observed between the indicator organisms, MST and CST markers detected in the rooftop debris (RD) samples versus the tank water samples (RW) with the significant correlations ($p < 0.05$) indicated in red.

	HF183 RD	Adeno-virus RD	Human mtDNA RD	Lachno-spiraceae RD	*E. coli* RD	Entero-cocci RD	Methyl-paraben RD	Caffeine RD	Salicylic acid RD	Aceta-minophen RD	Triclocar-ban RD	Triclosan RD
HF183 T	0.50	0.06	-0.05	-0.16	-0.12	0.19	-0.20	-0.11	-0.24	0.00	-0.20	-0.21
	p = 0.000	*p* = 0.639	*p* = 0.709	*p* = 0.226	*p* = 0.362	*p* = 0.164	*p* = 0.135	*p* = 0.433	*p* = 0.078	*p* = 0.994	*p* = 0.130	*p* = 0.123
Adeno-virus T	-0.15	0.09	-0.03	-0.08	-0.03	-0.05	-0.05	0.11	0.09	0.13	0.20	-0.06
	p = 0.263	*p* = .0528	*p* = 0.848	*p* = 0.553	*p* = 0.821	*p* = 0.741	*p* = 0.699	*p* = 0.429	*p* = 0.496	*p* = 0.338	*p* = 0.130	*p* = 0.635
Human mtDNA T	0.54	0.25	-0.04	0.23	-0.09	-0.07	-0.13	0.00	-0.11	-0.05	0.04	0.05
	p = 0.000	*p* = 0.067	*p* = 0.774	*p* = 0.095	*p* = 0.520	*p* = 0.604	*p* = 0.325	*p* = 0.985	*p* = 0.400	*p* = 0.708	*p* = 0.758	*p* = 0.729
Lachno-spiraceae T	0.11	-0.27	-0.09	-0.29	-0.12	0.26	-0.18	-0.07	-0.13	0.23	-0.16	-0.25
	p = 0.408	*p* = 0.047	*p* = 0.507	*p* = 0.028	*p* = 0.379	*p* = 0.056	*p* = 0.176	*p* = 0.585	*p* = 0.353	*p* = 0.086	*p* = 0.246	*p* = 0.062
E. coli T	0.37	0.31	-0.04	0.25	-0.19	0.23	-0.27	0.11	0.00	-0.01	0.34	0.06
	p = 0.006	*p* = 0.018	*p* = 0.768	*p* = 0.068	*p* = 0.166	*p* = 0.086	*p* = 0.044	*p* = 0.422	*p* = 0.977	*p* = 0.922	*p* = 0.010	*p* = 0.669
Entero-cocci T	-0.14	-0.13	0.64	-0.19	0.13	0.15	0.23	-0.12	0.07	0.18	-0.08	0.23
	p = 0.298	*p* = 0.353	*p* = 0.000	*p* = 0.159	*p* = 0.323	*p* = 0.283	*p* = 0.088	*p* = 0.369	*p* = 0.619	*p* = 0.189	*p* = 0.565	*p* = 0.086
Methyl-paraben T	-0.14	-0.06	-0.03	-0.10	-0.01	-0.05	0.05	-0.09	0.07	0.11	-0.15	-0.06
	p = 0.296	*p* = 0.652	*p* = 0.826	*p* = 0.457	*p* = 0.918	*p* = 0.732	*p* = 0.704	*p* = 0.502	*p* = 0.618	*p* = 0.413	*p* = 0.286	*p* = 0.668
Caffeine T	-0.08	0.02	0.62	0.02	0.10	-0.05	0.16	-0.09	-0.07	0.09	0.00	0.32
	p = 0.556	*p* = 0.884	*p* = 0.000	*p* = 0.858	*p* = 0.459	*p* = 0.740	*p* = 0.249	*p* = 0.489	*p* = 0.588	*p* = 0.520	*p* = 0.974	*p* = 0.016
Salicylic acid T	-0.24	-0.22	0.12	-0.28	0.05	-0.02	0.43	0.07	0.07	0.12	-0.06	0.34
	p = 0.070	*p* = 0.102	*p* = 0.376	*p* = 0.040	*p* = 0.740	*p* = 0.886	*p* = 0.001	*p* = 0.625	*p* = 0.618	*p* = 0.393	*p* = 0.635	*p* = 0.010
Aceta-minophen T	-0.07	0.05	0.07	-0.11	0.10	0.08	0.15	-0.01	0.06	-0.09	-0.12	0.01
	p = 0.596	*p* = 0.705	*p* = 0.612	*p* = 0.413	*p* = 0.453	*p* = 0.554	*p* = 0.258	*p* = 0.950	*p* = 0.667	*p* = 0.503	*p* = 0.387	*p* = 0.943
Triclo-carban T	-0.17	0.13	0.01	-0.12	0.01	-0.04	0.10	-0.07	0.05	0.10	-0.15	0.20
	p = 0.203	*p* = 0.335	*p* = 0.922	*p* = 0.399	*p* = 0.968	*p* = 0.751	*p* = 0.460	*p* = 0.614	*p* = 0.712	*p* = 0.461	*p* = 0.255	*p* = 0.142
Triclosan T	0.20	0.44	-0.13	0.31	-0.18	0.07	-0.29	-0.07	-0.26	-0.23	0.05	0.21
	p = 0.134	*p* = 0.001	*p* = 0.325	*p* = 0.019	*p* = 0.175	*p* = 0.632	*p* = 0.027	*p* = 0.616	*p* = 0.052	*p* = 0.092	*p* = 0.740	*p* = 0.125

Table A8 Correlations of the MST markers, CST markers and indicator organisms detected in the tank water samples versus the rainfall observed for the sampling period with the significant correlations ($p < 0.05$) indicated in red.

Organism/Marker	Rainfall
HF183	-0.12
	$p = 0.820$
Adenovirus	0.05
	$p = 0.926$
Human mtDNA	-0.45
	$p = 0.372$
Lachnospiraceae	-0.60
	$p = 0.206$
E. coli (qPCR)	-0.06
	$p = 0.912$
Enterococci (qPCR)	0.05
	$p = 0.928$
E. coli (culturing)	0.62
	$p = 0.193$
Total coliforms	0.74
	$p = 0.091$
Enterococci (culturing)	-0.05
	$p = 0.928$
Fecal coliforms	0.89
	$p = 0.018$
Heterotrophic bacteria	-0.60
	$p = 0.204$
Methylparaben	0.09
	$p = 0.865$
Caffeine	0.37
	$p = 0.468$
Salicylic acid	-0.15
	$p = 0.778$
Acetaminophen	-0.45
	$p = 0.373$
Triclocarban	0.41
	$p = 0.414$
Triclosan	0.34
	$p = 0.504$

Table A9 Correlations of the MST markers CST markers and indicator organisms detected in the rooftop debris samples versus the rainfall observed for the sampling period with the significant correlations ($p < 0.05$) indicated in red.

Organism/Marker	Rainfall
HF183	-0.52
	$p = 0.288$
Adenovirus	-0.32
	$p = 0.535$
Human mtDNA	0.04
	$p = 0.936$
Lachnospiraceae	0.45
	$p = 0.368$
E. coli (qPCR)	0.16
	$p = 0.765$
Enterococci (qPCR)	0.40
	$p = 0.431$
Methylparaben	0.15
	$p = 0.773$
Caffeine	0.83
	$p = 0.043$
Salicylic acid	0.16
	$p = 0.768$
Acetaminophen	-0.65
	$p = 0.162$
Triclocarban	0.83
	$p = 0.043$
Triclosan	0.61
	$p = 0.198$