Bacteroides species as indicators of faecal pollution in environmental water sources: A literature review

A Wolmarans

orcid.org 0000-0003-3568-0117

Dissertation submitted in fulfilment of the requirements for the degree *Master of Science in Environmental Sciences* at North-West University

Supervisor:	Prof CC Bezuidenhout
Co-supervisor:	Prof RA Adeleke

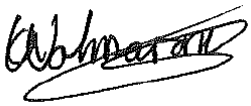
Graduation May 2019
21776539

Table of contents

DECLARATION

I declare that the dissertation submitted by me for the degree Magister Scientiae in Environmental Sciences at the North-West University (Potchefstroom Campus), Potchefstroom, North-West, South Africa, is my own independent work and has not previously been submitted by me at another university.

Signed in Potchefstroom, South Africa

Signature: 

Date: 20-11-18

Adéle Wolmarans

ACKNOWLEDGEMENTS

It is with immense gratitude that I wish to thank everybody involved in the process of completion of this dissertation. Thank you in particular to Prof. Adeleke for his insight and financial support via the Agricultural Research Council. I am grateful to Prof. Carlos Bezuidenhout for his guidance throughout. Thank you to colleagues: G. O'Reilly, J.J. Bezuidenhout and R. Coetzee and all others I have not listed whom assisted with brainstorming sessions, proofreading, answering countless questions and effortlessly resolving formatting issues. Lastly I would like to thank my family, particularly Christian, Leone and Lana for their never ending support, assistance and motivation; it means the world to me.

ABSTRACT

Faecal pollution of water has long been an issue of great concern due to the potential health risks associated with faeces. A lack of understanding the implications of sewage contamination in water combined with inadequate municipal management practices contribute toward deterioration of water infrastructure. Poor management practices, particularly in developing countries, combined with limited financial resources restrict the extent of water quality monitoring which takes place, especially in rural areas. Recent large-scale sewage contamination of the Vaal River in South Africa is an example of where microbial source tracking (MST) by means of *Bacteroides* host-associated assays could be applied advantageously in mitigation efforts. *Bacteroides* species make up the majority of mammalian intestinal bacterial communities. The intestinal community composition of hosts may however still vary regionally and per individual due to factors such as diet. Unlike conventionally applied faecal indicator bacteria (FIB), *Bacteroides* has very limited survival capabilities in the environment due to their anaerobic nature. Subsequently by detection of the host source and quantification of marker levels the extent of contamination as well as possible points of entry into an area of interest may be determined. Unfortunately, as with any developing technique, increased popularity and frequency of application have brought several shortcomings of *Bacteroides* related assays to light. A systematic literature review was performed identifying recurring themes with regard to challenges and limitations faced by researchers applying *Bacteroides* related assays to environmental water samples. It is evident from the literature that a lack of marker specificity both by host species and geographically hinders the application of these techniques worldwide. Markers that were thought of as host-specific were frequently detected in other animal host groups in the reviewed literature. Another shortcoming of these techniques is sensitivity to inhibitory substances commonly found in faeces and environmental water. Data interpretation according to different parameters such as a set sample volume regardless of DNA concentration or DNQ (detected but not quantified) samples in quantitative PCR assays being seen as either positive or negative may complicate data comparability between datasets or across similar studies. Cost reduction of these techniques will increase the opportunities for application in developing countries. These methods are best applied in a toolbox approach along with several other assays or markers in order to estimate the full extent of contamination in a sampling area. Standardisation of *Bacteroides* host-associated assays is crucial for successful application of these techniques, especially with regards to comparability of data.

Keywords: *Bacteroides*, HF183, eDNA, Environmental water, Challenges and limitations

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

AE buffer	Tris EDTA (TE) elution buffer
ANC	African national congress
CCE	Cell calibrator equivalents
CTAB	Cetyl trimethylammonium bromide
DA	Democratic alliance
ddPCR	Droplet digital polymerase chain reaction
DMSO	Dimethylsufoxide
DNA	Deoxyribonucleic acid
DNQ	Detected but not quantified
DO	Dissolved oxygen
EDS	Ebsco discovery service
ELISA	enzyme-linked immunosorbent assays
FIB	Faecal indicator bacteria
FISH	Fluorescence in situ hybridization
GDR	Green drop report
GI	Gastro intestinal
GITC	Guanidinium isothiocyanate-phenol
HPyVs	Human polyomaviruses
IBDG	Indoxyl-β-D-glucuronide
LH	Length heterogeneity
LLOD	Lower limit of detection
LOD	Limit of detection
ML	Megalitres
MST	Microbiological source tracking
MUG	4-Methylumbelliferyl-β-D-galactopyranoside
NGO	Non-government organisation
NWRS	National water resource strategy
OFM	Orange Free State radio (station name)
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PICO	Problem, indicator, comparison, outcome
PMA	Propidium monoazide
PRISMA	Preferred reporting items for systematic reviews and meta-analyses
qPCR	quantitative polymerase chain reaction
rDNA	recombinant DNA
rRNA	Ribosomal ribonucleic acid
RT-PCR	Real-time polymerase chain reaction

SANS	South African national standard
SAHRC	South African human rights commission
SDS	Sodium dodecyl sulphate
SIBU	Sannieshof inwoners belastingbetalers unie (residents taxpayers union)
SODIS	Solar disinfection
TLM	Tswaing local municipality
T-RFLP	terminal restriction fragment polymorphism
UQL	Upper quantification limit
UV	Ultraviolet radiation
VBNC	Viable But Non-Culturable
WHO	World health organisation
WWTP	Wastewater treatment plant
x-gal	5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside
%CV	Coefficient of variance percentages

CHAPTER 1 – PROBLEM STATEMENT, INTRODUCTION AND LITERATURE REVIEW

1.1 Problem statement

Global as well as local water quality issues have increased the need for development of rapid, cost effective detection, quantification and source tracking methods to aid in mitigation regimes. Faecal indicator bacteria (FIB) play an important role in the detection of faecal pollution in water sources (Sidhu *et al.,* 2012). In recent years, the focus has shifted from merely detecting organisms that can indicate the presence of potential pathogens, to focussing on methods for the detection of the source of contamination in order to facilitate management and prevention of the causes of faecal pollution (Okabe *et al.,* 2007; Ballesté & Blanch, 2010). Microbial source tracking tools have been developed in order to detect, quantify and subsequently trace the sources of faecal pollution that occur near contaminated water bodies (Mcquaig *et al.,* 2012*)*. The holistic ideal is to create the ability to manage the source of pollution in order to avoid contamination, saving on remediation costs and reducing human health risk.

In recent literature pertaining to South African water quality, ongoing sewage contamination of, for example, the Vaal River has impacted the health and safety of residents in surrounding areas, who are dependent on the river for potable, industrial or recreational use (Mackintosh & Colvin, 2003; Phakgadi, 2018). Members of the *Bacteroides* genus, more specifically the *Bacteroides fragilis* group, have been suggested as effective alternative faecal indicators by various researchers (Lee *et al.,* 2011). This is mainly due to the fact that they are gram negative, obligate anaerobic rods that do not form spores and are resistant to bile (Sinton *et al.*, 1998; Wexler, 2007). If faeces is analysed, roughly a third of the weight thereof is comprised of bacteria. If *Bacteroides* makes up 30-40% of the total faecal flora, 10% of the total faecal mass provides a target for the identification of faecal contamination in the environment (Layton *et al.,* 2006). Development of host-specific molecular assays making use of *Bacteroides* markers have paved the way for inclusion of these techniques in source tracking and pollution monitoring efforts (Shanks *et al.*, 2010; Nshimyimana *et al.*, 2017). Although several advantages to using *Bacteroides* PCR assays for water quality monitoring remain, it has recently become evident that researchers have faced several challenges whilst applying these techniques (Stewart *et al.*, 2013; Cao *et al.*, 2015). Limitations of *Bacteroides* related PCR assays will be highlighted and discussed by means of a systematic literature review in order to indicate whether the use thereof would be advantageous for water management in developing countries.

1.2 Aim and objectives

The aim of this research was to review is to evaluate the application of *Bacteroides* as an indicator of water quality and to systematically evaluate the literature for challenges with regard to detection, specifically when molecular methods are used.

Objectives:

1. To provide a literature overview of water challenges and why alternative faecal indicators are needed.
2. To provide an overview of *Bacteroides* spp. and why they are regarded as a suitable alternative to classical faecal indicator bacterial parameters.
3. To conduct a systematic literature review of the challenges associated with using *Bacteroides* spp., particularly DNA based methodologies.
4. To provide an overview and recommendations of inclusion (or not) of *Bacteroides* spp. into approaches used to study faecal pollution in developing countries.

1.3 Introduction and Literature Review

1.3.1 The importance of water and water quality

Water is a scarce resource and is not evenly distributed over the earth's surface. Projections made by Gundry *et al.* (2004) indicated that 48 percent of global population growth would occur in areas that already are, or will in future be experiencing water stress. It is currently estimated that more than of the world's population (which has quadrupled over the span of the past century) live in urban areas which are, or may soon be subjected to water stress (Wada *et al.,* 2017). Schlosser *et al.* (2014) stated that social and economic impacts paired with climate change could result in an additional 1.8 billion people living in areas with limited water resources due to overexploitation by the year 2050. The observation was also made that 80% of the population that would be living under water-stress conditions would be residents of developing countries (Schlosser *et al.,* 2014). This causes great concern due to the social and economic implications population growth and development would have in water scarce areas (Gundry *et al.*, 2004). Although there are various rainfall regions and annual rainfall patterns over the expanse of South Africa, the country is situated in a part of the world that is predominantly semi-arid (Rouault & Richard., 2003). Water, worldwide, is used for a vast range of social and economic activities on a daily basis (Roeger & Tavares, 2018). Proper water quality is therefore of great importance to ensure the health and safety of the consumers thereof (D'Inverno *et al.,* 2018). Contamination of aquatic environments may be detrimental to

their use in activities such as: contact recreation, fishing and shell fishing as well as household and potable use of water from specific sources (Sinton *et al.,* 1998; Ebdon *et al.,* 2007; Nnane *et al.,* 2011; Sidhu *et al.*, 2012).

In less economically developed countries, the management of faecal pollution may be especially challenging due to public resources being inadequate, and reliable information on the extent, sources, risks and severity of faecal pollution being very limited (Nnane *et al.,* 2011). Potential health risks caused by exposure to pathogens via faecal contamination should effectively result in low public acceptability of reduced water quality (Ahmed *et al.*, 2010; Edokpayi *et al.,* 2018; Getachew *et al.,* 2018). In reality many urban areas have inadequate infrastructure for the treatment of sewage, and in many rural areas, residential sewage is merely pumped into rivers and streams untreated (McLellan & Eren, 2014). These rivers and streams may also be impacted by agricultural runoff, resulting in a mixture of human and animal sources of faecal pollution evident in environmental water sources (McLellan & Eren, 2014). Failing global water related infrastructure poses a major problem with regard to water quality. Contributions to water contamination by means of leaking sewer lines, failing septic tanks and ageing stormwater drainage systems have been previously reported (Sidhu *et al.*, 2012). Layton *et al.* (2013) stressed the importance of municipalities investing in proper maintenance and management of sewage infrastructure. Limited management of wastewater infrastructure in densely populated areas of developing countries exacerbate the effects of contamination via these sources (Van De Werfhorst *et al.,* 2011; Nshimyimana *et al.*, 2017).

Studies related to water quality and consumption in developing countries have come to the conclusion that unaccounted-for water exacerbates issues related to the availability and distribution of safe drinking water (Lee & Schwab, 2005; Kumar, 2010). Unaccounted-for, or non-revenue water is the term used for the amount of water flowing in a distribution system that does not reach the consumer due to leakages within the system, unmetered use of water because of poor maintenance, or lack of meter registration as well as illegal connections to the system (Hurtgen, 1931; Lee & Schwab, 2005; González-Gómez *et al.,* 2011; Jang & Choi, 2017).

Urbanisation plays a major role in changes relating to the supply and demand of water. Schlosser *et al.* (2014) reported a fourfold increase in non-agricultural water demand in Africa, which was double that of other developing countries. According to Hudson (1964) the rise in population in areas that became more urbanised increased the per capita usage of water. With increasing development of urban and suburban areas, less land became available for the retention of surface runoff and less water gets absorbed into the underground supply (Hudson,

1964). Urbanisation's effects on water quality have increased as population size increased over the years, with little of the impact thereof being considered in hydrological models for future water use and availability predictions (Wada *et al.,* 2017). Surface as well as underground supplies of water were more readily polluted due to the increase in population in urban and suburban areas. These factors often lead to the depletion of an area's current water supply, creating the necessity to obtain a new supply, most often costing more than the original (Hudson, 1964; Kumar*,* 2010; Zheng *et al.,* 2013).

1.3.2 Reality regarding water quality and faecal pollution in Africa

The quality of water that is collected and stored in rural areas where access to piped water is lacking, may deteriorate due to the use of dirty containers or unhygienic handling of stored water e.g. dipping hands in the container to scoop water out or using an open topped storage vessels (Jensen *et al.,* 2002; Tumwine *et al.,* 2002; Gundry *et al.,* 2004; Lee & Schwab, 2005; Gorham *et al.,* 2017; Onyango *et al.,* 2018). In areas with intermittent water distribution more wastage occurs as households attempt to collect large quantities of water when it is supplied. In 1976 only 3% of East African households that had access to piped water stored water for later use, however that number increased to 90% by the year 2000 (Lee & Schwab, 2005). Getachew *et al.* (2018) reported a decrease in the percentage of the Ethiopian population with access to unsafe drinking water by 11%. Even with this decrease it was found that only 43.5% of households participating in the study had coliform counts below the WHO classification limit (Getachew *et al*., 2018).

An effective, extremely low cost disinfection alternative that may be used in rural areas is solar disinfection (SODIS). According to Conroy *et al.* (2001) the simplest form of solar inactivation was filling clear plastic or glass bottles and leaving them in direct sunlight until they reach a temperature of 40ºC and up. Reaching the correct temperature for disinfection is estimated at a minimum of 6 hours (Figueredo-Fernández *et al.,* 2017). Bottles could be placed on a dark background or be half blackened with paint in order to increase the heat absorption potential of the water in the bottles. Both heat and UV radiation can play a role in inactivation of bacteria in water. It has been reported that solar exposure may inactivate *E. coli*, *Salmonella enterica* serotypes *typhi, paratyphi* and *enteritidis, Shigella flexneri*, *Pseudomonas aeruginosa*, *Enterococcus faecalis,* as well as *Vibrio cholera* and various other viruses (Conroy *et al.,* 2001; Rojko, 2003). In a trial experiment 206 children aged 5-16 years old were given solar disinfected water to drink. A ten percent reduction in the incidence of diarrhoea was observed as well as a 24% decrease in the occurrence of severe diarrhoea over the twelve week study period (Conroy *et al.,* 2001). Although this disinfection method aided in the reduction of

bacteria present in water, there were some shortcomings in this technique, making it reliable but not one hundred percent effective (Conroy *et al.,* 2001; Rojko, 2003). If the correct temperatures are not reached, the sunlight is not strong enough and the water is not left out long enough before use, partial inactivation of microorganisms may take place (Conroy *et al.,* 2001; Rojko, 2003). Viruses do not possess the capability to perform photo-repair, but bacteria may correct the damage UV radiation has caused to their DNA via natural DNA repair mechanisms in the cells (Conroy *et al.,* 2001; Rojko, 2003). Spore forming organisms have the increased potential to survive extreme environmental conditions and may not be affected by solar inactivation. Lastly, turbidity of the water may hinder the penetration of sunlight into the water. Even though the particles in turbid water may help reach higher temperatures, the heat generally does not penetrate the entire volume of water, effectively shielding organisms from the effects of UV inactivation (Conroy *et al.,* 2001; Rojko, 2003).

Solar disinfection (SODIS) projects have been applied in rural African communities by various studies (Figueredo-Fernández *et al.,* 2017; Wang *et al*., 2017; Mac Mahon & Gill, 2018). An estimated 4.5 million individuals, mainly in developing countries, are making use of solar disinfection for their household water on a regular basis (Mac Mahon & Gill, 2018). Jadhav *et al.* (2017) and Wang *et al.,* (2017) noted the advantages of making use of solar energy as well as solar disinfection in African countries such as Zimbabwe and Ghana. Various factors such as the contamination level and physical-chemical composition of the water, environmental conditions and the type of SODIS container being used affect the efficiency of disinfection by this method (Figueredo-Fernández *et al.,* 2017). Samples were tested in Morocco where plastic bottles were replaced with plastic polymer bags as this is the most affordable improvement to be made to SODIS techniques. Other options for improved SODIS application include using light reflectors to increase the amount of UV exposure as well as making use of broadband semiconductors to accelerate the disinfection reaction when exposed to light (Figueredo-Fernández *et al.,* 2017). Results of the study indicate that simultaneous heat and light exposure increased the disinfection efficiency of SODIS techniques. Polymer bags increase the surface area exposed to UV radiation making them twice as effective as plastic bottles. Figueredo-Fernández *et al.,* (2017) concluded that SODIS techniques are an effective technique for creation of safe drinking water in areas with similar climate and UV exposure as Morocco. Mac Mahon and Gill (2018) reported several difficulties during the implementation of a novel water disinfection system in Kenya. Important aspects that Mac Mahon & Gill (2018) highlighted with regard to the building and implementation of alternative water disinfection systems, were the importance of funding and support for such systems by humanitarian and development agencies, community involvement in cleaning and maintenance of the system,

as well as careful consideration toward the position of the system with regard to water access in times of drought and avoiding damage during floods.

In many instances where intermittent and/or poor quality water was supplied worldwide, consumers were reluctant to pay for such services (Hudson, 1964; Lee & Schwab, 2005; Rusca *et al.,* 2017). Poorer residents of Malawi's capital city Lilongwe have raised complaints regarding unfair water supply to richer communities and high costs of water service delivery deterring poorer communities from using safe water supplies (Rusca *et al*., 2017).

According to Obiri-Danso *et al.* (2003) there was a drastic increase in the consumption of bottled or bagged water in Ghana between 1993 and 2003. The main driving forces behind the increase were: the consumption of designer water having become more popular, increased concern about the safety of piped water supplies in the country, as well as the presence of more people within major shopping areas that required high quality drinking water (Obiri-Danso *et al.,* 2003). Plastic bagged water was the most popular option in the city of Kumasi. Although many commercial manufacturers used multi-candle pressure filters to remove sand, rust, metal sediments, algal biofilms as well as bacteria from water via activated carbon, many hand-filled bagged waters were also available on the market (Obiri-Danso *et al.*, 2003). Depending on the hygiene practices within the factories some bottled and bagged waters still contained bacteria after production. Bottled water contained total viable counts of bacteria ranging between 1 and 4670 ml^{-1}, however no total or faecal coliforms were present (Obiri-Danso *et al.*, 2003). Factory bagged water was heat sealed into the sachets, but even so, ten samples contained total viable counts of heterotrophic bacteria between 2 and 6.33 x 10^5 ml^{-1}. Four of the eighty-eight samples tested contained total coliforms and two contained faecal coliforms (Obiri-Danso *et al.,* 2003). Lastly, due to the low start-up cost of becoming a bagged water vendor, many individuals in poverty stricken areas that did not have the knowledge or means to implement basic human hygiene practices, managed to sell their products in local markets across Ghana. Hand-filled bagged water ran the risk of becoming contaminated at all steps in the process, as bags were inflated by blowing air into them, filtration took place via unsterilized foam being attached to a hosepipe, merely filtering out large particles such as sand. No other method of sterilisation of the water took place, and the bags were tied by hand, possibly contributing to contamination of the water within (Obiri-Danso *et al*, 2003). Ten selected samples of hand-filled, hand-tied water contained heterotrophic plate count bacteria at numbers ranging from 2.23 x10^3 to 7.33x10^{12} ml^{-1}. Seventeen of the forty hand-filled water bags that were tested were positive for large numbers of total coliforms and nine were positive for faecal coliforms. Two samples were also positive for *Enterococci* (Obiri-Danso *et al.*, 2003). It is evident from the tests performed on the water obtained from various vendors in the city of

Kumasi in Ghana, that although the quality of bottled or commercially bagged water is higher than that of the tap water in the area, low-end hand filled water bags and manufacturers of commercial brand bagged water who do not practise disinfection steps and do not make use of aseptic bagging techniques have the potential to spread waterborne pathogens via handling or selling of low quality, bagged tap water (Obiri-Danso *et al.*, 2003).

1.3.3 Reality regarding water quality and faecal pollution in South Africa

In Africa and South Africa the natural availability of water in several river-basins is far exceeded by the demand for water in these areas (NWRS, 2013). According to the World Bank (Lee & Schwab, 2005) the estimated average rate of unaccounted-for water in developing countries was between 37 and 41 percent. South African levels of non-revenue water were recently estimated at 36.8% (Mckenzie *et al*., 2012). Nationally, one third of water supplied in South Africa is lost as non-revenue water annually (Mckenzie *et al.,* 2012). The occurrence of unmetered or illegal water connections made it difficult for municipalities to determine the true relationship between supply and demand of water within their districts (Mackintosh & Colvin, 2003 Lee & Schwab, 2005; González-Gómez *et al.,* 2011). Due to the fact that illegal users were not held accountable, financially or otherwise, it was apparent that there was no incentive for conservation of the resource among them (Lee & Schwab, 2005; Gouws *et al.,* 2010; González-Gómez *et al.,* 2011).

Water quality may be influenced at numerous stages of use, and poor management may lead to water not being up to standard before distribution (Mckenzie *et al.,* 2012; Roeger & Tavares, 2018). Water may also be contaminated or compromised during distribution due to contamination within the distribution system as a result of natural ageing of the system, corrosion, pressure changes due to intermittent supply, and deterioration of general infrastructure (Lee & Schwab, 2005; González-Gómez *et al.,* 2011; Rusca *et al.,* 2017). It is generally accepted that piped water is of high quality, yet chemical and biological deterioration of water quality may take place within distribution systems (Chalchisa *et al.,* 2017). Edokpayi *et al.,* (2018) estimated that 2.11 million South African residents still lack access to safe water infrastructure. Lee & Schwab (2005) reported that 14–16% of water samples obtained from households in Johannesburg, South Africa contained coliform counts greater than that of samples taken at treatment plants directly after chlorination. Even so, this water was of relatively high quality. For example, in Pietermaritzburg, South Africa, inadequate chlorination lead to the presence of coliforms within the distribution system in increasing numbers as the distance from the water plant increased (Lee & Schwab, 2005). Through this, the importance of proper maintenance at the plant was highlighted, including how crucial knowledge regarding

the capacity and distance of the distribution system was to the supply of high quality drinking water (Samir *et al.,* 2017; Potgieter *et al.,* 2018; Roeger *&* Tavares*,* 2018)

In two separate incidents, illegal water connections in the Free State province of South Africa came to light in popular news. According to Refilwe Mochoari, reporting for the Vrystaat News page on the Facebook social media platform, it was found that an illegal water connection had been set up by the mayor of Clocolan, Thediso Jakobo, in 2012. During this time water shortages in the Clocolan, Senekal and Marquad areas were evident as the Cyferfontein, De-put as well as Moperi dams in the area had all run dry. After public outcry, an investigation was launched into the accusations that illegal water was being used whilst the townsfolk had poor quality or no water at all. Illegal water was being supplied from a municipal feed to a kraal containing fifty pigs belonging to the mayor. In the Mandelapark suburb of Clocolan, illegal water connections were also found in the homes of several of the mayor's friends. Unfortunately no further comment or outcomes of the investigation were published (Mochoari, 2012). In December 2017 the Democratic Alliance (DA) political party laid charges against an African National Congress (ANC) councillor named Jerry Moitse for illegal water connections running to his home in the town of Fauresmith in the Free State Province. As reported by Josca Human on OFM Radio (2017) DA spokesperson James Letuka stated that Fauresmith residents had been without water for two months due to the local municipality failing to pay their supplier, Bloem Water. After thorough investigation four houses were found to have illegal connections running straight from the reservoir to these houses. Councillor Moitse alleged that the pipelines were laid legally by the Kopanong District municipality to his and six other families' homes and that legal action would be taken against the DA regarding the claims they have made regarding the water supply (Human, 2017).

In the Swartland district of the Western Cape audits on water losses brought several illegal water connections to light. The municipality charged land owners R1 500 tampering fee as well as R 5 000 fee for damages to municipal property and took legal action against those who did not comply after an amnesty period was given in which residents could check the legality of their water connections and have them converted to legal connections accompanied by the installation of water meters (Anon 3, 2017). Arrests have been made in Johannesburg on two separate accounts of fraud and corruption linked to illegal water connections. In February 2018, a resident of the Westedene suburb reported a man fraudulently claiming to be a chairperson of Johannesburg Water to the authorities after the man attempted to solicit bribes, claiming that the payments from the resident would ensure water supply to his property would not be cut off due to late payment (Anon 4, 2018). Two property developers had been arrested in Woodmead after investigations revealed that 90% of water meters in the Waterfall City

development were installed illegally and were not linked to the Johannesburg Water billing system (Bornman, 2018). It was estimated that an annual 5 to 8 billion rand is lost to revenue leakages and electricity and water accounts being deleted off of the city billing system (Bornman, 2018).

In an investigation by Mackintosh and Colvin (2003) the water supply in the poorer, more rural Eastern and relatively urbanised, wealthier Western Cape provinces of South Africa were compared. Five of seven groundwater supply schemes in the Eastern Cape were not functional at the time of the investigation, as opposed to all being functional in the Western Cape (Mackintosh & Colvin, 2003. Due to the high turbidity of surface water in the Eastern Cape, all schemes visited practiced additional flocculation and settling steps during water purification. A high number of schemes in both provinces had no disinfection capabilities (Mackintosh & Colvin, 2003). With regards to the microbiological water quality of drinking water provided to these provinces, approximately 50% failed the maximum allowable limit for heterotrophic plate counts in Eastern Cape where 62% failed in the Western Cape (Mackintosh & Colvin, 2003).

The Save the Vaal Project is aimed at protection and maintenance of the Vaal River between the Vaal dam and the town of Parys. The river is the main source of water for the Witwatersrand area and water from the river is applied for commercial, agricultural and personal use (Anon 1, 2013). Due to ongoing discharge of industrial effluent, mine and municipal wastewater into the river, the quality of water within the system has deteriorated to the point of being classified as potentially life-threatening. Communities along the Vaal have complained about water quality since 2008 (Phakgadi, 2018). According to The save the Vaal Website (Anon 1, 2013) the second edition of the NWRS published in 2013 identified numerous threats to South African water resources, but offered no solutions or strategies that could be applied realistically in order to manage and rehabilitate deteriorating and diminishing water supplies. The aims of the Save the Vaal organisation are to: (i) gather information on water quality in the Vaal River system and make sure it is publically available; (ii) be in contact with the department of Water Affairs as well as municipalities whose infrastructure are of poor quality and that are subsequently contributing to the pollution of the Vaal River; (iii) raising community awareness in order to alert residents in the vicinity of the Vaal River to the quality of their water; (iv) take legal action such as revisiting court orders obtained against municipalities who are not being held accountable for their contribution to pollution of the Vaal River as well as (v) taking necessary steps in order to assist in the prevention of further pollution of the river (Anon 1, 2013). Even though the Save the Vaal project was launched in order to assist in restoration of water quality, recent news indicates further deterioration of

water quality in the area. The South African Human Rights Commission (SAHRC) was quoted in the media stating that the sewage contamination is suspected to originate from WWTPs which have been managed poorly by the Emfuleni municipality (Anon 2, 2018). Allegations state that an estimated 150 ML of untreated sewage is spilling into the river on a daily basis. A site inspection was conducted in September 2018 following these allegations after which a statement was made by the SAHRC that the contamination in the river violates human rights of access to clean environment, clean water and human dignity (Pijoos, 2018). The inquiry by the SAHRC aims to elucidate whether the government failed to protect communities from exposure to sewage contamination. Investigations into whether impacts of short and long term exposure to contamination were taken into consideration form part of the enquiry (Phakgadi, 2018).

In 2007 the town of Sannieshof in the North-West province declared a dispute with the Tswaing Local municipality as water service delivery in the town was virtually non-existent (Gouws *et al.,* 2010). Ratepayers in the town formed a group (SIBU) that literally took over the functions of the local government in the town. Members of SIBU decided to withhold their municipal rates and taxes as no service delivery was taking place, when they were paying local government to do so. The Tswaing Local Municipality (TLM) took SIBU to court for various infringements including trespassing on municipal property and taxes being in arrears. According to Gouws *et al.* (2010), the municipality felt that the community was interfering in the work of municipal officials, where the community was irate due to the total lack of service delivery. The main cause of the water service crisis in Sannieshof was the lack of management of infrastructure as the wastewater treatment plant had not seen any upkeep in ten years and the town's reservoirs allegedly hadn't been cleaned in twenty (Gouws *et al.,* 2010). In addition to the poor infrastructure a major influx in residents, specifically in the poorer township areas due to governmental promises of free housing, exacerbated the problem as the system was never built to deal with the extra capacity. Furthermore the informal settlement area (Phelindaba) of the township (Agisanang) had intermittent and/or little to no water and sanitation services whatsoever, forcing residents to use pit latrines, buckets and shallow holes in the open veld as toilets. Along with overflow from the failing wastewater treatment plant, storm water runoff would carry raw sewage directly into the Harts River. Gouws *et al.* (2010) stated that raw sewage was also being dumped directly into the river causing both major environmental and health concerns.

These examples demonstrate that faecal pollution of water is a real problem in South Africa, that could however be addressed by implementing suitable solutions. It also underscores the importance of applying appropriate monitoring mechanisms and methods.

1.3.4 South African Blue and Green Drop water status

The Blue drop report is a governmental certification programme in South Africa that indicates the national microbiological compliance of South African tap water as measured against the National Standard (SANS, 2015). Green Drop certification indicates the quality of wastewater and is meant as an initiative to aid in the progressive improvement of wastewater quality, in order to decrease the environmental impact of wastewater being discharged into natural water bodies for recirculation (Burges; N.D.). These are regulatory programmes that assist in the assessment of water quality management in South Africa (Burges, N.D.). According to the 2012 Blue Drop Report the North West province of South Africa had an overall score of 78.7% with Tlokwe City Council as the best performing municipality with a score of 98.45%. Certain districts had continual high scores due to proper management being implemented and infrastructure maintained. Improvements in infrastructure showed significant annual increases in the score of other developing districts. Unfortunately certain areas have had continual low scores due to a lack of infrastructure. Low scoring areas therefore have poorer water quality. However, environmental conditions such as floods and droughts may also influence the scores of districts with poorer infrastructure. In the 2014 report the overall score of the North West province was brought down to 63% by districts with poor performance within the province (GDR, 2011).

In order to obtain Green Drop Certification a score of 90% or higher has to be obtained against the assessment requirements laid out in a standardised scorecard format by the Department of Water Affairs. A green drop score is obtained for each wastewater treatment system assessed, and a cumulative risk rating is calculated after the design capacity of the specific plant as well as all the plants within the municipal system are assessed. A stringent site inspection takes place during which the physical condition of the plant is critically evaluated (Burges, N.D.). Only if all the criteria are met and the score is above 90% will a municipality be granted green drop certification. The principle of the Green Drop strategy is that municipalities that are performing poorly should be identified via customers, the media, political classes and Non-Governmental Organisations (NGOs) in order to make them aware of poor performance via public outcry. They should consequently identify the shortcomings in their system and correct the issues being addressed in order to be “rewarded” for excellence, rather than penalised for failures (Burges, N.D.).

In the 2011 Green Drop Report, North West province scored an overall 50% with Tlokwe municipality having had the best performance score of 97% (GDR, 2011). In the 2013

Executive Summary of the Green Drop Report it was stated that 21 of the 35 wastewater treatment plants in North West Province had Green drop scores of below 30%.

In a project launched by Afriforum (South African NGO supporting minority groups), basic water quality parameters have been tested twice annually in all nine provinces in order to create community based awareness of the safety of water in each province, with the goal of imploring local municipalities to improve water quality and service quality in affected areas (Pawson & Boshoff, 2017). An Afriforum Blue and Green drop Report that compares the average of the two annual samples for each sampling site across the provinces is available. In the 2017 report recurring problems with drinking water quality have been indicated in Stella, Coligny, Witbank and Hertzogville (Pawson & Boshoff, 2017). With regard to wastewater treatment, sites that had recurring problems or that were tested for the first time in 2017 and failed are as follows: 13 of 21 in Gauteng; 7 of 16 in Western Cape; 3 of 11 in Northern Cape; 7 of 8 in Eastern Cape; 6 of 12 in Free State; 11 of 18 in Mpumalanga; 7 of 17 in North West; 6 of 11 in Limpopo and 3 of 8 in Kwa-Zulu Natal. As can been seen from these results it is clear that there are significant issues surrounding the infrastructure and management of wastewater treatment plants and the quality of water released into the environment, compared to drinking water quality in South Africa (Pawson & Boshoff, 2017).

1.3.5 Indicator organisms and potential human health risk

In a study by Kurokawa *et al.* (2007) gut microbes from 13 healthy Japanese individuals were compared to one another, as well as to samples from two American adults, via comparative metagenomic analysis. High inter- individual variation in taxonomic and gene composition of a relatively simple intestinal flora system was found in unweaned infants, where adults and children (weaned), had a more complex intestinal flora composition, which showed high functional uniformity regardless of the individuals' age or sex (Kurokawa *et al.*, 2007). Individuals taking part in this study had no dietary restrictions beyond avoiding antibiotics, pro-biotics and fermented foods for four weeks prior to sampling, and none of the individuals had a history of gastrointestinal disorders or unusual eating habits (Kurokawa *et al.,* 2007). In adults, the major constituents of the gut microbiota were always *Bacteroides* followed by genera belonging to the Firmicutes division including *Eubacterium*, *Ruminococcus* and *Clostridium* as well as the genus *Bifidobacterium*. In infants, *Bifidobacterium* and/or a few genera from the *Enterobacteriaceae* family such as *Escherichia*, *Raoultella* and *Klebsiella* were the major intestinal flora constituents in the gastro-intestinal tract (Kurokawa *et al.,* 2007). According to Kurokawa *et al.* (2007) it should be noted that, although data from two individuals isn't sufficient to understand the structure and functional capabilities of gut microbiomes and

the intrinsic and environmental factors that may affect them, there was a significant difference in the composition of the microbial communities between Japanese and American adults.

Faecal indicator organisms that reside in the gastrointestinal tract of mammals are used to assess the microbial quality of water for drinking and recreational use. Bacterial water quality standards are aimed at reducing health risks associated with water. Standard faecal indicator bacteria (FIB) that are monitored include: total coliforms, faecal coliforms, *E. coli*, as well as *Enterococci* (Ahmed *et al.*, 2008[A]; Schriewer *et al.*, 2010; Raith *et al.*, 2013; McLellan and Eren, 2014). Culture-based methods making use of *E.coli* and *Enterococci* are generally applied as the "traditional" method of water quality management due to low cost and ease of use (Lee *et al.,* 2011; Sauer *et al.*, 2011). Although culture methods are relatively inexpensive, the standard 24 to 48 hour incubation period, lack of specificity, as well as the poor detection of slow growing or viable but non-culturable (VBNC) microorganisms limit the success of culture techniques as an effective method of indicating recent faecal pollution, particularly from a specific source (Rompré *et al.*, 2002; Savichtcheva & Okabe, 2006; Lee *et al.,* 2011). Other shortcomings of *E. coli* as an indicator of faecal pollution include genetic diversity within the species which may cause false negatives if strains do not ferment lactose. Oshiro (2002) and Layton *et al.* (2006) further noted the occurrence of false positives due to the presence of *E. coli* from non-faecal sources. Additionally, colloidal or suspended particulate matter may clog membrane filters, preventing proper filtration of water samples prior to incubation on selective media (Oshiro, 2002; Layton *et al.*, 2006). Contamination may yield false positive results as well. Certain other bacterial species such as the *Corynebacterium, Clostridia* and certain members of the *Streptococci* and some *Bacteroides* strains also produce beta-glucuronidase enzymes which could lead to overestimation of the presence of *E. coli* (Nakamura *et al.*, 2002; Hall *et al.*, 2003; Stringer *et al.*, 2008; Pollet *et al.*, 2017). Further limitations of using standard FIB to represent potential pathogens in water include their ability to multiply in the environment and fact that the absence of FIB does not necessarily mean that there are no pathogens present in the water (Ahmed *et al.*, 2008[A]; Sauer *et al.*, 2011; Sidhu *et al.*, 2012; Toledo-Hernandez *et al.*, 2013). Furthermore FIB generally do not distinguish between sources of contamination, and thus potential health risks due to the presence of human-specific pathogens may be grossly underestimated (Nnane *et al.,* 2011; Sidhu *et al.*, 2012; Ridley *et al.*, 2014).

Four basic approaches are used when investigating the possibility of using a microbe as an indicator of a source of faecal pollution. These approaches are: Speciation- certain microbial species are linked to a specific host or source, Biochemical reactions- sources can be differentiated from one another due to speciation and the resulting biochemical differences

between microbes, Assemblages and ratios- the type of microbes as well as amount of microbes present may shed light on the source of faecal contamination and DNA profiles- generally genetic differentiation of species is more reliable than the previously mentioned phenotypic methods (Sinton *et al.,* 1998).

Assemblages and ratios, although more effective than coliform counts, may still be unreliable as most faecal indicators including *E.coli*, enterococci, coliforms and faecal coliforms are present in the intestinal tract of both humans and animals and are thus not effective source differentiators (Sinton *et al*., 1998). Faecal streptococci–faecal coliform ratios have been used to distinguish between human and animal sources of faecal pollution as streptococci are more abundant in animals where coliforms are more commonly found in humans. The dispersal of faeces into the environment may however change the ratio and give false results due to environmental factors affecting the survival and detection rates of these species differently (Kreader, 1995). Before an organism or a specific gene can be considered as a source tracking marker, it has to be evaluated to determine whether it is specific to one host group. In other words: the ability of an assay to exclude non-target faeces (host specificity) and whether it is present in all members within the target host group (host sensitivity) (Ridley *et al.*, 2014). Furthermore source tracking markers should preferably be geographically and temporally stable within the host group and the decay rate of these markers should be similar to that of the pathogens it is used to detect (Green *et al.*, 2011).

The ideal for FIB is that: they should be exclusively associated with the intestinal tract of humans or other specified mammals, they should not survive in the environment for extended periods of time, and they should be associated with the occurrence of human enteric pathogens (Toledo-Hernandez *et al.*, 2013). Additionally FIB should ideally be present in high concentrations in polluted sources and in much lower concentrations or absent in unpolluted environments (Ahmed *et al.*, 2008[B]; Schriewer *et al.*, 2010).

The health risk posed by faecal contamination of water can only be accurately evaluated if the source of the pollution can be determined (Shanks *et al.*, 2010; Newton *et al.*, 2011). The prevention of outbreaks of water-borne diseases such as cholera in rural communities may be achieved by improved management of water sources and supply, as well as being able to quickly and effectively identify the source of pollution in order to remedy the problem. The use of faecal bacteria to determine a host animal source is based on the assumption that certain strains of bacteria are associated with specific host species, and that they can be differentiated from one another by means of genotypic and phenotypic markers (Layton *et al.*, 2006). Making use of *E. coli* and other “traditional” indicator species as a source identifier proves to be

problematic due to the previously mentioned shortcomings as well as geographic and temporal variability within host and bacterial species (Layton *et al.*, 2006; Okabe *et al.*, 2007; Ahmed *et al.*, 2008[B]; Newton *et al.*, 2011). Identifying a source of pollution is therefore rather difficult and outbreaks or incidences of disease may continually affect communities where alternative faecal indicators are not used to identify the source of faecal contamination within a water source or system (Newton *et al.*, 2011).

One of the most commonly used detection methods, especially for the detection of coliforms, is the membrane filtration method in conjunction with chromogenic and fluorogenic identification (Edberg & Edberg, 1988; Oshiro, 2002; Rompré *et al.,* 2001). The basis of identification in this test is specific enzymatic activity due to most enteric bacteria (such as *E.coli*) possessing β-Galactosidase. This enzyme is critical for hydrolysis of the disaccharide Lactose into easier to process monosaccharides: Glucose and Galactose (Henne & Karcher, 2007). Fluorogenic identification makes use of agar containing a 4-methylumbelliferyl-β-D glucuronide enzyme substrate (MUG agar). During hydrolysis of this substrate, methylumbelliferone (a strongly fluorescent compound) is liberated and the amount of free methylumbelliferone (and thus the amount of substrate hydrolysed) may be quantitatively analysed via a flourimeter (Dahlén & Linde, 1973; Oshiro, 2002). Chromogenic analyses make use of agar containing 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside (X-gal), a modified galactose sugar, as well as an activator called isopropyl- β -D-thiogalactopyranoside (IPTG). When microbes process these sugars by means of galactosidase enzymes, the product they produce is blue (Tortora *et al., 2010*). MI agar may also be used for a combined fluorogenic/ chromogenic analysis. This medium contains both 4-Methylumbelliferyl-β-D-galactopyranoside (MUG) and Indoxyl-β-D-glucuronide (IBDG) substrates (Oshiro, 2002). After water has been filtered through a nitrocellulose membrane filter, it is placed on the prepared agar medium and incubated for 24 hours at 37°C. The presence of *E. coli* is indicated by the presence of blue colonies where the breakdown of MUG by total coliforms is indicated by fluorescence under long-wave ultraviolet light (Oshiro, 2002).

Alternative detection techniques as well as alternative indicator organisms are continually being researched in order to create inexpensive, accurate, and rapid methods for detection of faecal pollution in water sources. An alternative, relatively simple and inexpensive detection technique for faecal pollution, specifically from human sources, is the use of bacteriophages (Jofre *et al.*, 2014). The host range for *Bacteroides* species may not be restricted to humans, but it seems that the bacteriophages that infect these bacterial strains are almost exclusively present in human faeces (Ebdon *et al.,* 2007).According to Ebdon *et al.* (2007) and Ervin *et al.* (2013) the determination of the origin of faecal contamination, whether human or animal,

is essential for the estimation and assessment of public health risk. The facilitation of remediation methods and treatment to affected communities or individuals, as well as resolving the legal implications and responsibility for remediation efforts, also rely on source identification (Ebdon *et al.,* 2007; Ervin *et al.*, 2013). Ranges of microbial source tracking tools have been developed by researchers to help distinguish between human and animal sourced faecal pollution (Ahmed *et al.,* 2012). Source tracking markers that are used to detect and quantify faecal pollution in environmental waters include: anaerobic bacterial gene markers, bacterial toxin gene markers as well as viral markers, and certain chemical compounds such as sterols (Shah *et al.*, 2007; Ahmed *et al.,* 2012). Other culture based and molecular methods for analysis of faecal pollution such as antibiotic resistance patterns of faecal streptococci and *E. coli* ribosomal DNA tracking exist, but these methods are labour intensive and rather unreliable as they do not quickly and accurately identify the source of pollution. Molecular markers that do not require culturing of organisms may prove to be more effective (Bernard & Field 2000). More approaches include: Restriction endonuclease analysis -cleavage of DNA at specific sites. PCR-biochemical amplification of specifically selected gene sequences, media for recovery and enumeration of streptococci based on ability to grow in presence of azide, fermentation of carbohydrates to lactic acid and media for enterococci based on ability to hydrolyse complex carbs (esculin) in the presence of bile salts (Sinton *et al.,* 1998). Eukaryotic mitochondrial DNA differentiation of sources of faecal pollution, chemical makers such as caffeine, fragrance substances, and fluorescent whitening agents have also been considered (Jofre *et al.,* 2014).

Molecular methods such as polymerase chain reaction (PCR), *in-situ* hybridisation (FISH), length heterogeneity marker identification (LH), terminal restriction fragment polymorphism (T-RFLP) as well as immunological techniques such as enzyme-linked immunosorbent assays (ELISA) have been evaluated for their usefulness as alternative methods for detecting faecal pollution in water (Ahmed *et al.,* 2008[B]; Rompré *et al.,* 2001; Savichtcheva & Okabe, 2006). The use of computational methods such as oligotyping may give insight into patterns of host association (McLellan & Eren 2014). In recent years the use of DNA recombinant methods pertaining to microbial community structure and function have gained popularity as they offer a rapid method for the comparative analysis of populations and offer a “fingerprint” view of a bacterial population from a specific source or environment. These methods, however, do not identify individual organisms within these populations (Hill *et al.,* 2002).

Individual members of microbial communities may be identified via library-independent methods (Ridley *et al.,* 2014) such as PCR and direct sequencing (for example members of the *Bacteroides* genus, by making use of the Bac32 and Bac708 primer set- Okabe *et.al.*

2007) or the creation of clone libraries and the subsequent sequencing of specific targets within the microbial genome, the most frequent target being the 16S rRNA gene (Hill *et al.,* 2002; McLellan & Eren, 2014). The usefulness of the abundance ratio of human *Bacteroidales* to total *Bacteroidales* has also been discussed by Sauer *et al.* (2011). Abundance ratios were also used by Newton *et al.* (2011) to identify sources of faecal pollution in an urban harbour in Milwaukee, Wisconsin USA. In their study they found that the abundance ratio of human to total *Bacteroidales* was the similar to that of untreated sewage, suggesting that sewage overflow is the main cause of faecal pollution in the harbour (Newton *et al.,* 2011).

Detection of faecal indicator organisms may differ between samples of similar origin in different parts of the world due to differences in human faecal flora caused by diet and other factors (Ebdon *et al.,* 2007), suitable hosts would thus need to be isolated for different geographic regions (Jofre *et al.,* 2014; McLellan & Eren, 2014; Payan *et al.,* 2005). Realistically, it is unlikely for any marker to be absolutely host specific, therefore a specificity of <0.8 or 80% may not be useful for microbial source tracking studies (Oshiro, 2002; Ahmed *et al.,* 2012). Terminology to better distinguish patterns of occurrence of faecal indicators across different species has recently been suggested by McLellan and Eren (2014). Cosmopolitan taxonomic units do not show any preference for a specific host; host–preferred distribution describes organisms that are significantly more abundant in one host species, although it may still occur in other hosts in lower abundances; host-associated taxonomic units occur in only one host species, but not always in all individuals of the host species, and host-specific organisms may be seen as core members of the intestinal biota of a host species as they are present in every sample from that same host species (McLellan & Eren, 2014).

Faecal anaerobes such as *Bacteroides* and *Bifidobacterium* species have been suggested as alternative indicators due to the fact that they make up the majority of flora within the mammalian gastrointestinal tract (Bower *et al.*, 2005; Okabe *et al.*, 2007; Lee *et al.,* 2011; Ridley *et al.*, 2014).

The majority of the intestinal bacterial population is anaerobic or facultatively anaerobic, and of these organisms about 25% is made up *Bacteroides* species (Wexler, 2007). *Bacteroides* are gram negative, obligate anaerobic rods that do not form spores and are resistant to bile (Sinton *et.al.* 1998; Wexler 2007). These organisms may be stimulated by the presence of 20% bile, which is inhibitory to most other anaerobes (Wexler, 2007) and show distinctive growth when cultured on selective media such as *Bacteroides* Bile Esculin agar (BBEA). When faeces in particular is analysed, roughly a third of waste material (by weight) is comprised of bacteria (Layton *et al.,* 2006). Because of their physiological characteristics, these organisms

may be present within the intestinal tract at a hundred to a thousand times greater density than members of the coliform group (*E. coli* in particular), effectively making *Bacteroides* an alternative and quite possibly more sensitive indicator of faecal pollution (Sinton *et al.,* 1998; Bower *et al.,* 2005; Ahmed *et al.,* 2008[B]; Lee *et al.,* 2011; Toledo-Hernandez *et al.,* 2012).

1.3.6 Suitability of *Bacteroides* as alternative faecal indicator

Due to their anaerobic characteristics, *Bacteroides* species should have a low survival rate outside of the gastrointestinal tract (Kreader, 1998; Ahmed *et al.,* 2008[B]; Bell *et al.,* 2009; Toledo-Hernandez *et al.,* 2012). Predictable concentrations of organisms are reportedly present in faecal matter, and human/bovine *Bacteroides* species are seldom detected in faeces of the other group or other animals such as domestic pets. If so, detection in low numbers occurs due to horizontal transfer because of humans and animals coexisting in close proximity to one another (Kreader, 1998; Ahmed *et al.,* 2008[A]; Ahmed *et al.,* 2008[B]; Ahmed *et al.,* 2012). Documented host specificity and geographic stability have led to these assays being incorporated into microbial source tracking studies in Japan and Europe (Bell *et al.,* 2009; Ahmed *et al.,* 2010; Schriewer *et al.,* 2010). Because of predictable concentrations of *Bacteroides* markers being present in faecal matter, quantitative analysis of samples via qPCR is possible (Bell *et al.,* 2009; Ridley *et al.,* 2014).

Although sewage associated *Bacteroides* species are assumed to be host specific and geographically stable, (Ahmed *et al.,* 2010), biotic and abiotic environmental parameters affect their survival and persistence as well as the biochemical degradation of their DNA in the environment. Some of these factors include: protozoan grazing, infection by bacteriophages, temperature fluctuations, solar irradiation as well as sedimentation. Sunlight inactivation may have different effects on microorganisms depending on their oxygen protection mechanisms. Definite correlations between temperature and DNA degradation exist, regardless of species. This may cause problems with detection (misdetection of viable organisms or over/underestimation of risk) (Ballesté & Blanch, 2010). *Bacteroides* in raw sewage have similar survival rates to coliforms, but die off more rapidly in water (Sinton *et al.* 1998). Different members of the *β-fragilis* group react differently to environmental parameters such as temperature and dissolved oxygen content of water. For example: *B. fragilis* persists longer in cooler temperatures and *B. thetataomicron* longer in summer (Ballesté & Blanch, 2010). Dissolved oxygen content has a bigger influence on *B. thetataomicron* where high temperatures are more of a threat to the survival and persistence of *B. fragilis*. *B. fragilis* may be more tolerant (survival from 48 to 72 hours) to oxygen if they possess enzymes that detoxify oxygen where *B. thetataomicron* is less tolerant of aerobic conditions. Higher levels of

predation during warmer seasons may also play a role in the survival and persistence of *Bacteroides* species in environmental waters (Ballesté & Blanch, 2010). Presence of predators and temperature are thought to be the two main factors influencing PCR detectable DNA persistence of anaerobes in environmental waters (Kreader, 1998). The window for detection of *Bacteroides* DNA (before degradation) in water may be as little as a day or two in summer and up to two or three weeks in winter (Bell *et al.,* 2009). Detection rates in the summer may further be lowered by protozoan predators being more active in warmer weather (Bell *et al.,* 2009). According to Ebdon *et al.* (2007) adsorption to sediments as well as UV irradiation may, along with the previously mentioned factors, affect the survival rates of faecal organisms in aquatic environments. Temporal shifts linked to seasonal changes may occur in *E. coli* and *Lactobacillus* populations, where *Bacteroides* and *Bifidobacteria* groups seem to remain steady through time, with no seasonal variation in their distribution (Jofre *et al.*, 2014). Meteorological conditions; particularly storm events, which seem to be increasing in frequency due to climate change, result in greater incidences of faecal indicator organisms and presumably high pathogen loads in environmental water sources (Nnane *et al.,* 2011). The incidence of waterborne diseases after storm events may be attributed to elevated water levels creating excess runoff which includes: fresh faecal pollution from animal sources, sewage leaks or overflows as well as disturbed sediments and re-suspended bacteria that have bound to particles and settled out of suspension (Sidhu *et al.,* 2012).

In a study performed by Gawler *et al.* (2007) where an effort was made to validate host specific *Bacteroidales* gene markers in Pacific Rim countries of the European Union, variable levels of sensitivity and specificity of the HF183 as well as CF128 markers were found. Faecal samples were taken from human volunteers, pigs, cattle and chicken in France, Ireland, Portugal and the United Kingdom. The HF183F, Bac708R Primer pair showed a lower range of sensitivity with an average of 86% (Gawler *et al.,* 2007). The average specificity however was higher at 97%. Regional variations in detection specificity occurred and in some instances did not differentiate between human and animal faeces. The CF128F, Bac708R primer pair used in the same study had a sensitivity of 72%. The specificity however varies from 41% in Portugal to 96% in Ireland. The marker failed to differentiate between pig (physiologically the intestinal tract and diet are more similar to that of humans as mentioned earlier) and bovine faeces in France and the UK. It was further reported that chicken and human faeces both carried CF128F marker genes in Portugal. The results of this study illustrate the possibility that horizontal transfer of genetic material as well as geographic instability of genetic markers may necessitate the identification of a local strain or species of *Bacteroides* for use as an indicator of faecal pollution in that specific region (Gawler *et al.,* 2007).

In more recent studies, however, researchers have been faced with certain challenges regarding the use of *Bacteroides* molecular markers as indicators of faecal pollution (Ahmed *et al.,* 2010; Aslan & Rose, 2013; Layton *et al.,* 2013; Napier *et al.,* 2017). These challenges may stem from limited knowledge regarding *Bacteroides* and include factors such as survivability and lack of geographic stability in the environment.

If applied correctly *Bacteroides* species may be used successfully as extremely effective indicators of faecal pollution, specifically when source tracking is required. From the information in this section one can see that limited knowledge on *Bacteroides* species is available. It is thus suggested that localised validations be performed in the area of interest to confirm whether *Bacteroides* application will be advantageous to the user. Environmental and sampling conditions will have to be taken into account for accurate data interpretation. Certain parameters, and classifications such as sample volumes, whether DNQ data will be accepted as positive or negative etc. will have to be clearly defined by the researcher in order to render comparable data when making use of *Bacteroides* as a faecal indicator. These extra precautions, should however not discourage the use of *Bacteroides* as their host-associated characteristics and anaerobic nature still render them more specific than traditional indicators such as *E. coli*.

CHAPTER 2 – SYSTEMATIC LITERATURE REVIEW OF CHALLENGES ASSOCIATED WITH MOLECULAR METHODS USED FOR DETECTION OF *BACTERIOIDES* SPP.

2.1 Introduction

Recent literature suggests that *Bacteroides* markers seem to be either host-preferred or host-associated instead of host-specific as previously believed (Layton *et al.,* 2006; Zheng *et al.,* 2013; McLellan & Eren, 2014; Odagiri *et al.,* 2015). It has also been found that faecal bacteria, including members of the *Bacteroidales*, are likely to differ physiologically from cultured organisms (Ballesté & Blanch, 2010). The components of faecal matter may affect the persistence of these organisms. Some experiments showed that 10 micrograms of faeces was necessary for strong detection of *B. distasonis* in a 1mL sample (10% of which was analysed by PCR) (Kreader, 1998). Horizontal transfer of *Bacteroides* species or markers may also have occurred due to close contact of animals and humans (farming communities or household pets) (Ahmed *et al.*, 2010).

When considering cost, the application of molecular techniques in itself may be a limiting factor for the effective application of *Bacteroides* based assays. Sauer *et al.* (2011) remarked that the selection of assays for microbial source tracking purposes was based on cost. Furthermore, time constraints and inadequate budgets may curb the initial number of samples researchers are able to include in their studies (Sauer *et al.*, 2011). In the opinion of Zheng *et al.* (2013) the cost of running tests is directly linked to their complexity, which in turn reflects the knowledge researchers have about current methodology, the water type being sampled as well as the purpose of the data they are aiming to collect.

When all these factors are taken into account, it becomes clear that besides the advantages of making use of host specific *Bacteroides* markers, there are limitations to any assay or technique that have to be thoroughly explored before recommending the use of any particular method for microbial source tracking.

The aim of this systematic literature review is to evaluate the application of *Bacteroides* as an indicator of water quality by systematically evaluating available literature for challenges and limitations with regard to detection and quantification of *Bacteroides* markers, specifically when molecular methods are used. This aim will be achieved by completing the following objectives: Providing an overview of the need for alternative indicators; briefly highlighting the

advantages of making use of *Bacteroides* related markers alongside or instead of traditional FIB; and tabulating key points regarding the shortcomings of *Bacteroides* molecular markers and associated assays to the end of providing recommendations on the application of *Bacteroides* related molecular assays for detection and quantification of faecal pollution in environmental water sources, particularly in developing countries.

2.2 Materials and methods

2.2.1 Eligibility criteria and search strategy

Search terms used were selected by means of the PICO framework: Problem- faecal pollution in water sources, Indicator- making use of traditional faecal indicators such as *E. coli* not specific enough to discern sources of faecal contamination; Comparison- making use of *Bacteroides* as indicators of the sources and/or magnitude of faecal contamination in water sources; Outcome- Effective faecal source identification and measurement of potential human health risk. A total of 206 records were found on EDS with the search term: “Bacteroid*” AND "environmental water*" AND (“PCR” OR "molecular marker*" OR "gene* marker*") AND (“HF183” OR “CF128” OR “BAC32”). The investigation was further refined by selecting full text articles in English, peer reviewed publications, as well as publications in academic journals as limiting factors. In total 75 articles were screened for application in the systematic review. Figure 1 illustrates the process for literature screening and article selection.

Studies from 2007 up to and including 2017, as well as those that made use of the following primers were included: HF183, HF134, BAC32, CF128, BAC708R and GenBac. Studies in environmental water sources, along with studies focussing on the use of direct PCR methods without prior culturing of organisms were included. Studies written prior to 2007 were not included in the systematic literature review. Studies using clone library or culture techniques for *Bacteroides* were excluded. Studies using novel or other *Bacteroides* primers than those already mentioned were not included. Studies of a medical nature i.e. any infections or illnesses not directly linked to faecally contaminated water were also excluded.

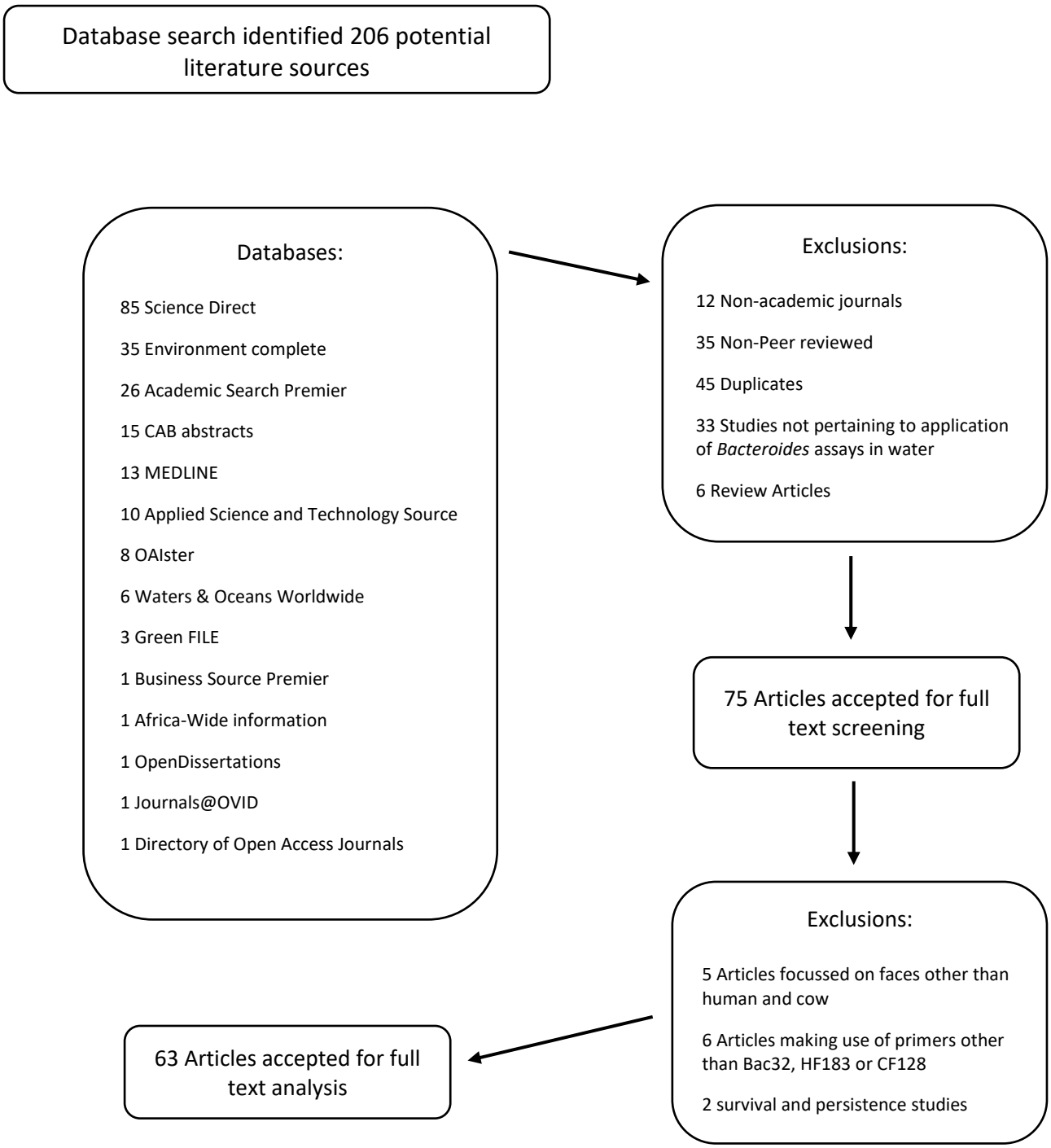


Figure 2.2-1. Flow chart of systematic search strategy

2.2.2 Data collection process

Mendeley Citation manager version 1.17.9 was used to collect and analyse literature during the selection process for the systematic review. Data was presented in table format, summarising relevant themes and findings within the selected articles regarding the use of *Bacteroides* species as indicators of faecal pollution in environmental water sources. Data was organised under the following headings: References – indicating the authors of the studies as well as the countries in which analysis took place; Aim – indicating the main or relevant purpose of the studies; Approach – stipulating sample type, sample volume and assay type were applied; Extraction method – summarising which type of DNA extraction was performed, either with a commercial kits or otherwise; Primers – listing the relevant primers and reverse primers used, if listed; Results – giving a short overview of the main findings regarding the application of relevant *Bacteroides* assays and Challenges and limitations – highlighting the themes for discussion in the review.

2.2.3 Quality assessment criteria

Literature summarised in Table 2.2-1 was assessed for applicability by placing references in one of three categories. High quality articles specifically focussed on the development, optimisation or application of *Bacteroides* host-specific assays for indication of faecal pollution. Medium quality articles compared the use of *Bacteroides* markers to traditional faecal indicators currently in use as part of a toolbox strategy or comparative marker where *Bacteroides* assays were not the main focus of the studies. Low quality articles explored the survival and persistence of *Bacteroides* within the environment, or compared secondary factors such as geographic stability which could have influenced the use of *Bacteroides* related assays. Table 2.2-1 summarises which articles fall into each of the abovementioned categories.

Table 2.2-1. Quality assessment of literature applied in the systematic review

High	Medium	Low
Gawler *et al.,* 2007	Dorai-Raj *et al.,* 2009	Santoro & Boehm, 2007
Gourmelon *et al.,* 2007	Harwood *et al.,* 2009	Walters & Field, 2009
Ahmed *et al.,* 2008[B]	Bonkosky *et al.,* 2009	Ballesté & Blanch, 2010
Ahmed *et al.,* 2008[A]	Ballesté *et al.,* 2010	Green *et al.,* 2011
Lee *et al.,* 2008	Wade *et al.,* 2010	Sercu *et al.,* 2011
Ahmed *et al.,* 2009[A]	Kirs *et al.,* 2011	Haack *et al.,* 2013
Ahmed *et al.,* 2009[B]	Chase *et al.,* 2012	Ahmed *et al.,* 2014
Jenkins *et al.,* 2009	Green & Field, 2012	Heaney *et al.,* 2014
McLain *et al.,* 2009	McQuaig *et al.,* 2012	Boehm *et al.,* 2015
Ahmed *et al.,* 2010	Staley *et al.,* 2012[B]	Cao *et al.,* 2015
Lee *et al.,* 2010	Boehm *et al.,* 2013	Napier *et al.,* 2017
Shanks *et al.,* 2010	Ebentier *et al.,* 2013	Seidel *et al.,* 2017
Sauer *et al.,* 2011	Ervin *et al.,* 2013	
Van De Werfhorst *et al.,* 2011	Gordon *et al.,* 2013	
Ahmed *et al.,* 2012	Layton *et al.,* 2013	
Staley *et al.,* 2012[A]	Liang *et al.,* 2013	
Aslan & Rose, 2013	Raith *et al.,* 2013	
Green *et al.,* 2014	Riedel *et al.,* 2015	
Riedel *et al.,* 2014	Sidhu *et al.,* 2013	
Odagiri *et al.,* 2015	Toledo-Hernandez *et al.,* 2013	
Kabiri *et al.,* 2016	Zheng *et al.,* 2013	
Kirs *et al.,* 2016	Mika *et al.,* 2014	
Mayer *et al.,* 2016	Nshimyimana *et al.,* 2014	
Nshimyimana *et al.,* 2017	Shahryari *et al.,* 2014	
	Stea *et al.,* 2015	
	Sowah *et al.,* 2017	
	Wiegner *et al.,* 2017	

2.3 Results and discussion

Literature selected within the constraints of the search criteria has been summarized in Table 2.3-1. The number of potentially useful literature sources, databases used, excluded as well as accepted sources is provided in Figure 2-1. Sixty three studies were used to extract the information presented in Table 2.3-1. This table, (Table 2.3-1) is followed by a detailed discussion of general trends and themes that have been identified by the author pertaining to the application of *Bacteroides* host-specific assays as indicators of faecal pollution in environmental water sources, the suitability of use in developing countries, and the challenges and limitations that have been experienced by researchers applying these methods. The citation period is from 2007 to 2017 and discussions are conducted in a thematic order.

Table 2.3-1. Summary of relevant *Bacteroides* literature

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Gawler *et al.,* 2007 (EU- France, Portugal & Ireland, UK)	Host specific *Bacteroides* 16S rRNA gene for source tracking	Faecal samples (250 mg). PCR	Q-Biogene FastDNA Spin kit for soil or Quiagen QiAmp blood mini kit	HF183F CF128F (Bac708R)	• HF183 86% sensitive & 97% specific • CF128 100% sensitive & 72% specific	• Markers present in various animal species across different countries (low specificity) • Survival and persistence influenced by environmental factors such as temperature • Location based validation required for optimal use of assays
Gourmelon *et al.,* 2007 (Brittany & Normandy, France)	Optimisation & validation of *Bacteroides* 16S rRNA for MST	Environmental water & waste water (filtration 100 mL). Sludge and faecal samples (250 mg). PCR	Water: Qiagen DNeasy Blood & Tissue kit Faeces: MP Biomedical FastDNA® SPIN Kit for soil	HF183F HF134F CF128F BAC32F (Bac708R)	• HF183 98% sensitive & 94% specific • HF134 84% sensitive & 99% specific • CF128 100% sensitive & 69% specific • Bac32 100% sensitive & 95% specific	• CF128 present in cow, sheep and pigs • Faecal input and marker persistence may influence results • Comparison of different detection methods can cause a lack of correlation in data • Different extraction kits were used • Assays best applied in toolbox approach
Santoro & Boehm 2007 (California, U.S.A.)	Effect of tides & waves on FIB & *Bacteroides*	Seawater (filtration 1L). Hot SDS/ proteinase K lysis. PCR	Qiagen DNeasy Blood & Tissue kit.	HF183F (Bac708R)	• HF183 present in 25% of samples from 2005 • Present in 50% of samples from 2006	• Pre-filtration of samples could have led to underestimation of HF183 • Comparison of different detection methods can cause a lack of correlation in data • Not for individual samples but good for trends
Ahmed *et al.,* 2008[A] (Queensland, Australia)	Validation of HF183, HF134 and CF128 application	Stormwater, sewage & septage (filtration 500 mL). PCR	Qiagen DNeasy Blood & Tissue kit.	HF183F HF134F CF128F BAC32F (Bac708R)	• HF183 100% sensitive & 100% specific • HF134 100% sensitive & 95.5% specific • CF128 95% sensitive & 93% specific • BAC32 100% sensitive & 100% & specific	• Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog) • Survival and persistence in the environment have to be studied • Comparison of different detection methods can cause a lack of correlation in data
Ahmed *et al.,* 2008[B] (Queensland, Australia)	Sensitivity & Specificity of HF183 & HF134 *Bacteroides* markers	Sewage & Septage water (filtration 100 mL). Faecal samples (200 mg). PCR	Quiagen QiAmp blood mini kit	HF183F HF134F (Bac708R)	• HF183 100% sensitive & 100% specific • HF134 97.3% sensitive & 95.5% specific	• Presence/ absence PCR cannot indicate magnitude of contamination • 35% of dog faeces contained HF134 marker • Detection levels were high due to sampling after storm events No field testing of markers to eliminate false positives was done

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Lee *et al.,* 2008 (Georgia, U.S.A)	Faecal exposure of 2 watersheds impacted by cattle farms under different management practices	Environmental water (filtration 100 mL). Faecal and sediment samples (0.2 –to 0.25 g). PCR	Mo-Bio UltraClean ® Soil kit	CF128 BAC32F (Bac708R)	• HF183 was only detected once 1 site in watershed 1 • CF128 detection frequency 97% for watershed 1 and 100% for watershed 2 • BAC32 detection frequency 97% for watershed 1 and 100% for watershed 2	• Amplification efficiency and copy number of targeted gene affects assay sensitivity • *Bacteroides* may grow in anaerobic soil up to 24 hours after oxygen exposure • Interpretation of data are influenced by parameter characterisation
Ahmed *et al.,* 2009[A] (Queensland, Australia)	Lower detection limits of human associated faecal pollution markers in spiked water samples	Fresh water, sea water, distilled water (filtration 125 mL). RT-PCR	Qiagen DNeasy Blood & Tissue kit.	HF183F CF128F	• HF183 100% sensitive & 98% specific • HF183 detected up to 10^{-9} dilution • CF128 used as inhibition control only	• Higher concentrations (3-4 magnitudes) of HF183 markers in the gut make correlation with presence of enteric viruses and pathogens difficult • Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog & seagull)
Ahmed *et al.,* 2009[B] (Queensland, Australia)	Specificity of 5 sewage associated *Bacteroides* markers in relation to 11 host groups	Sewage influent (centrifugation 10 mL). Faecal samples (150 – 200 mg). PCR	QiAmp stool DNA kit	HF183F	• HF183 100% sensitive & 99% specific • No inhibition was evident • LOD 1 gene copy	• Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog & sheep) • Location based validation required for optimal use of assays
Bonkosky *et al.,* 2009 (Puerto Rico)	Applicability of MST to address non-point source faecal contamination in coral reef habitats	Marine water (filtration 100mL). PCR	Mo-Bio PowerSoil® kit	HF183F HF134F BAC32F (Bac708R)	• HF183 present in 28% of samples • No amplification with HF134 • Bac32 Present in 42% of samples	• Extended marker persistence cannot differentiate recent and old pollution • HF134 developed in temperate climates, not suited to tropical environment • Tides and current affect distribution in coastal areas • MST methods can be used to monitor coral reef health

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Dorai-Raj *et al.* 2009 (Galway, Ireland)	Novel PCR assays for ruminant specific markers in contaminated rural water	Primary effluent, raw water & piped water (filtration 1L). Sewage & faecal samples (20 mg). qPCR	Mo-Bio PowerSoil® kit	HF183F CF128F BAC32F (Bac708R)	• HF183 88% sensitive and 100% specific • CF128 95% sensitive & 94% specific • Novel BAC32 assays with RumD1R reverse primer 91% sensitive & 100% specific • Novel BAC32 assays with RumD2R reverse primer 100% sensitive & 95% specific	• Sensitivity and specificity of different assays may vary across the same sample set • Location based validation required for optimal use of assays
Harwood *et al.,* 2009 (Florida & Mississippi, U.S.A.)	Standardisation and evaluation of 3 MST methods for sewage detection across 3 laboratories	Influent spiked sterile buffered water, fresh water & marine water (filtration 500 mL). Faecal samples (0.25 - 0.3 g). PCR	Mo-Bio PowerSoil® kit	HF183F BAC32F (Bac708R)	• HF183 100% sensitive & 96% specific for faecal samples • BAC32 used as inhibition control only	• 25% of samples exhibited PCR inhibition (cow faeces) • Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog, chicken and seagull) • Small sample size for limit of detection experiments • Location based validation required for optimal use of assays • Assays best applied in toolbox approach
Jenkins *et al.,* 2009 (Njoro, Kenya) (California, U.S.A.)*	Determining if American and European qPCR assays for faecal pollution can be used in surface water in African countries	River water (filtration 50 mL). Raw sewage & faecal samples (0.5 mL). qPCR	QiAmp stool DNA kit	HF183F CF128F (Bac708R)	• HF183 65% sensitive & 100% specific • CF128 with BacCow-305r 94% sensitive & 100% specific	• CF128 present in cow, goat, horse and sheep (overestimation of presence) • Sensitivity and specificity of different assays may vary across the same sample set • Limited resources for sample processing in Kenya • DNA pasteurised before use
McLain *et al.,* 2009 Arizona, U.S.A.)	Cross-amplification of human-specific *Bacteroides* 16S rRNA markers in 4 freshwater fish species	Fish faeces (0.5 g). PCR. qPCR	Zymo Research Faecal DNA kit	HF183F HF134F	• HF183 with HF264R detected in human and trout faecal samples • HF134 with HF654R did not amplify any fish faecal samples	• Limited information on fish-*Bacteroides* associations exist • Primer design considerations will have to be made to exclude cross amplification • Misinterpretation of results in the presence of fish faeces

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Walters & Field 2009 (Oregon, U.S.A.)	Survival of *Bacteroides* cells and persistence of 4 faecal *Bacteroides* markers	River water (filtration 100 mL). Human faecal slurry (200 mg), cow faecal slurry (220 mg). qPCR	Qiagen DNeasy Blood & Tissue kit	HF183 HF134 CF128 (Bac708R)	• Human markers not detected after day 7 • Cow marker detected up to day 14	• Occurrence of growth for 4 days after microcosm inoculation • PCR inhibitory substances in cow faeces may interfere with DNA extraction & PCR • Proportion of bacteria differs between hosts • Strain related survival capabilities may influence detection
Ahmed *et al.,* 2010 (Dhaka, Bangladesh)	Assessing the magnitude of sewage pollution in an urban lake	Lake water (filtration 100 mL). Faecal samples (0.2-0.5 g). qPCR	QiAmp stool DNA kit	HF183F CF128F (Bac708R)	• HF183 87% sensitive & 93% specific • CF128 75% sensitive & 100% specific	• Lake water contained PCR inhibitors • Small sample size may affect specificity results • Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog) • Detection rates rapidly decreased within 1 km radius • Detection levels were high due to sampling after storm events
Ballesté & Blanch 2010 (Barcelona, Spain)	*Bacteroides* survival and persistence in the environment	River water (filtration 200 mL). PCR. RT-PCR	Quiagen QiAmp blood mini kit	BAC32F (Bac708R)	• *B. fragilis* survives longer in winter & *B. thetaiotaomicron* longer in summer • Environmental strains persist longer than laboratory strains	• Variables in onsite environmental experiments are difficult to monitor, control and replicate • Environmental factors (temperature & predation) influence persistence & detection of *Bacteroides* markers in water • Misinterpretation of results may occur due to differing DNA degradation rates • Molecular techniques in this study do not distinguish between viable, culturable and non-culturable cells
Ballesté *et al.* 2010 (Barcelona, Spain)	Combinations in which the least number of markers yields highest rates of source discrimination	Waste water, faecal samples & faecal slurries (200 µL). PCR	Quiagen QiAmp blood mini kit	HF183 HF134 CF128 (Bac708R)	• HF183 50% sensitive & 71% specific • HF134 30% sensitive & 81% specific • CF128 26% sensitive & 100% specific	• Frozen samples were used* • Sensitivity in combination with specificity are important for marker efficiency • Markers have been developed from genomic library information for non-cultured species • Geographic & environmental factors may influence sample composition and assay results • Location based validation required for optimal use of assays • Assays best applied in toolbox approach

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Lee *et al.,* 2010 (Ontario, Canada)	Characterisation of faecal sources of water pollution using host-specific *Bacteroides* TaqMan RT-PCR assays	River water, raw influent & final effluent (filtration 100 mL). Faecal samples (homogenised in PBS & mixed with lysis buffer 1:2 v/v). RT-PCR	Qiagen DNeasy mini spin columns	BAC32F (Bac708R)	• BAC32 used as assay calibrator only • Markers increased significantly between upstream sampling points and suspected point of inflow, remained high up to 60 m downstream	• Cow specific marker present in deer faeces • Test conditions and parameters have to be clearly defined to ensure correct interpretation of results • Geographic stability has to be field tested • Multiple sources of faecal contamination may make results difficult to analyse (baseline ratio method suggested)
Shanks *et al.,* 2010 (Ohio, U.S.A.)	Assessment of 5 end-point and 10 real-time *Bacteroides* PCR assays to indicate presence of human faecal pollution	Primary influent (filtration 25 mL). Faecal samples (500 mg –kit instructions). PCR. qPCR	MP Biomedical FastDNA® SPIN Kit for soil.	HF183F HF134F (Bac708R) (BthetP1)	• HF183 100% sensitive & 95% specific • HF134 100% sensitive & 81% specific • On average low numbers of HF183 marker (0.06 $\log_{10}$ copies) found in non-target sources	• Differences in distribution of *Bacteroides* in hosts groups lead to geographic instability • Test conditions and parameters have to be clearly defined to ensure correct interpretation of results • Data is limited by small sample sizes in studies • Prolonged storage is detrimental to sample quality (less than 12 months)
Wade *et al.,* 2010 (Alabama, Mississippi & Rhode Island, U.S.A.)	Relationships between swimming-related illnesses and FIB at 3 marine beaches	Marine water (filtration 50-100 mL). qPCR	AE buffer, bead beating & centrifugation	Genbac 3	• When CCE$\Delta\Delta$ values were 2 530 CCE/100 mL 12% of swimmers and 6% on non-swimmers reported GI illness • When CCE$\Delta\Delta$ values were below 542 CCE/100 mL only 4% of swimmers reported GI illness	• Statistical significance of markers among swimmers was not consistently associated with GI symptoms • Point sources in the vicinity may affect results • Birds may be the predominant source of faecal pollution at beaches • This study may not be geographically stable (not suited for use in humid tropical environments)
Green *et al.,* 2011 (Oregon, U.S.A.)	Marker persistence in marine and fresh water by *Bacteroidales* ratios	Sewage spiked river water (filtration 50 mL). qPCR	Qiagen All Prep DNA/RNA Micro Kit®	HF183Taqman HF183 GenBac3 (HF303R)	• Lag phase was 3.1 days longer in salt water • rDNA measured by Genbac3 had a higher decay rate in light compared to dark	• Higher levels of predation in freshwater may influence results • Microcosms cannot accurately represent complex natural interactions • Results may be biased due to lack of replicate microcosms

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Kirs *et al.,* 2011 (Nelson, New Zealand)	Application of PCR based assays using 8 MST markers including *Bacteroidales*.	Primary influent & oxidation pond water (filtration 10-25 mL). Faecal samples (0.2 g). PCR	Zymo Research Faecal DNA kit	HF183 CF128 BAC32 (Bac708R)	• HF183 100% sensitive & 97% specific • CF128 85% sensitive & 65% specific • BAC32 99% sensitive & 100% specific indicating absence of inhibition	• CF128 present in wastewater & human faecal samples • Cross reaction of markers with brush-tailed possum, marsupial, gull & rabbit faeces • HF183 markers are either giving false positives as there is no correlation with other human markers used or HF183 is more sensitive than the other markers used • Location based validation required for optimal use of assays • Larger sample volumes containing more host groups is necessary for method validation (specificity influenced by number of individuals per taxon analysed)
Sauer *et al.,* 2011 (Wisconsin, U.S.A.)	Degree of human faecal pollution in urban areas from stormwater in wet weather	Stormwater (filtration 200 mL). Sewage (filtration 100 mL). PCR. qPCR	MP Biomedical FastDNA® SPIN Kit for soil	HF183 BAC32F (Bac708R)	• HF183 100% sensitive & 83-95% specific • BAC32 100% sensitive & 100% specific • 57% of all samples taken contained HF183 markers • Out of 9.8 x 10^8 copy numbers of BAC32, 4.8 x 10^7 copy numbers were HF183	• Rainfall, pre-storm conditions and water system infrastructure can influence sites • Upstream contamination points may not be detected by outfall samples • In-line samples contained more *Bacteroides* than grab samples, but human markers remained the same • Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog)
Sercu *et al.,* 2011 (California, U.S.A.)	Wet weather influence on dry weather pollution sources in creeks	Creek water, stormwater (filtration 2L). qPCR	Mo-Bio UltraClean ® Water kit	HF183F	• Variability in indicator numbers was lower in wet weather • No wet weather samples were positive for HF183	• Extra runoff in wet weather makes determination of point sources difficult • 1000 fold less HF183 was detected in wet weather • Filtering larger volumes of water may increase detection rates • Characterisation of wildlife faecal impacts have not been characterised

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Van De Werfhorst *et al.,* 2011 (California, U.S.A.)	Host specificity of 2 human-specific *Bacteroides* markers	Sewage & septage water (filtration 50 mL-kit instruction). Faecal samples (0.25 g-kit instruction). qPCR	Water: Mo-Bio PowerWater® kit & ethanol precipitation Faeces: Mo-Bio PowerSoil® kit & ethanol precipitation	HF183SYBR (Bac708R)	• HF183 100% sensitive & 100% specific for sewage • HF183 100% sensitive & 67% specific for septage • HF183 100% sensitive & 63% specific for faecal samples	• Different extraction kits were used • Level of treatment of water can influence detection of MST markers • Sample variability (size/volume/number of individuals) can influence detection of MST markers • Assays best applied in toolbox approach
Ahmed *et al.,* 2012 (Queensland, Australia)	Consistency of markers compared to 2007 study	Primary & secondary influent (100 mL). Faecal samples (100-200 mg). qPCR	Water: centrifugation in filter columns, Qiagen DNeasy Blood & Tissue kit Faeces: QiAmp stool DNA kit	HF183F CF128 (Bac708R)	• HF183 95% sensitive & 94% specific • CF128 used as inhibition control only	• Different extraction kits were used • Marker prevalence in as in not well known • Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog, chicken & bird) • Specificity of less than 80% is not acceptable for MST
Chase *et al.,* 2012 (Florida, U.S.A.)	Sources of faecal pollution during high-flow periods in the New River watershed	River water & diluted sediment samples (filtration 500 mL). PCR. qPCR	Mo-Bio PowerSoil® kit	HF183F CF128 (Bac708R)	• HF183 PCR reaction efficiency 94.6% • Detection threshold 10 gene copies per 100 mL • CF128 not detected upstream of areas containing ruminants	• Marker levels must be compared over wet and dry seasons • Agricultural and wildlife sources may contribute to faecal load • Soil and sediments must be taken into account as potential reservoirs of faecal indicators or markers
Green & Field 2012 (Washington, U.S.A.)	Application of Kinetic Outlier Detection methods increased qPCR sensitivity	Sewage spiked marine water (filtration 100 mL). qPCR	Qiagen DNeasy Blood & Tissue kit	HF183F HF134F (Bac708R)	• HF183 qPCR not greatly affected by ethanol during DNA extraction • High water turbidity negatively correlated with qPCR detection • HF183 more susceptible to humic acid inhibition than other markers in the study	• qPCR may be inaccessible due to lack of efficiency and funding • Presence of inhibitors and their effects on certain samples is not well understood • Accuracy of assays may be impaired due to the inclusion or exclusion of data (such as inhibited samples)

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
McQuaig *et al.,* 2012 (California, U.S.A.)	Water quality of two marine beaches based on MST markers and FIB	Beach water (filtration 500 mL). Touchdown PCR	Modified MoBio PowerSoil® & Quiagen DNeasy DNA extraction protocol	HF183F (Bac708R)	• HF183 was detected in 23.6% of samples • HF183 was detected more frequently than other markers in the study	• Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog, chicken & cat) • Composite or community based samples are necessary for effective detection of faecal markers and FIB are not correlated with one another • Assays best applied in toolbox approach
Staley *et al.,* 2012[A] (Florida U.S.A.)	Origin of faecal pollution, link between human markers and pathogens and effects of stormwater runoff	Lake water (filtration 500 mL). PCR	Mo-Bio PowerSoil® kit	HF183F (Bac807R)	• HF183 detected most frequently (10.1% of samples) • Co-occurrence of markers only during rainfall events	• Pathogen presence not correlated to HF183 occurrence • Pathogens tested can be from animal hosts • Larger water volumes that are concentrated more efficiently are needed to determine the relationship between HF183 and pathogens • Stormwater significantly increases FIB and human faecal markers
Staley *et al.,* 2012[B] (Florida, U.S.A.)	HF183 and HPyVs to determine specificity and LOD in 5 types of inland water	Swamp, river beach and lake water (filtration 500 mL). Sewage samples (700 µL). Faecal samples (0.2-0.3 g). qPCR	Mo-Bio PowerSoil® kit	HF183F GenBac3	• HF183 81.25% specific with SSHBacR reverse primer • HF183 detectable up to 10^{-6} dilutions, only quantifiable up to 10^{-2} • Detectable dilutions were inconsistent across sites and sampling dates • Water from 2 sites contained HF183 markers prior to being spiked	• Presence/ absence PCR cannot indicate magnitude of contamination • 36% of undiluted cattle faecal samples exhibited PCR inhibition • Interpretation of results are influenced by parameter characterisation • Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog, chicken & duck)

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Aslan & Rose 2013 (Michigan, U.S.A.)	Specificity of HF183 compared to another human specific marker and *B. thetaiotaomicron* housekeeping gene for cross reactivity detection	Secondary & tertiary effluent (filtration 1L). Pelleted influent (400 µL). Faecal samples (200 mg). PCR	Water: 2mL concentrate and 180 µL buffer solution mixed with 20 µL Proteinase K. Incubated at 65°C for 30 minutes prior to extraction. Faeces: QiAmp stool DNA kit	HF183F (Bac708R)	• HF183 100% sensitive & 71% specific • Housekeeping gene had 97% specificity	• Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (cow, dog, cat, goose, duck, chicken, gull, horse, turkey & sheep) • Geographic stability has to be field tested • PCR inhibitors are present in animal faeces
Boehm *et al.,* 2013 (California, U.S.A.)	Challenging 41 MST methods over 27 laboratories with 64 blind samples that contain 12 possible faecal sources	Faecal slurries, sewage and septage water- (filtration 50mL or 5mL for 1:10 dilutions). PCR and/or qPCR	Different kits by different laboratories: Generite DNA-EZ, Qiagen DNEasy, Quiagen QIAmp, MP Biomedical FastDNA spin, MoBio Power-Soil and Phenol:chloroform extraction	HF183F HF183Taqman HF183SYBR CF128 (Bac708R)	• 2 labs reported 50% sensitivity for HF183. 5 labs reported ≥80%. • All labs reported specificity for HF183 of ≥92% • HF183SYBR sensitivity ≥97% • HF183SYBR specificity reported as 28% by 1 lab and ≥85% by the other 3 • HF183Taqman had varying results across the 5 labs and was classified as non-specific • CF128 ≥80% sensitive and ≥80% specific by 1 lab and 40% specificity by the other	• HF183SYBR present in deer, goose and horse faeces at similar or higher levels than target faeces • CF128 was not detected in target faeces by 1 lab • Environmental factors may influence abundance and detection of markers • The age or storage of samples may decrease the potential for marker detection • Data is limited by small sample sizes in studies

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Ebentier *et al.,* 2013 (California, U.S.A.)	Repeatability and reproducibility of MST methods across 10 labs to emphasise the need for standardization of protocols	Faecal slurries, sewage and septage water- (filtration 50mL or 5mL for 1:10 dilutions). qPCR	Different kits by different laboratories: Generite DNA-EZ, Qiagen DNEasy, Quiagen QIAmp, MP Biomedical FastDNA spin, MoBio Power-Soil and Phenol:chloroform extraction	HF183Taqman	• Highest coefficient of variance observed with 1:0 diluted singleton and the minor contributors in doubleton samples • HF183Taqman and two other protocols were most reproducible	• Unequal sample sizes influence statistical comparison of data • Faecal concentrations that are likely to be found in the environment should be tested • Environmental factors such as the nature of water matrices should be taken into account • Lower %CV may be caused by using only solid faecal samples for animal sources and methods
Ervin *et al.,* 2013 (California, U.S.A.)	Effects of physical, culture-based and qPCR characterisation methods on interpretation of results	Composite faecal slurries (filtration 0.3-180 mg/ 100 mL wet weight). qPCR	No DNA extraction method listed	GenBac3	• Least correlation between wet faecal mass and total DNA mass compared to culture based and qPCR measurement units • DNA mass and *Bacteroides* qPCR had a 5% (significant) correlation level	• Results can only be interpreted meaningfully if the intended use of the method of measurement is clear • More research on ranges of variability of markers are required • Horizontal transfer and non-point source pollution may influence results (gulls & birds)
Gordon *et al.,* 2013 (Florida, Mississippi, Louisiana & Texas U.S.A.)	Relationship between *Enterococci* and human associated MST markers in the Gulf of Mexico	Fresh and marine surface water (filtration 500 mL). PCR	Mo-Bio PowerSoil® kit	HF183F (Bac708R)	• HF183 detected most frequently (20.3%) • HF183 less specific than *M. Smithii* and HPyVs but present in higher concentrations in sewage	• Assays best applied in toolbox approach • Relationship between *Enterococci* and HF183 statistically significant in this study • Location based validation required for optimal use of assays
Haack *et al.,* 2013 (Michigan, U.S.A.)	Environmental factors, detection of FIB, swimming-associated disease and MST genes over similar geographic regions	Beach water samples (filtration 100 mL). qPCR	Extraction from 100-400 µL of glycerol stock sample. Qiagen DNeasy Blood & Tissue kit	HF183F (Bac708R)	• HF183 detected in 25% of samples • HF183 detected at every beach	• No single source of faecal pollution influences any or all of the beaches • Environmental factors influence beaches on a regional and local scale • Extreme variability in samples make application of data on other beaches impossible

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Layton *et al.,* 2013 (California, U.S.A.)	Performance of 10 human faecal anaerobe-associated PCR assays under different performance metrics by 17 labs	Faecal slurries, sewage and septage water-(filtration 50mL or 5mL for 1:10 dilutions). PCR and/or qPCR	Different kits by different laboratories: Generite DNA-EZ, Qiagen DNEasy, Quiagen QIAmp, MP Biomedical FastDNA spin, MoBio Power-Soil and Phenol:chloroform extraction	HF183F HF183Taqman HF183SYBR GenBac3 (Bac708R)	• HF183 88% sensitive & 100% specific • HF183Taqman 100% sensitive, specificity not reported • HF183SYBR 91% sensitive, specificity not reported	• Assays performed worse than originally reported • HF183SYBR present in chicken faeces, HF183Taqman present in chicken and dog faeces • Higher specificity if DNQ samples were considered negative • Septage had highest microbial abundances • HF183 end-point less sensitive than HF183 qPCR for sewage samples
Liang *et al.,* 2013 (Florida U.S.A.)	Frequency of detection of pathogens and indicators in association with land use, rainfall and physico-chemical properties of water	Retention pond, river water, canal water, lake water (filtration 500 mL). PCR	Mo-Bio PowerSoil® kit	HF183F CF128F (Bac708R)	• 8,7% of samples positive for HF183 • 19.3% of samples positive for CF128	• Higher HF183 frequency in residential areas • CF128 found at higher frequency than HF183 in all samples • No seasonal differences in detection • HF183 present in samples with lower oxygen levels
Raith *et al.,* 2013 (California, U.S.A.)	Performance of 2 endpoint PCR assays and 5 qPCR assays associated with ruminant or cow faeces, when applying different performance metrics by 11 labs	Faecal slurries, sewage and septage water-(filtration 50mL or 5mL for 1:10 dilutions). PCR	Different kits by different laboratories: Generite DNA-EZ, Qiagen DNEasy, Quiagen QIAmp, MP Biomedical FastDNA spin, MoBio Power-Soil and Phenol:chloroform extraction	CF128F (Bac708R)	• CF128 33-100% sensitive & 88-100% specific • qPCR using 16S rRNA is more sensitive than ribosomal based markers	• Interpretation of results are influenced by parameter characterisation (benchmark threshold selection) • Environmental factors, lab methods and instrumentation as well as proficiency of technicians can influence results • Endpoint PCR data cannot be normalised which make interpretation of result difficult

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Sidhu *et al.,* 2013 (Victoria, Queensland & New South Wales Australia)	Frequency of sewage contamination of stormwater runoff in urban catchments	Stormwater - inline sampling, concentration of 20L samples to 10 mL) (filtration 10 mL). RT-PCR	Mo-Bio PowerSoil® kit	HF183F	• HF183 most frequently detected at 96% • HF183 had 76% concurrence with other sewage markers	• 99% probability of HF183 in stormwater indicating direct sewage contamination • Aging infrastructure and poor management result in faecal contamination of stormwater • Assays best applied in toolbox approach • Low cost treatment options such as retention, wetland filtration and UV disinfection can be considered before reusing water
Toledo-Hernandez *et al.,* 2013 (Puerto Rico) (Ohio, U.S.A.)*	Spatial and temporal applicability of host-specific marker specificity in tropical areas	Environmental water & Waste water (filtration 100 mL). Faecal and raw sewage samples (0.25 - 0.30 g). PCR	Mo-Bio PowerSoil® kit	HF183F CF128F BAC32F (Bac708R)	• HF183 100% specific, present in 75% of sewage samples • CF128 64% specific	• CF128 present in turkey, pig, chicken & WWTP samples • DNA sample dilution was necessary to prevent inhibition • Survival of *Bacteroides* in tropical regions & seasonal variability of marker presence is not well documented • Location based validation required for optimal use of assays
Zheng *et al.,* 2013 (Illinois, Kentucky, Missouri & Wisconsin, U.S.A.)	Distinguishing human and animal faecal pollution in karst aquifers	Spring and well water (filtration 1L). qPCR	Bead beating, CTAB, Phenol chloroform extraction followed by isopropanol precipitation and ethanol wash	BAC32F	• Small sample size showed no significant difference • 49 of 62 samples contained *Bacteroides* markers with BAC303R primer • Human markers were higher in wells during springtime	• Geographical and geological differences limit comparisons between water from different regions • Chlorination of well during sampling period may influence results • Comparison of different detection methods can cause a lack of correlation in data • Assays best applied in toolbox approach
Ahmed *et al.,* 2014 (Queensland, Australia)	Relative inactivation times of HF183, FIB and Human adenoviruses in sewage spiked microcosms	Fresh water (drinking) & sea water (40 mL reduced to 200 µL). qPCR	Centrifugation in filter columns, Qiagen DNeasy Blood & Tissue kit	HF183F (Bac708R)	• T_{90} inactivation time in fresh water was 3.5 days • T_{90} inactivation time in sea water was 2.7 days • HF183 inactivation time similar to FIB but shorter than adenoviruses	• HF183 may be sensitive to salinity • Comparison of two different approaches for determining inactivation time • HF183 presence may be due to presence in non-target animal faeces due to horizontal transfer • Assays best applied in toolbox approach

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Green *et al.* 2014 (Ohio, U.S.A.)	Assay Modification excluding amplification by-products, incorporation of new strategies and reagents for monitoring of inhibition	Influent (filtration 25 mL). Faecal samples (1g). qPCR	GITC buffer slurry, 800 µL used for bead beating GeneRite DNA-EZ kit	HF183F	• HF183 detected in all sewage samples by both assays • Limit of detection was tenfold lower when using (BacR287) reverse primer • Primer dimers were present in all samples using (BFDrev)	• Assays may yield different results when adapted and optimised • Integration into management programmes are dependent on method standardisation
Heaney *et al.,* 2014 (Alabama & Rhode Island, U.S.A.)	Factors influencing measures of FIB in beach sand including FIB in water, weather and environment	Beach water & sand slurries (filtration 100 mL). qPCR	AE buffer with salmon testes DNA, bead milling & centrifugation	GenBac3	• *Bacteroidales* were present in 68% of samples • *Bacteroides* (faecal) were present in 53% • Highest correlation between sand and water samples for *Bacteroides*	• PCR inhibitors in sand and turbid water may influence results • Limited sand samples were taken as part of a separate study • Sample matrix difference such as mineral content may influence the survival, persistence and detection of microorganisms in sand and water
Mika *et al.,* 2014 (California, U.S.A.)	Distribution and frequency of FIB and HF183, relationships between them and effects of rainfall on their concentrations	Storm drain, channel, harbour & marina water (filtration 1-2L). qPCR	Mo-Bio UltraClean ® Faecal kit	HF183SYBR	• HF183 detected in 58% of samples for Santa Monica Canyon channel, none were quantifiable • HF183 detected in 14% of samples at Ventura Marina Dock, 4.6% thereof was quantifiable • Samples with HF183 also had higher FIB levels.	• Detection levels were higher in wet weather • No seasonal variation in levels • Upstream and downstream levels did not correlate • Complex relationship with environmental factors (dilution, UV inactivation) may influence presence and persistence of genes • Multiple parameters may need to be tested for source identification
Nshimyimana *et al.,* 2014 (Singapore, Asia)	Determining relationships between FIB, HF183 markers and land use	Fishpond water, onsite treated effluent & residential WWTP influent (filtration 500 mL-1L). Touchdown & semi-nested PCR. qPCR	PCR: Mo-Bio UltraClean ® soil kit qPCR: Mo-Bio UltraClean ® plant kit	HF183F (Bac708R)	• HF183 levels higher in farming/horticultural areas relative to residential and undeveloped areas • 72% of samples with high *E.coli* levels had high HF183 levels as well	• Different extraction kits were used for PCR and qPCR DNA • PCR inhibition that could not be remedied by dilution was present in many of the samples • Dilution made touchdown PCR less reliable • No threshold value for HF183 in environmental water has been established as of yet • Markers have been developed from genomic library information for non-cultured species

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Riedel *et al.,* 2014 (California, U.S.A.)	Costs and capabilities of environmentally relevant PCR and qPCR assays	Spiked artificial, marine and creek water (filtration 10 mL of 1:20 dilution). PCR. qPCR	GeneRite DNA-EZ kit ST1	HF183F HF183Taqman HF183SYBR (Bac708R)	• Water type did not affect assay performance significantly • HF183Taqman had the lowest limit of detection followed by HF183SYBR then HF183	• HF183SYBR yielded inconsistent results even with the same reagents as HF183Taqman • qPCR may be inaccessible due to lack of efficiency and funding • dye sensitivity for PCR and qPCR may influence LOD data
Shahryari *et al.,* 2014 (Isfahan province, Iran)	Traditional FIB, and *Bacteroidales* markers for detection of faecal contamination in drinking water sources	Well, spring, aqueduct and river water (filtration 500 mL). Nested PCR	Liquid nitrogen freezing and reheating in boiling water x3 Promega Wizard® genomic DNA kit	BAC32F (Bac708R)	• *Bacteroides* detected most frequently (68%) • Only 56.1% of samples contained *Bacteroides* along with coliforms or other bacteria	• Comparison of different detection methods can cause a lack of correlation in data • Environmental *Bacteroides* may be detected along with faecal *Bacteroides* by this method • Inhibitors and environmental factors that affect persistence influence the results
Boehm *et al.,* 2015 (California, U.S.A.)	Risk assessment of Human-associated qPCR marker concentrations to occurrence of GI illness among swimmers	-	-	HF183Taqman	• HF183 marker concentrations double that of HumM2 marker • If 4200 copies/ 100 mL is present rate of 30 in 1000 swimmers is predicted	• Environmental factors, lab methods and instrumentation as well as proficiency of technicians can influence results • Study is modelled on raw sewage contamination only
Cao *et al.,* 2015 (California, U.S.A.)	Development of duplex digital droplet PCR for simultaneous analysis of general and human faecal indicators	Beach water, Storm-water (filtration). Faecal samples. qPCR and digital qPCR	GeneRite DNA-EZ kit MST1	HF183Taqman with HEX instead of FAM label	• Lowest LOD for OSTD1 plasmid was at 10^{-6} dilution with 3 plasmid copies per reaction • Duplex ddPCR 81% sensitive and 59% specific • Simplex qPCR 84% sensitive and 68% specific • Simplex ddPCR 75% sensitive and 50% specific	• Duplex qPCR did not yield results or caused underestimation of presence of one or either markers due to interference between them • Both dPCR techniques and simplex qPCR can be applied for HF183 detection and quantification

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Odagiri *et al.,* 2015 (Odisha, India)	Performance of selected assays developed in U.S.A and Europe for application in India	Human faeces, pooled animal faeces & waste water (180-220 mg). qPCR	QiAmp stool DNA kit	HF183Taqman HF183SYBR GenBac3	• HF183Taqman 80% specific • HF183SYBR 89% sensitive • GenBac 95% sensitive	• Geographic & environmental factors may influence sample composition and assay results • Data is limited by small sample sizes in studies • Cross reactivity with animal faecal samples (HF183Taqman: dog & chicken; HF183SYBR: cow, buffalo, goat, sheep, dog & chicken) • Large variations in the abundance of *Bacteroides* markers in individuals • Prolonged storage (3 months)
Riedel *et al.,* 2015 (California U.S.A.)	Investigation of sources of faecal pollution and applicability of using MST in Topanga Watershed	Creek water, lagoon water marine water (filtration 200 mL). qPCR	GeneRite DNA-EZ kit ST1	HF183Taqman	• Elevated human markers present in creek in wet weather • HF183 detected in 64% of creek samples • 2200 copies/100 mL highest detected level in lagoon samples • HF183 detected four times at sea level or below at 1000 copies/100 mL	• No seasonal variation of HF183 markers in lagoon samples • HF183 levels in marine water only detected in winter • Human markers were not detected in the lagoon during winter • No correlation between lagoon and marine contamination events
Stea *et al.,* 2015 (Nova Scotia, Canada)	Relationships between enteric pathogens, water quality measures, mitochondrial DNA based faecal source tracking and human and ruminant *Bacteroides* markers	Lake water & river water (filtration 500 mL). Faecal samples (250 mg). qPCR	Mo-Bio PowerSoil® kit	HF183F	• HF183 100% sensitive & 91% specific • Detected in all urban baseflow samples and 20% of samples in rural areas after storm events	• Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (rabbit, raccoon & porcupine) • Faster runoff rates in urban areas can lead to misinterpretation of pollution incidences • Faecal input and marker persistence may influence results

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Kabiri *et al.,* 2016 (Arizona, U.S.A.)	Characterisation of variable gene regions within the *Bacteroides*16S rRNA gene sequence via alignment of fish and animal samples against human samples	River water, canal water, drinking water plant intake, reclamation pond water (1L). Faecal samples (0.5 g). PCR	Zymo Faecal DNA extraction kit	Hf183F HF134F (Bac708R)	• Samples positive for both markers or only HF183 were considered positive • Sequence similarity between *B. fragilis* and human faecal sequences 99%	• Horizontal transfer of *Bacteroides* between animals and humans can cause human markers to amplify animal faeces (dog) • Small sample size limiting data, markers are not present in all individuals of a species • Primers detect different species and/or strains of *Bacteroides* (HF183: *B. dorei*; HF134: *B. dorei, B. vulgatus & B. massiliensis*) • Assays best applied in toolbox approach
Kirs *et al.,* 2016 (Oahu, Hawaii)	Presence of sewage pollution in recreational waters in Hawaii by means of human faecal molecular markers	Raw sewage (filtration 25 mL). Influent before UV treatment (filtration 50 mL). Microcosm water (100 mL). Marine and fresh water (300 mL). Faecal samples (0.3g). qPCR	10 minute vortexing replaced by bead beating for 2 min Mo-Bio PowerSoil® kit	HF18Taqman	• HF183 100% sensitive for wastewater • HF183 80% sensitive for faecal samples • Overall specificity 78% • HF183 concentration decreased by 62,5% after UV disinfection • HF183 concentration decreased by 99.,99% after secondary treatment with no disinfection • Markers quantifiable for 11-15 days in both types of water, detectable for twice as long	• Different sample volumes were compared • Cross reaction of markers with dog, mongoose & cat faeces • Treatment processes may influence faecal indicators differently • UV disinfection can make interpretation of data difficult • Location based validation required for optimal use of assays

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Mayer *et al.,* 2016 (Vienna, Austria)	Prevalence and abundance of human-associated *Bacteroides* markers over 1 year in untreated and treated wastewater	Influent (filtration 20 mL). Effluent (filtration 50 mL). Filtered effluent (1.5 L). qPCR	Phenol: chloroform extraction	HF183Taqman	• Active sludge treatment did not greatly affect marker concentration (log_{10} 8.6 down to log_{10} 6.4) • Significant relationships exist between BacHum, JCPyV and HF183 marker concentrations	• Random grab samples are not as effective as in-line volume proportional sampling systems • PCR requires more preparation than MPN • Environmental factors, mobility, persistence, specificity of assays and indicators, lab methods and instrumentation can influence results • No correlations found between *Bacteroides* and FIB • No statistical difference between markers and different water treatment systems or sampling during different seasons • HF183 targets only *B.dorei*
Napier *et al.,* 2017 (Ohio, U.S.A.)	Associations between self-reported illness by swimmers and 4 human-*Bacteroides* markers at 6 marine and freshwater beaches	Beach water (filtration 1L). qPCR	Bead milling, concentration using 96-well silica column from Quiagen DNEasy kit along with buffers from a Generite DNA-EZ kit for purification	HF182Taqman GenBac3	• >98% samples contained *Bacteroides* markers • 4-49% of samples contained HF183 • 8 in 1000 swimmers reported respiratory illness in the presence of *Bacteroides* markers • Higher risk of illness when 2-3 of 4 markers were present in the same sample	• Prolonged storage of samples (2-6) years influence detectability of markers and accuracy of quantification • Lower abundance of human markers in the environment may require larger sample sizes • Environmental factors influence the persistence and detection of markers by beach • Detection varied by assay • Quantitative better for risk assessment
Nshimyimana *et al.,* 2017 (Singapore, Asia)	Best assays for detecting human faecal pollution in Singapore	Sewage samples (filtration 1L). Faecal samples (2 mg). qPCR	Mo-Bio PowerFecal® kit	HF183Taqman HF183SYBR	• Both HF183 markers 100% sensitive for sewage • HF183 Taqman 60% sensitive for stool samples • HF183 SYBR 63% sensitive for stool samples • HF183Taqman 90% specific • HF183SYBR 88% specific	• Cross reactivity of markers with rabbit, chicken, boar and monkey faeces • Geographic differences exist between assay performance for Singapore, Bangladesh & India • Molecular methodology may be unaffordable to developing countries • Location based validation required for optimal use of assays

References	Aim	Approach	Extraction method	Primers	Results	Challenges & Limitations
Seidel *et al.,* 2017 (North Rhine-Westphalia, Germany)	Improvement of live-dead differentiation during qPCR by Propidium monoazide or dimethylsufoxide (DMSO) addition	River water, creek water, influent & effluent (filtration 500 mL). PCR. qPCR	Selected treatment with heat, PMA and/or DMSO Mo-Bio PowerSoil® kit	HF183F HF183Taqman (Bac708R)	• Bac708R used to obtain longer amplicons • All assay 100% specific for their target • PMA did not affect fresh faeces • Depending on long or short sequence target assays water stored for a week at 4°C showed a drop in intact cells	• DNA sample dilution was necessary to prevent inhibition • DMSO (not tested on *Bacteroides*) may increase solubility of faecal matter • qPCR had higher specificity • Detection of longer sequences showed higher relative proportions of membrane damaged cells
Sowah *et al.,* 2017 (Georgia, U.S.A)	Impact of Septic systems on faecal pollution load in suburban streams	Stream water (filtration 100 mL). qPCR	Mo-Bio PowerFecal® kit	HF183F	• HF183 quantifiable in 57% of samples • HF183 detected at a maximum of 3.7 $\log_{10}$ copies per 100 mL	• Assays best applied in toolbox approach • Highest bacterial yields in spring specifically in high density areas • Inadequate information on copies of markers in different hosts • Positive correlation between HF183 and FIB indicate a continuous source of pollution such as a septic system • Location based validation required for optimal use of assays
Wiegner *et al.,* 2017 (Hawaii, U.S.A.)	Variability of microbial pollution spatially and temporally on a tropical estuary	River water, ground water, bay water (filtration 1L). PCR	Mo-Bio PowerSoil® kit	HF183F BAC32F (BAC708R)	• 55% of samples contained BAC32 • 85% of positive samples contained HF183	• 15% increase in positive samples after storms at surface water inputs • UV and tides affect short term FIB concentrations

Footnotes:

- "Faecal samples" include samples of human and animal origin
- BAC708R is the universal primer assumed to have been used in most studies unless specifically listed or listed otherwise
- Asterisk (*) indicates countries where tests were performed in a different country as to where samples were taken

2.3.1 Sample collection and preparation

Raw sewage, sewage tank seepage, and solid faecal matter were often used as a positive control or as a comparative sample when testing environmental water samples. The assumption was usually made that the target molecular marker or faecal indicator organism would definitely be present in these samples. Faeces was preferably collected as fresh as possible before a slurry was made from which DNA could then be extracted (Walters & Field, 2009; Raith *et al.*, 2013). The desired amount of water from an environmental source was most commonly collected via a grab method as used by 93.5% of studies listed in Table 2.3-1. The literature reviewed here indicates that electronic water sampling systems as well as programmable in-line systems were also made use of (Nshimyimana *et al.*, 2014). In-line systems were able to collect or filter water proportionally to the source, flow etc. and were therefore believed to be more effective by Sauer *et al.* (2011) and Mayer *et al.* (2016). Electronic or automatic sampling systems could be applied when samples were more difficult to obtain, as was the case with Nshimyimana *et al.* (2014) when collecting reservoir water, Zheng *et al.* (2013) taking samples of water from wells and Sidhu *et al.* (2013) collecting storm water runoff from storm drains during rainfall events.

Depending on the type of study and water type researchers were interested in, certain pre-preparations were necessary prior to DNA extraction. Water samples in particular could be be pre-filtered through glass fibre filters to remove rough particulate matter before membrane filtration takes place, as performed by Santoro and Boehm (2007) to remove dinoflagellates from marine samples taken after an algal bloom. Water samples were generally membrane filtered prior to DNA extraction. Pre-filtration could however lead to the underestimation of the presence of a particular marker, as bacterial cells which are attached to particulate matter in the water sample were excluded (Santoro & Boehm, 2007). The smallest volume filtered was 10 mL of sample dilutions by Riedel *et al.* (2014). Studies making use of different filtration volumes depending on sample type or sampling site included those by the following authors: Wade *et al.* (2010); Sauer *et al.* (2011) and Mayer *et al.* (2016). In multi-laboratory studies performed by four authors including Boehm *et al.* (2013), the volume of liquid that was filtered was used as an estimation of the bacterial content. A sample volume of 5 mL filtered was thus equivalent to a 1:10 dilution as seen in Table 2.3-1. The most frequently occurring filtration volume for the studies in Table 2.3-1 was 100 mL, being applied by 26.8% of researchers. The second most filtered volume in Table 2.3-1 was 500 mL, being used by 21.4% of studies, followed by 1 L used by 16.1% of researchers. In Table 2.3-1, 10.7% of studies filtered 50 mL of water prior to DNA extraction. Kirs *et al.* (2011); Mika *et al.* (2014) and Nshimyimana *et al.* (2014) filtered samples until no more liquid was being drawn through the membrane, after which the sample volume was recorded, ranging between 10 mL and 2 L across these studies.

Precautions should be taken when comparing samples with different starting volumes as this may affect the efficiency of assays and give different results (McQuaig *et al.*, 2012; Mayer *et al.*, 2016). The average volume filtered across all studies was 438.1 mL. Microcosm studies were applied in 6.4 % of the studies in Table 2.3-1. Raw water was used for microcosms by Walters and Field (2009); Ballesté and Blanch (2010) and Ahmed *et al.* (2014). Additionally, Ballesté and Blanch (2010) autoclaved a volume of river water intended to be used as a matrix in some of the microcosms in order to compare the effects of predation in environmental waters on the survival and persistence of *Bacteroides* markers. Artificial freshwater was one of the water matrices used in a limit of detection (LOD) study by Riedel *et al.* (2014). According to 9.7% of studies in Table 2.3-1, the use of artificial water, or performing small-scale microcosm studies, could not properly illustrate the complex relationships between *Bacteroides,* human-specific markers and the natural environment.

2.3.2 DNA extraction kits can influence downstream application

Diverse DNA extraction methods were applied in the studies listed in Table 2.3-1. A single type of kit was used for DNA extraction from both faecal reference samples (samples of known origin) and water samples by 70.5% of the studies. In 11.5% of studies listed in Table 2.3-1 samples were pre-treated for more efficient cell lysis and DNA extraction. Pre-treatments performed by researchers in Table 2.3-1 included, but were not limited to: physical lysis such as heat-cold cycles (Shahryari *et al.*, 2014), bead milling (Kirs *et al.*, 2016), as well as chemical treatments such as the use of propidium monoazide (PMA) (Seidel *et al.*, 2017). Six studies in Table 2.3-1 made use of two different kits for their faecal and water samples (Gawler *et al.*, 2007; Gourmelon *et al.*, 2007; Van De Werfhorst *et al.,* 2011; Ahmed *et al.,* 2012; Aslan & Rose, 2013). A Mo-Bio UltraClean ® soil kit was applied by Nshimyimana *et al.* (2014) for DNA used in endpoint PCR assays whilst a Mo-Bio UltraClean ® plant kit was applied for samples intended for qPCR analysis.

Certain DNA extraction kits are optimised for use with a specific sample type in order to efficiently extract DNA. When using a kit with samples it was not designed to extract and purify DNA from, it could limit the efficiency of the extraction process (Ballesté & Blanch, 2010; Nshimyimana *et al.*, 2014). It was thus important that researchers consider the environment from which samples were taken and bear knowledge of the effects of inhibitory substances on DNA extraction efficiency before selection of kits.

Making use of different extraction kits and protocols may in turn have affected assay results (Ahmed *et al.*, 2009[A]; Zheng *et al.*, 2013). Post-extraction DNA treatment with ethanol for the

removal of inhibitors was performed by Van De Werfhorst *et al.* (2011). Although this step sometimes resulted in more efficient DNA product, ethanol itself could also inhibit molecular assay performance and had to be applied with caution (Green & Field, 2012). As seen in Table 2.3-1, McQuaig *et al.* (2012) and Napier *et al.* (2017) combined elements from two different extraction kits in their studies in order to maximise the yield and quality of the DNA product.

Multi-laboratory studies were performed by Boehm *et al.* (2013), Ebentier *et al.* (2013), Layton *et al.* (2013) and Raith *et al.* (2013), where participating laboratories made use of different extraction methods including commercial kits and physical-chemical extraction methods such as phenol chloroform extraction. Although kit extraction seemed to be the more popular and more efficient DNA isolation method, direct chemical extraction also held its merits. Participating laboratories did not report the same results in all cases where the same extraction method was used across laboratories. Ebentier *et al.* (2013) reported notably higher inter-laboratory coefficient of variance percentages (%CV) when methods and reagents were not standardised across participating laboratories.

When studying Table 2.3-1 one can see that various extraction kits were used in the different studies. The two most commonly used extraction kits were the Qiagen DNeasy Blood & Tissue kit and the Mo-Bio PowerSoil® kit, being applied in 28% of studies. The second most commonly used kit in Table 2.3-1 was the QiAmp stool DNA kit which was used in 10.52% of studies in Table 2.3-1. Concentration and quality of extracted DNA varied greatly across studies, however stool and soil kits containing reagents which aid in removal of humic acids and other potential inhibitors were favoured by most authors in Table 2.3-1 when working with environmental water and faecal samples.

Wade *et al.* (2010) and four other authors in Table 2.3-1 did not use commercial DNA extraction kits in their studies. Bead milling and centrifugation techniques were applied by Wade *et al.* (2010) and Heaney *et al.* (2014). A combination of physical bead milling as well as chemical extraction was applied by Zheng *et al.* (2013), while Mayer *et al.* (2016) extracted DNA via chemical means making use of phenol chloroform extraction. Although these DNA extraction methods were cost effective and quick and easy to apply, the advantages came at the cost of extraction efficiency and DNA quality. Depending on the intended use of the DNA after extraction, the type of extraction kit or method had to be carefully considered. There was thus a growing need for the optimisation of extraction protocols and assays used for the detection and quantification of *Bacteroides* markers (Staley *et al.*, 2012[A]; Ebentier *et al.*, 2013). This was an important step that may have resulted in biases on the data produced as well as the interpretation thereof, and the value derived from comparative studies.

2.3.3 Sample size

Data may be limited by small sample sizes in studies, making definitive conclusions difficult. The studies summarised in Table 2.3-1 had a wide range of sample sizes. Sample size may differ by the physical weight and volume of individual or composite samples as well as the amount of samples occurring in a sample set. Larger sample volumes were necessary for effective assay performance, particularly when water samples were being filtered (Sercu *et al.*, 2011; Kabiri *et al.*, 2016; Napier *et al.*, 2017). Individual faecal samples were assessed separately by 31 of the 62 studies in Table 2.3-1.

Sample freshness, particularly pertaining to faeces, may also affect the total DNA compilation and yield from a sample. Assay efficiency were highest when making use of fresh faeces (Layton *et al.*, 2013) that had not been exposed to environmental factors for a prolonged period of time (Van De Werfhorst *et al.*, 2011). Making use of older samples could thus mean that twice the sample amount would be necessary to obtain similar data as when fresh faeces was used, due to natural degradation of DNA. Composite faecal samples, wastewater, or environmental water from various sources were assessed by 74.2% of studies in Table 2.3-1. It was concluded that making use of combined samples increased the chances of more accurate marker detection within a community as these markers were not necessarily present in all individuals of a species (Kabiri *et al.*, 2016).

Odagiri *et al.* (2015) stated that *Bacteroides* markers varied in presence and abundance between individuals of the same host group. Due to varying marker abundance, authors such as Santoro and Boehm (2007) suggested making use of composite faecal samples for indicating trends in persistence, detection and relationships between *Bacteroides* and traditional FIB in water. Small sample sizes (referring to the number of samples available) have been listed as a limiting factor in *Bacteroides* associated marker studies by several authors in Table 2.3-1 (Harwood *et al.*, 2009; Shanks *et al.*, 2010; Boehm *et al.*, 2013; Odagiri *et al.*, 2015; Napier *et al.*, 2017). Ebentier *et al.* (2013) found that the comparison of unequal sample sizes, as used in the majority of studies in Table 2.3-1, significantly influenced the statistical comparison of data pertaining to *Bacteroides* markers. On the other hand, small sample sizes made no significant difference to data obtained by Zheng *et al.* (2013). Small sample sizes were seemingly not a problem in the study as a wide range of physical-chemical data was also collected for each site (Zheng *et al.*, 2013). In order to eliminate errors in data from varying sizes of samples and sample sets, standardisation of these assays had to occur. Additionally more efficient filtration methods for concentration of larger water samples needed

to be developed to aid in effective marker detection from potentially contaminated water bodies (Staley *et al.*, 2012[B]).

The amount of DNA obtained from the samples could vary, depending on the original concentration of *Bacteroides* DNA present in the sample (Kabiri *et al.*, 2016), as well as the extraction method that was used, and whether inhibitory substances such as humic acids were present (Staley *et al.*, 2012[A]). DNA was extracted from an average faecal weight of 285.2 mg of faeces per sample in 29 of the studies in Table 2.3-1 that made use of fresh faecal samples. The smallest sample weight used for DNA extraction in Table 2.3-1 was 2 mg (Nshimyimana *et al.*, 2017) where the largest was 1 g (Green *et al.*, 2014).

2.3.4 DNA quality and quantity

Prolonged storage of samples or storage under incorrect conditions may be detrimental to sample quality. Although sample collection and preparation, sample size, as well as choice of extraction method (kit, physical or chemical) may impact DNA quality and quantity, they were not the only influential factors. Inhibitory substances present in environmental water and other sample types (soil, faeces, and plant material) adversely affected molecular based assays. The relationship between inhibitors and their effects on certain samples was not all that well understood (Green & Field, 2012), but they may have interfered with both DNA extraction and downstream applications such as PCR (Walters & Field, 2009).

Table 2.3-1 indicates that Aslan and Rose (2013) found inhibitory substances in animal faecal samples, as did Heaney *et al.* (2014) in sand and turbid water samples. The effects of inhibition could have been negated or lessened by making use of specific DNA extraction kits formulated for plant or soil samples (for the removal of humic acids), or if additional ethanol wash steps were performed on extracted DNA to the same end (Van De Werfhorst *et al.*, 2011). Staley *et al.* (2012[A]) stated that tenfold or hundredfold DNA dilutions were an acceptable means of reducing inhibition as the samples were not extensively altered by this process. Relatively high levels of PCR inhibition were observed in 25% of faecal samples tested by Harwood *et al.* (2009). The majority of the inhibited samples were from cattle faeces (Harwood *et al.*, 2009). Staley *et al.* (2012[A]) experienced similar difficulties with 36% of their cattle faecal samples showing evidence of inhibition. Performing a tenfold dilution on the aforementioned samples, as well as samples taken for studies by Toledo-Hernandez *et al.* (2013) and Seidel *et al.* (2017) reportedly remedied the inhibition problem. Raw animal faecal samples tested by Ahmed *et al.* (2009[A]) also contained PCR inhibitory substances. The effects of these inhibitors were lessened or removed by making use of tenfold dilutions of the DNA prior to use. Ahmed *et al.* (2010) found it necessary to make use of hundredfold DNA dilutions to negate the effects

of PCR inhibitors present in their environmental samples. Thus, several authors used dilutions to negate PCR inhibition caused by co-extracted substances.

Table 2.3-1 indicates that PCR inhibition of samples tested by Nshimyimana *et al.* (2014) could not be resolved by dilution as PCR yields became undetectable during the dilution process. Internal amplification controls such as salmon testes DNA may be used to determine whether qPCR inhibition will occur. Salmon DNA was added to the samples in known quantities and therefore served as reference for relative efficiency of DNA recovery from a specific sample (Wade *et al.*, 2010; Aslan *et al.*, 2015). Four studies made use of Salmon testes DNA in Table 2.3-1. Finally, calibrators or external controls may also be applied to determine the quality of DNA with regard to recovery and purity. As with internal controls a known amount of DNA template was used, but is in this case treated as a sample with regard to further analysis (Wade *et al.*, 2010; Aslan *et al.*, 2015). It was recommended that both an internal as well as external control be used, as internal controls would produce a signal even if the target sequence was not present in the sample (Aslan *et al.*, 2015). Wade *et al.* (2010) as well as Heaney *et al.* (2014) made use of calibrators in their research. It is further reported that HF183 markers for end-point PCR seemed to be more susceptible to inhibition by humic acids than the other comparative markers applied in a study performed by Green and Field (2012).

Sample freshness played an integral role in analyses where microbial quantification data was collected. Ideally samples should have been collected and stored on ice for analysis within 6 hours of collection. Samples had to be analysed within a maximum of 24 hours unless special precautions were taken for shipment to other laboratories, where sampling had to take place within a maximum of 48 hours (Venter, 2000). DNA quality was influenced by physical and chemical factors, and naturally deteriorated over time, even if stored under the proper conditions. Prolonged storage influenced detectability of molecular markers as well as the accuracy of assays being performed on samples that had been stored for long periods of time, such as the samples taken by Napier *et al.* (2017), which were stored between 2-6 years before analysis. Shanks *et al.* (2010) analysed all DNA extracts in their study within 12 months of collection. Although it was generally accepted that samples may be stored for a maximum of three months (Odagiri *et al.*, 2015), performing assays on freshly extracted DNA was the ideal.

2.3.5 Analysis method (PCR/qPCR)

Sensitivity and specificity of different markers and assays may vary across the same sample set. Several factors could have influenced the detection and quantification of genetic markers. The occurrence and abundance of markers within a certain host group influenced results

depending, in turn, on other factors such as data interpretation methods and environmental parameters (which will be discussed later on). When studying Table 2.3-1 one can see that 48.4% of studies made use of quantitative PCR. Presence-absence data was collected via singleplex end-point PCR by 30.64% of the studies. Both end-point as well as qPCR assays were applied by 8.06% of the authors in Table 2.3-1. A risk assessment was performed by Boehm *et al.* (2015) making use of previously collected qPCR data. Touchdown PCR was performed by 2 studies in Table 2.3-1 (McQuaig *et al.*, 2012; Nshimyimana *et al.*, 2014). Two authors, Nshimyimana *et al.* (2014) and Shahryari *et al.* (2014), performed Nested PCR assays in their respective studies. The only study in Table 2.3-1 making use of Droplet digital PCR was Cao *et al.* (2015).

Different assays were selected by individual researchers as the final purpose for the application of their collected data was not the same. Presence-absence data collected via end-point PCR were intended to indicate the presence of certain molecular markers in water, but did not yield any quantitative data. According to Raith *et al.* (2013) the inability to normalise end-point PCR data complicated interpretation of results. Conventional PCR assays have been modified in several ways to increase assay specificity or product yield. Touchdown PCR is a method which may be applied when assays yield smaller non-specific bands during amplification. In order to save time and reagents, Touchdown PCR may be performed instead of testing all the separate parameters of a PCR protocol to eliminate the effects of mispriming samples or inhibition, thus ensuring better sample yields (McQuaig *et al.,* 2012). Touchdown PCR was applied by Harwood *et al.* (2009) in an attempt to cancel out external factors such as PCR inhibition caused by different sample matrixes during their study.

The nested PCR technique negated the need to run multiple PCR assays with increasingly specific primers, as the product of the first assay is used as the template for the second (Hernández-Rodríguez & Ramirez, 2012). There was also reduced risk of contamination of reaction mixtures when using nested PCR (Hernández-Rodríguez & Ramirez, 2012). Although both a touchdown and semi-nested PCR approach was applied by Nshimyimana *et al.* (2014), the nested PCR yielded better results. Dilution of samples to reduce PCR inhibition negatively affected the touchdown PCR assay, resulting in a higher proportion of samples tested by Nshimyimana *et al.* (2014) yielding negative results. In order to determine whether *Bacteroidales* markers would be applicable to drinking water monitoring, FIB were compared with a nested PCR approach (Shahryari *et al.*, 2014). Shahryari *et al.* (2014) found that the nested PCR approach was tenfold more sensitive than the culture method it was compared to. The bacterial primer set 27F and 1492R was used as the extended primer to confirm the presence of bacterial DNA, whereafter the universal *Bacteroidales* 16S primer set (BAC32F

and BAC708R) was applied by Shahryari *et al.* (2014) to confirm the presence of *Bacteroides* species. Shahryari *et al.* (2014) noted that the presence of environmental *Bacteroidales* species in water samples could lead to higher, erroneous, detection rates than that of indicators of faecal origin in the samples.

Fluorescent probe markers HF183Taqman and HF183SYBR for the detection of human associated *Bacteroides* species have been developed to detect and quantify and human associated faecal pollution. A total of 36 studies in Table 2.3-1 made use of qPCR techniques. Regarding the quantification and limits of detection of qPCR assays the following observations were made: Samples may have yielded a positive result, but did not give a strong enough signal during amplification in order to quantify the copy numbers present. According to Riedel *et al.* (2014) the LOD for quantitative PCR methods was usually lower than that of end-point PCR, but the LOD could be linked to other factors influencing detection of genetic markers and was generally assay specific. In another example found in Table 2.3-1, HF183 markers were only quantifiable up to a 10^{-2} dilution by Staley *et al.* (2012[A]) but were still detectable up to 10^{-6}. Although 58% of samples from one of two sampling sites studied by Mika *et al.* (2014) contained HF183 markers, none were quantifiable. Furthermore only 14% of samples from the other site tested by Mika *et al.* (2014) contained HF183 markers, of which a mere 4.6% were quantifiable. DNA degradation could be a possible explanation for samples being detectable but not quantifiable by qPCR. In both marine and freshwater samples, Kirs *et al.* (2016) found that markers could be quantified for up to 11-15 days after microcosms were inoculated. Markers were however still detectable for approximately twice as long (Kirs *et al.*, 2016). Green *et al.* (2014) reported that the LOD in the 95% of their samples that contained HF183 markers, was tenfold lower when making use of the BacR287 reverse primer, than when the BFDrev primer was applied. Increased sensitivity of HF183/BacR287 allowed the primer pair to bind to markers from a larger number of *Bacteroidetes, Bacteroides* and uncultured *Bacteroides* species due to more hybridisation taking place (Green *et al.*, 2014).

Cao *et al.* (2015) found that *Enterococcus* and HF183 markers could not be duplexed in conventional qPCR, however no interference between targets occurred when making use of duplex ddPCR. Droplet Digital PCR (ddPCR) partitions the sample by mixing it with oil to create smaller sample “droplets”. As this method counted the frequency of positive results within the partitions, no comparative quantification cycle data was necessary, potentially increasing precision of this technique when compared to qPCR. Morisett *et al.* (cited by Cao *et al.*, 2015) believed that this type of assay may be a cost effective technique for simultaneous quantification of multiple targets. Furthermore ddPCR may potentially be more robust against inhibition making it ideal for use with wastewater, faecal and environmental water samples

(Cao *et al.*, 2015). As with any method, ddPCR assays had certain limitations. Dilution of highly polluted samples were still necessary as ddPCR had a lower upper quantification limit (UQL), determined by the number of droplets present per reaction. Furthermore Cao *et al.* (2015) noted that the detection limits of ddPCR varied per reaction as more droplets increased sensitivity of a ddPCR run. Lastly faulty partitioning due to genes possessing or sharing multiple operons could lead to underestimation of copy numbers if the DNA segment containing both operons were partitioned into the same droplet (Cao *et al.*, 2015). Partitioning issues however only posed problems if operons occured within less than 50 kb from one another on the genome (Cao *et al.*, 2015).

Two studies in Table 2.3-1 reported that assay sensitivity may be influenced by amplification efficiency as well as the copy numbers of targeted genes (Lee *et al.*, 2008; Jenkins *et al.*, 2009). Along with sensitivity, specificity of assays could also vary across the same sample set (Dorai-Raj *et al.,* 2009). Inclusion or exclusion of certain data may have affected the accuracy of assays selected for specific studies (Green & Field, 2012). Layton *et al.* (2013) found that all assays applied in their study performed worse than originally reported in previous literature on the subject. An assay seemed effective until it was optimised or adapted for specific samples, whereafter the same assay yielded different results (Green *et al.*, 2014). Multiplex assays with an internal control were run by Green *et al.*, (2014), but they suggested the use of singleplex format to further test assay efficiency. Although the same reagents were used for both HF183SYBR and HF183Taqman assays by Riedel *et al.* (2014), results for the HF183SYBR assays were inconsistent. Napier *et al.* (2017) reported variations in marker detection depending on which assay was applied. Eleven authors in Table 2.3-1 suggested using assays in a toolbox approach in order to optimise the usable data obtained from a specific study (Gourmelon *et al.*, 2007; Harwood *et al.*, 2009; Ballesté *et al.*, 2010; Van De Werfhorst *et al.,* 2011; McQuaig *et al.,*2012; Sidhu *et al.*, 2012; Zheng *et al.*, 2013; Gordon *et al.*, 2013; Ahmed *et al.*, 2014; Kabiri *et al.*, 2016; Sowah *et al.*, 2017) It is indicated in Table 2.3-1 that Napier *et al.* (2017) deemed the use of quantitative methods better than detection methods for risk assessment when looking at beach-water quality and microbiological safety of recreational water to swimmers.

Correlations or relationships between host-specific *Bacteroides* markers and traditional Faecal Indicator Bacteria (FIB) including *Enterococci*, *Escherichia coli* and enteric viruses were also applied regularly to evaluate the detection and quantification efficiency of a particular assay. High detection efficiency was demonstrated in a study by Ahmed *et al.*, (2009) where HF183 markers could still be detected in a 10^{-9} dilution in which no culturable FIB could be detected. Significant correlations of the presence or absence of HF183 marker to *E. coli* and *Enterococci*

concentrations were found by Ahmed *et al.* (2008[B]). Riedel *et al.* (2015) observed strong correlations between HF183 and BacHum markers and relatively strong correlations between HF183 markers and total coliform as well as *Enterococci* presence. Bonkosky *et al.* (2009) reported significant, but poor relationships between *Enterococci*, faecal coliforms and HF183 markers near a coral reef in Puerto Rico. Similarly very poor to no correlation between the occurrence and/or abundance of FIB and *Bacteroides* markers existed in assays performed by 20.63% of studies in Table 2.3-1.

Layton *et al.* (2013) concluded that the presence of other markers did not seem to affect the sensitivity of HF183 markers. Assay specificity was higher when samples that could not be quantified by Layton *et al.* (2013) were considered negative. The presence of similar concentrations of FIB and *Bacteroides* markers in water samples was relayed by 33% of studies in Table 2.3-1. Detection of HF183 was the same as that of the levels of *E. coli* present in samples tested by Chase *et al.* (2012). The occurrence of gastrointestinal illness reported to Wade *et al.* (2010) by swimmers on selected beaches was higher on days where the *Bacteroides* calibrator cell equivalent values were high. This pattern was observed for high levels of *Enterococci* as well (Wade *et al.*, 2010). Higher detection of *Bacteroides* markers were also documented by both Chase *et al.* (2012) and Kirs *et al.* (2016), who detected HF183 more frequently than human-polyomavirus (HPyV) and found the markers more sensitive than detection and quantification of *Enterococci* and *E. coli*, respectively.

Singleplex qPCR assays making use of SYBR intercalating dye were performed in 16.7% of the studies applying qPCR techniques in Table 2.3-1. As little as one gene copy per reaction could be detected by the HF183 assays performed by Ahmed *et al.* (2009[A]). Human factor markers could be detected in spiked freshwater samples in their study up to a 10^{-8} dilution. The detection and quantification of markers and faecal indicators by Ahmed *et al.* (2009) may have been influenced by the use of different DNA extraction methods and PCR cycling parameters, underestimating the effectiveness of marker detection in fresh sewage with certain markers. Varying concentrations of markers within the gut of host species and the environment, complicate the use of these markers for correlations between them and known enteric viruses and pathogens used as traditional faecal indicators (Ahmed *et al.,* 2009[B]). Variations in marker abundance in assorted sample types were represented by the following studies which appear in Table 2.3-1: Of 20 environmental water samples taken after rainfall events, which were tested by Ahmed *et al.* (2010), 14 contained HF183 markers in numbers ranging between 3.9×10^4 to 6.3×10^7 genomic units per 100 mL of water. Only 16 of 22 human faecal samples tested for HF183 markers by Ahmed *et al.* (2012) yielded positive results. Ahmed *et al.* (2012) found the smallest number of HF183 markers in secondary effluent at 1.8

x 10^2 to 4.6 x 10^3 gene copies per 10 mL followed by 7.0 x 10^5 to 1.5 x 10^7 gene copies per 10 mL for primary effluent. The largest number of marker gene copies obtained by Ahmed *et al.* (2012) was that of raw, wet weight human faeces at 63.8 x 10^6 to 9.8 x 10^7 gene copies per 10 mL. Quantitative PCR assays performed by Zheng *et al.*, (2013) had lower efficiencies than previously reported results, possibly due to primer structure leading to complex secondary structures when binding with the product during amplification, thus causing false negative results. A large number of samples were analysed by Zheng *et al.* (2013) via singleplex qPCR, but multiplex was suggested as a faster, more effective alternative.

Singleplex qPCR assays making use of Taqman hydrolysis probes were applied in 26.6% of the studies in Table 2.3-1 making use of qPCR techniques. Quantitative PCR was performed on HF183 markers by Chase *et al.* (2012). The detection limit for samples in this study was 10 gene copies per 100 mL. Some of the sampling areas were heavily polluted, with obvious sewage odours, indicating the potential source of contamination. The highest number of HF183 detected at a particular site was 2360 copies per 100 mL. Chase *et al.* (2012) found that the frequency of detection was strongly linked to the specific locations selected for sampling. Of the samples tested by Ebentier *et al.* (2013), 88.9% of human faecal samples taken fell within the range of quantification. The remaining percentage was not detected. Various parameters for the measurement of faecal contamination in water samples were compared by Ervin *et al.* (2013) who found significant variation in the results, depending on which performance metrics were used for their assays. Performance metrics will be discussed in greater detail later on. Haack *et al.* (2013) found similar variations in marker concentrations depending on sampling area and environmental factors of that region. The average number of gene copies per 100 mL of water was 3797 ± 8938 for HF183 which represented roughly 14% of all samples taken by Haack *et al.* (2013) which tested positive for universal *Bacteroides* markers. Heaney *et al.* (2014) collected 144 sand samples and tested for *Bacteroides* species and *Bacteroidales* markers, yielding positive results in 68% and 53% of samples respectively. Correlations between the sand and their corresponding water samples from each site were highest for *Bacteroides* species cell calibrator equivalents (CCE) (Heaney *et al.*, 2014). The fourth highest correlation between sand and water samples taken by Heaney *et al.* (2014) was that of *Bacteroidales* species CCE (R^2 = 0.30; $p < 0.0426$). Due to potential difficulties with optimisation of qPCR methods via the inhibiting effect of sand and turbid water samples, the methods applied by Heaney *et al.* (2014) were based on relative measurements of faecal indicators between the samples.

Simultaneous use of two assays, one applying hydrolysis probe (Taqman), and the other intercalating dye technology (SYBR), were performed by the remaining 6.7% of studies

making use of qPCR methods in Table 2.3-1. Assays using Taqman reagents were deemed to be more reproducible than other protocols used in a study by Ebentier *et al.* (2013). Riedel *et al.* (2014) found Taqman assays to have the lowest limits of detection in their study. Both Layton *et al.* (2013) and Odagiri *et al.* (2015) reported relatively low host specificity for both Taqman and SYBR probes, where Boehm *et al.* (2013) reported varying specificity for SYBR. The second LLOD reported by Riedel *et al.* (2014) in Table 2.3-1 was for SYBR assays.

The problem however, was that no distinction could be made between viable and non-viable cells with conventional PCR or qPCR (Ballesté & Blanch, 2010). Propidium monoazide (PMA) could be applied for this goal (Seidel *et al.,* 2017). Due to the variability in qPCR applications, marker sensitivity and specificity, and other factors influencing the outcomes of molecular assays a toolbox approach making use of more than one genetic marker or assay was recommended (Ahmed *et al.*, 2014). A relativity scale (baseline-ratio comparison) may also be applied to samples taken upstream, at the suspected point of contamination, and downstream thereof (Lee *et al.*, 2010). Riedel *et al.* (2014) suggested that basic presence-absence be used to indicate faecal pollution until such time that assays were standardised and comparison and interpretation of results simplified.

2.3.6 Survival factors

The detection of host specific markers from the environment may vary significantly due to differences in persistence of the bacteria and their genetic material. It was found that host-associated markers were not always present in all individuals of a species. It can be seen in Table 2.3-1 that Kabiri *et al.* (2016) also concluded that markers were not present in all individuals of a host species. Ahmed *et al.* (2009[A]) and Ahmed *et al.* (2012) noted that *Bacteroides* made up the larger portion of gut microbes in mammalian species; consequently HF183 markers may have occurred in concentrations of 3 to 4 orders of magnitude higher than non-target markers. Abundance in the gut may be influenced by the host's intestinal physiology, diet, and geographic location (Ebdon *et al.*, 2007; Ervin *et al.*, 2013). Large variation in marker abundance in individuals was reported by Stea *et al.* (2015). Physical and chemical parameters within the environment could have further influenced the abundance and survival of *Bacteroides*. Chemical parameters in particular, influenced the presence and abundance of *Bacteroides* in water in a study performed by Mayer *et al.* (2016) During this study the effects of chlorination were not investigated (Mayer *et al.*, 2016), as chemical water disinfection does not take place consistently across Europe. Lee *et al.* (2010) noticed that markers persisted for quite some distance before dilution or other factors came into play and rendered them undetectable (Lee *et al.*, 2010). In a study comparing human and animal faecal pollution in karst water aquifers across Midwest U.S.A., it was found that human markers

occurred more frequently in wells where bovine markers occurred more frequently in spring-water (Zheng *et al.*, 2013). Alarmingly Mayer *et al.* (2016) detected similar concentrations of *Bacteroides* markers prior to, as well as after activated sludge treatment of sewage.

2.3.7 Factors influencing survival, persistence and detection

Low marker specificity may be caused by horizontal transfer of *Bacteroides* species or markers between different host species. Although intestinal physiology and host diet played a role in host specificity, homology in host physiology could lead to survival of *Bacteroides* species in other hosts (Zheng *et al.*, 2013). Gawler *et al.* (2007) stated that marker specificity may vary across numerous animal species in different countries and geographical regions. Markers from target species (human and bovine) as well as a range of non-target animal markers were tested by 49.2% of studies in Table 2.3-1. The highest occurrence of non-target amplification across the studies comparing target and non-target markers in Table 2.3-1 was at 61.54% for dog faeces. Chicken faeces amplified markers in non-target sources in 45% of studies in Table 2.3-1. Animal species where horizontal transfer had been observed in Table 2.3-1 include, but are not limited to: sheep, horses and other ruminants, pigs, seagulls and other birds, raccoons, marsupials, porcupines and mongoose. Five studies in Table 2.3-1 reported the presence of CF128 in non-target faeces (Gourmelon *et al.*, 2007; Jenkins *et al.*, 2009; Kirs *et al.*, 2011; Toledo-Hernandez *et al.*, 2013) Interestingly both Kirs *et al.* (2011) and Toledo-Hernandez *et al.* (2013) documented the presence of CF128 marker in human wastewater samples. On the other hand Gourmelon *et al.* (2007) and Odagiri *et al.* (2015) both reported the presence of human markers in bovine faecal samples. It is evident from Table 2.3-1 that 3 of 5 studies making use of both Taqman and SYBR assays reported the presence of these markers in non-target faeces. Although sensitivity and specificity values for human associated primers varied greatly in Table 2.3-1, SYBR assays were less specific than their Taqman counterparts across the board. The highest reported sensitivity and specificity value in Table 2.3-1 was 100% with the lowest being 30% for sensitivity and 28% for specificity of host associated primers.

Bacterial transfer between host and non-target species may be due to the close proximity people have to house pets and domestic animals faeces samples from humans (sewage) and domestic animals were easiest to obtain and therefore most commonly sampled by studies in Table 2.3-1, however, a few studies compared rarer faecal samples such as porcupine and mongoose to wastewater and other faecal sources in their research. Kirs *et al.* (2011) and Ervin *et al.* (2013) highlighted the necessity of validating methods with a wider range of faecal sources which may occur in a specific environment, in order to eliminate potential false positive results and increase assay performance. McLain *et al.* (2009) mentioned that cross

amplification of non-target sources could be excluded if care was taken to design specific primers and give careful consideration to the DNA extraction process as well as to the type of PCR and PCR parameters that were intended to be used in a study. Evaluating qPCR assays according to sensitivity and specificity of the markers and the relative reliability of these assays highlighted the fact that great variation existed. With regard to *Bacteroides* associated host specific markers, sensitivity referred to the presence of a marker in faeces of all or the majority of hosts within the host group, and specificity referred to the ability of an assay to exclude non-target faeces (Ridley *et al.*, 2014). High sensitivity (and specificity) was reported by 7 authors in Table 2.3-1. High specificity was documented by Sauer *et al.* (2011). Low sensitivity and high specificity occurred in results published by Boehm *et al.* (2013), which varied across laboratories involved in the study. It was observed that HF183 primers specifically detected *B. dorei*, while HF134 primers were less specific and detected *B. dorei, B. vulgatus* and *B. massiliensis*, leading to the conclusion that a wider range of host species may be detected (Kabiri *et al.*, 2016). Table 2.3-1 indicated that Kirs *et al.* (2011) found HF183 markers to be more sensitive than other markers used in their study. Although assays performed by Shahryari *et al.* (2014) were effective, the use of the general BAC32F primer may have led to overestimation of risk as environmental strains of *Bacteroides* could be detected along with the faecal strains of interest.

The researcher's choice or design of reverse primer may also have affected the sensitivity and specificity of an assay. Dorai-Raj *et al.* (2009) reported reversed sensitivity and specificity values for their assays depending on which one of their two newly designed ruminant reverse primers was applied. Jenkins *et al.* (2009) found that pairing the CF128 primer with BacCow-305r yielded higher sensitivity results for cow associated assays. Gourmelon *et al.* (2007) provided evidence for the application of CF128 as a universal ruminant marker instead of a cow specific marker. This claim was supported by other studies such as the multi-laboratory study performed by Boehm *et al.* (2013). Sensitivity and specificity of Simplex qPCR was higher than that of both simplex and duplex ddPCR's performed by Cao *et al.* (2015), however all three assays would prove to be effective for detection and quantification of HF183 markers in environmental water.

It must be noted that no assay was 100% sensitive or specific, and variations in results and interpretation thereof may have been influenced by many factors (Shanks *et al.*, 2010). Ahmed *et al.* (2012) affirmed that, according to the US-EPA, a marker with a specificity of less than 80% was undesirable for application in microbial source tracking assays. When different types of human faecal pollution were compared, the HF183SYBR assay was most sensitive for sewage, followed by faeces, then septage (Van De Werfhorst *et al.,* 2011). Similarly Layton *et*

al. (2013) reported variations in sensitivity by sample type, however they found both HF183SYBR and HF183Taqman markers to be most sensitive for faeces, followed by septage, then sewage. The most reproducible assay performed by Ebentier *et al.* (2013) was the HF183Taqman assay.

2.3.8 Environmental factors

The relationships between environmental factors within environmental water cannot be replicated precisely in microcosm studies making comparison between the two difficult. A multitude of factors may influence the presence, persistence and detection of genetic markers from environmental samples. Some of these factors are: temperature, physical-chemical parameters of water, oxygen levels, UV exposure, predation and storm events. The effects of temperature on *Bacteroides* in environmental water are generally associated with different seasons. Lee *et al.* (2010) reported that all *Bacteroides* species were influenced by temperature and related environmental factors differently. Temperature was the main factor linked to natural DNA degradation regardless of species or specific markers being analysed (Ballesté & Blanch, 2010; Seidel *et al.*, 2017). The following authors in Table 2.3-1 indicated seasonal related differences in the detection of *Bacteroides* markers: Stea *et al.* (2015) reported increased levels of HF183 and ruminant markers in July to August at water temperature of 19°C. *Bacteroides thetataomicron* levels were higher during summer months at water temperature of 22-29°C (Ballesté & Blanch, 2010). *Bacteroides fragilis* was detected more often in winter. Marker persistence however, was half that of winter months at 24-72 hours for both markers tested by Ballesté and Blanch (2010). Similarly Lee *et al.* (2008) reported increased detection of Bac32F markers in summer and elevated levels of CF128 markers in winter. Total *Bacteroides* numbers in a study pertaining to the influence of septic systems on faecal pollution were higher in summer while no seasonal variation was observed for human markers across the seasons (Sowah *et al.*, 2017). Riedel *et al.* (2015) found no seasonal variation in HF183 detection in their lagoon samples, but only detected HF183 in marine water during winter. Walters and Field (2009), citing Kreader 1998, stated that predation by protozoans in water was linked to temperature, with predation having a more profound effect in warmer conditions. Nshimyimana *et al.* (2014) came to the conclusion that HF183 markers, although effective, have the potential to survive longer in anoxic tropical environments which mimic the mammalian gut, which in turn could potentially lead to over estimation of risk. Toledo-Hernandez *et al.* (2013) concluded that seasonal variability was not well documented, particularly in tropical regions. Santoro and Boehm (2007); Sauer *et al.* (2011) and Mayer *et al.* (2016) reported no significant seasonal variation in detection and persistence of *Bacteroides* and associated markers with regard to environmental parameters of water in their studies. Green *et al.* (2011) focused attention on the fact that the intricate

relationships between physical and chemical parameters of water within the environment could not readily be replicated accurately within microcosm studies. Even with the limitations researchers experienced when applying microcosm studies, the generalisation could be made that increased temperatures were linked to higher DNA decay rates in water (Green *et al.*, 2011).

Several factors were probably linked to one another, resulting in greater variation within the environment. The ability of water to carry oxygen is influenced by the temperature of the water. Dissolved oxygen (DO) levels are generally higher in warmer water (Ballesté & Blanch, 2010). DO values and water temperatures for samples taken by Liang *et al.* (2013) containing CF128 and CF193 markers did not vary as much as in samples where the markers were absent. *Bacteroides* species were suggested as an alternative to FIB mainly due to their anaerobic nature and the assumption that they would thus have lower survival rates in oxygenated waters (Gourmelon *et al.*, 2007; Kirs *et al.*, 2011). Liang *et al.* (2013) documented higher detection rates of HF183 in water with lower levels of dissolved oxygen. When comparing the persistence of *B. fragilis* and *B. thetaiotaomicron*, Ballesté and Blanch (2010) found that *B. fragilis* persisted longer in colder months. They concluded that this occurrence was due to greater sensitivity to DO concentrations in warmer water. Walters and Field (2009) reported prolonged survival of *Bacteroides* species within anoxic conditions in the water column and potentially in sediments, where no mixing of the water took place. Survival within anoxic niches in the environment may have led to over or underestimation of risk due to differences in survival and persistence, and ultimately detection, of specific *Bacteroides* species depending on their direct environment (Walters & Field, 2009).

When predation was considered, temperature and salinity influenced the presence of predators. The effects of predation were markedly higher in warmer months due to predators being more active (Ballesté & Blanch, 2010). Persistence of *Bacteroidales* markers in the presence of predators varied greatly as reported by Walters and Field (2009). Markers were merely detectable for 1 day at 30°C, but up to 2 weeks at 4°C. Furthermore Walters and Field (2009) observed that removal of the predators via autoclaving the water, removed the related temperature effect described by Kreader *et al.* 1998.

The pH of water did not seem to have a profound effect on the persistence and detection of *Bacteroides* markers. Five studies in Table 2.3-1 lowered the pH of their samples to increase extraction efficiency of total DNA or viral DNA as a comparison to *Bacteroides* markers. In contradiction with the previous observation, Lee *et al.* (2008) stated that pH negatively affected

the presence of Bac32F markers in their study on the exposure of two watersheds to cattle faeces from farms with different management regimes.

Salinity of water samples had no major effect on the persistence and detection of *Bacteroides* markers in studies performed by Santoro and Boehm (2007) and Ballesté and Blanch (2010). In contrast with these results, five authors from Table 2.3-1 reported various effects salinity had on detection of faecal indicators. Wiegner *et al.* (2017) noted that all faecal indicators applied in their study increased with decreasing salinity. McQuaig *et al.* (2012) observed that FIB concentrations were low in samples with increased salinity or UV exposure, relative to the frequency of detection human associated markers. Walters and Field (2009) suggested that host specific markers should be evaluated in water with various levels of salinity as it had a negative effect on the persistence of CF128 markers in previously published field studies. Mika *et al.* (2014) noted that all of their samples which tested positive for HF183 markers had significantly lower salinities than those which tested negative. Interestingly Green *et al.* (2011) describe saline samples in their study as being more likely than fresh water samples to contain detectable levels of HF markers. Higher detection rates under saline conditions may be ascribed to slower decay of *Bacteroides* cells due to the effect of salinity on the survival of predators within the water (Green *et al.*, 2011). The persistence of HF183 markers in freshwater and seawater microcosms was compared to that of FIB by Ahmed *et al.* (2014). The amount of marker decreased from 8.2×10^6 genomic units per 100 mL in raw sewage samples to under the detectable limit over a number of days with no significant difference measured between fresh and seawater by Ahmed *et al.* (2014).

As with temperature, exposure to ultraviolet (UV) light can influence the survival and persistence of bacteria and molecular markers in environmental water. Mayer *et al.* (2016) remarked that UV exposure had no marked effect on the detection of human associated *Bacteroides* markers in wastewater samples. Mika *et al.* (2014) and 6 other authors in Table 2.3-1 mentioned UV irradiation as a factor which may influence the persistence and detection of *Bacteroides* markers in water. Walters and Field (2009) found CF128 markers to be more sensitive to photoinactivation at various wavelengths than the other markers applied in their study. According to Ballesté and Blanch (2010) reactive oxygen in water may be increased by the effects of sunlight inactivation, which in turn affected species with low aerotolerance. Both Green *et al.* (2011) and Kirs *et al.* (2016) concluded that UV irradiation has a greater effect on the culturability of bacterial species than directly damaging or modifying molecular markers used for detection purposes. Although most faecal indicator species were susceptible to the effects of solar radiation, Haack *et al.* (2013) noticed that increased turbidity of water, such as

the suspension of debris on the shoreline of beaches, could have protected bacterial cells and molecular markers from the effects of UV light.

Presence of markers in a water body may further have been influenced by the nature of the surfaces runoff water was exposed to, as well as the effects of storm events on a receiving water body pertaining to the volume and alteration of physical-chemical properties thereof. Of the 62 studies summarised in Table 2.3-1, 13 made mention of storm water or surface runoff influencing their ability to detect host specific *Bacteroides* markers. Higher detection rates of *Bacteroides* and *Bacteroidales* markers occurred at shin-depth water than at waist-depth water in a beach-water assessment study performed by Heaney *et al.* (2014), likely due to localised variations in temporal and environmental parameters within the study sites. Reeves *et al.* (cited by Wiegner *et al.*, 2017) reported that an estimated 99% of annual indicator organism load in water, specifically in drier climates, was due to stormwater runoff. Similar to the effects of waves on beaches noted by Green *et al.* (2011), the addition of large volumes of surface runoff potentially altered the physical-chemical parameters of the receiving water body. Debris or contamination from non-point sources after rainfall contributed to poor water quality with regard to physico-chemical parameters such as pH, turbidity, etc. as well as microbial load (Haack *et al.*, 2013; Nshimyimana *et al.*, 2014). Results from a study performed by Stea *et al.* (2015) highlighted an interesting trend where human markers were found, as expected, in base flow from urban areas whilst also being present in rural watersheds in periods with increased runoff. Stea *et al.* (2015) concluded that the time it took for runoff to reach the receiving water body had an influence on the type of markers present. Ahmed *et al.* (2008[A]) recorded an increase in false positive results after storm events as runoff washed large numbers of bacteria into the creek they were studying. Jenkins *et al.* (2009) and two other authors in Table 2.3-1 reported increased detection of cow specific markers of up to 40% after rainfall events. Poor or deteriorating infrastructure may further have contributed to the exacerbating effects storm drains had on water quality (Mika *et al.*, 2014). Sercu *et al.* (2011) noted that storm water runoff had a homogenising effect on water as large volumes of surface water carried FIB and host specific markers into water sources. Curiously, wet weather may also have caused the detection of lower concentrations of FIB and host specific markers. Dilution or dispersion effects were observed in a closed system such as a stagnant pool in an ephemeral riverbed, where additional runoff diluted the microbial population. Surfaces which are more porous absorbed more runoff water, consequently less contaminated water ended up in the receiving water body (Heaney *et al.*, 2014). Alternatively, the addition of large amounts of water to an impermeable surface such as concrete created large volumes of runoff which picked up possible sources of contamination along the way to the receiving water body, concentrating the amount of bacterial markers one could potentially detect. More recent

contamination could be detected after storm events, however source tracking could not be applied as readily as non-point sources of contamination would play a greater role under these environmental conditions (Kirs *et al.*, 2011).

2.3.9 Geographic stability

Before the application of an assay using a specific molecular marker, field testing for efficiency of the assay was paramount (Gawler *et al.*, 2007). Geological, environmental and aquatic parameters that vary across different regions influenced the abundance, survival and persistence of markers in specific geographic locations. Although some authors in Table 2.3-1 reported geographic stability of markers in their particular study regions, the need for location based validation during environmental studies was highlighted by 20.63% of studies in Table 2.3-1. Not only did gut microbial communities differ from country to country, but Nshimyimana *et al.* (2017) observed differences in assay performance across these countries as well. Some studies validated in tropical areas however propose a link between geography and latitude in favour of the widespread use of these markers (Nshimyimana *et al.*, 2017). Haack *et al.* (2013) found that regional factors influencing Great Lakes beaches play an equally important role in the microbiological quality of these lakes to the local factors influencing individual lakes. Areas of the Great Lakes display natural variations in climate, groundwater tables, and geological differences such as sand dunes or pebble beaches (Haack *et al.*, 2013). Furthermore, anthropogenic influences which varied from location to location as well as regionally had to be taken into account when studying a water body (Haack *et al.*, 2013). Although Wade *et al.* (2010) performed validation for use of selected assays within their area of study, they asserted that these methods may not be stable enough to apply as is in other climate types or where different contamination sources were present.

2.3.10 Data interpretation

Comparison of different detection methods can cause a lack of correlation in data. The use of a specific method, as well as the interpretation of collected data towards a specific end is key (Ervin *et al.*, 2013). In table 2.3-1 Gourmelon *et al.* (2007); Santoro and Boehm (2007); Ahmed *et al.* (2008[B]); Zheng *et al.* (2013) and Shahryari *et al.* (2014) directly compared data from different detection methods to one another. Total and faecal coliform ratios were compared to the presence or absence of HF markers by Santoro and Boehm (2007). Enumeration methods for FIB were applied by the remaining four authors from Table 2.3-1. Two of the authors compared different detection methods for the following indicator bacteria in their studies: *E.coli,* total coliforms, faecal streptococci, and *Enterococci*. Gourmelon *et al.* (2007) additionally compared the detection of F-specific coliphages to FIB and *Bacteroides* related

data collected in their study. Localised differences in samples from the same sample set, for example water from a creek being compared to raw wastewater, could influence the sensitivity and specificity of various assays performed on these samples (Dorai-Raj *et al.,* 2009; Jenkins *et al.*, 2009). Pollution from faecal sources could vary up to five orders of magnitude depending on which method of measurement was applied to the samples (Ervin *et al.*, 2013). Raith *et al.* (2013), Lee *et al.* (2008), and three other authors in Table 2.3-1 detailed the importance of clearly defining the conditions and parameters of the sampling process followed, as well as assays selected for data extraction. Parameter characterisation may also have been referred to as a performance, baseline, or benchmark threshold selection. Regardless of the terminology used, it was crucial to the successful interpretation of data from a study to select and clearly define these criteria beforehand (Lee *et al.*, 2010). When multiple parameters such as rainfall, physical chemical data and microbial quality such as tested by Mika *et al.* (2014) were compared, the inclusion or exclusion of data made a significant difference to the interpretation of the results. Data disregarded by Mika *et al.* (2014) due to recurring inhibition of samples served as a basic example of exclusion. Spiked water samples containing less than 8×10^4 copies per μL^{-1} of HF183 marker were reanalysed after being diluted twofold by Mika *et al.* (2014). Samples that still did not amplify after three repetitions of the dilution step were disregarded for the remainder of the study. From one site, 6 of 86 samples, and 24 of 176 from the other were excluded by Mika *et al.* (2014) due to the presence of inhibition. While considering another parameter, such as DNQ samples as either positive or negative, the following observation was made by Ahmed *et al.* (2012): if the interpretation of the DNQ results were taken as positive, none of their assays met the 80% benchmark for sensitivity and specificity. Similarly Layton *et al.* (2013) found that counting DNQ values as negative decreased the sensitivity for their assays slightly, but made a significant improvements to specificity across their data set.

False negative results may occur due to primer specificity being very high. Ebentier *et al.* (2013) reported lower reproducibility of HumM2 assays due to the primers targeting a single-copy functional gene instead of conserved genes within the 16S rRNA region. The likelihood of the presence of a specific genetic marker in selected water samples had to be considered before classification of results as true positive or negative (Jenkins *et al.*, 2009). Assay interference due to the presence of inhibitors in certain sample types, such as humic acids in samples tested by Green and Field (2012), also could have resulted in false negative results (Kirs *et al.*, 2011). On the other hand, false positive results may also have complicated data interpretation. According to Ahmed *et al.* (2012) markers which were not sensitive and specific enough to detect a target host group, may have led to the occurrence of higher levels of false positive results. Cross reactivity of markers between samples containing genetic material from

closely related hosts, or hosts with similar intestinal physiology, were a relatively regular occurrence (Layton *et al.*, 2006). Horizontal transfer of marker-bearing *Bacteroides* between hosts who coexisted in close proximity (such as pets and domestic animals) also caused false positive results (Ahmed *et al.*, 2008; Lee *et al.*, 2010). False positive results were reported by 41.3% of authors in Table 2.3-1, of which cross reactivity with dog faecal samples was the most common. Not only did pets have to be considered as sources of false positive results, but wild animals such as porcupines (Stea *et al.*, 2015) and mongoose (Kirs *et al.*, 2016) were examples of area specific species which had to be considered as possible faecal contamination sources before accepting results as true positive or negative. Furthermore, environmental strains of *Bacteroides* more tolerant to physical parameters such as oxygen also made it difficult for researchers to collect usable data (Walters & Field, 2009). These organisms may have the ability to survive for longer than expected in tropical environments or isolated niches such as anoxic sediments, and could have contributed further to false positive results, adversely influencing microbial source tracking data (Nshimyimana *et al.*, 2014).

Viability of cells played a large role in the occurrence of DNQ data. Bonkosky *et al.* (2009), Staley *et al.* (2012[B]), Nshimyimana *et al.* (2014), Sowah *et al.* (2017), and three other authors in Table 2.3-1 documented the fact that both culture based and molecular techniques had shortcomings related to the detection and quantification of organisms that may be viable but non culturable (VNBC), or have the ability to replicate in the environment. Culture based techniques would not detect organisms that became dormant due to damage or stress factors (Nshimyimana *et al.*, 2014). Molecular techniques may detect VNBC organisms, but could not distinguish between them, possibly yielding DNQ results for real-time PCR (Ballesté & Blanch, 2010). Additionally genetic material from cells that died, but had not yet undergone genetic decay, would be incorporated in quantifications as most molecular methods did not distinguish between live and dead cells (Ahmed *et al.*, 2009[B]; Green *et al.*, 2011). The development of PMA and similar treatments by Seidel *et al.* (2017) aimed to remedy this shortcoming of PCR based assays.

Finally, data interpretation was dependent on factors such as proficiency of laboratory technicians running the assays (Raith *et al.*, 2013), extraction and evaluation methods used (Mayer *et al.*, 2016), as well as the instrumentation individual labs were equipped with (Boehm *et al.,* 2015). Layton *et al.* (2013) reported worse overall performance than was reported in prior literature for all assays applied in their study. Inconsistencies in detectable dilutions which occurred across different sites and sampling dates were reported by Staley *et al.* (2012[A]). Even when making use of the same reagents across their sample set, Riedel *et al.* (2014), documented inconsistent results when making use of HF183SYBR, which made collection of

data difficult. Researchers needed to keep in mind that adaptation of methods and optimisation of assays could have significant effects on the data collected in their studies (Green *et al.*, 2014).

2.3.11 Suitability and limitations of application in developing countries

A growing need for water quality monitoring in developing countries was urging researchers to evaluate the potential for use of *Bacteroides* assays in developing countries. There were however certain limitations that would have to be overcome before successful application could take place. Financial and technical limitations were of the main concerns when exploring the possibility of validation of *Bacteroides* assays in developing countries. The need to set up a reference library for library-dependent studies increased a researcher's expenses (Ahmed *et al.,* 2009[B]). Furthermore certain steps, such as dilution of DNA, were necessary to obtain usable results, at the expense of a low budget as well as impairment of limits of detection or other data that may potentially be collected (Green & Field, 2012). Interestingly Riedel *et al.* (2014) noted that, although the initial start-up as well as instrument maintenance costs of qPCR methods were more expensive than those of endpoint PCR, running costs of SYBR based assays were lower than Taqman and endpoint PCR. Various other considerations pertaining to the decision to purchase PCR equipment was given in their study. Kabiri *et al.* (2016) determined that the cost of running conventional PCR assays were 3 to 4 times lower than making use of a toolbox approach (in 2016, when the study was conducted). Although Toolbox strategies however yielded more reliable results, they were estimated to cost up to 525 US dollars per sample at that time, depending on whether commercial facilities were contracted for their services (Kabiri *et al.*, 2016).

Even though adaptations such as dilutions or the use of other reagents could have been made to assay protocols, this may also have affected the suitability of these assays for application in certain geographic regions. Lee *et al.* (2010); McQuaig *et al.* (2012), Aslan and Rose (2013) and Gordon *et al.* (2013) reminded us that bacteria have the ability to adapt to their environmental conditions in order to survive. This in turn created the necessity for validation of assays in different geographic regions and under various environmental conditions before use. The importance of validating assays before application was stressed by 38% of authors in Table 2.3-1. Kirs *et al.* (2011) and Jenkins *et al.* (2009) both determined that it would be necessary to validate samples against a larger number of host groups as well as against more individuals within these host groups commonly found in the region being studied. Similarly, Ahmed *et al.* (2012) observed that the sample type should also be taken into account, for instance: making use of wastewater or composite animal faecal samples, as these were more

likely sources of pollution than faeces from a single individual. Ebentier *et al.* (2013) and Stea *et al.* (2015) remarked that even though these methods were assumed to be geographically stable and host specific, inter-laboratory validation of methods was crucial for effective application and collection of useful data.

As mentioned previously, validations were performed in order to determine whether an assay is applicable in a specific geographic region. Eight authors in Table 2.3-1 made mention of geographic stability in their studies. Ahmed *et al.* (2009[B]) reported stability of HF183 across different water matrices in their study area, but found that the presence of *Bacteroides* markers were not an effective indicator for enteric viruses in environmental water. Enteric viruses were only detected up to 10^{-5} dilution where HF183 markers could still be identified in 10^{-9} dilutions, indicating a large difference in concentration of viruses and *Bacteroides* in fresh sewage (Ahmed *et al.*, 2009[B]). As already stated, a combination of assays may prove to be the most effective means of microbial source tracking. Gourmelon *et al.* (2007) recommend the use of multiple animal-specific *Bacteroides* markers for identification of hosts, but deemed HF183 as sufficient for identification of human pollution sources. Moreover Gourmelon *et al.* (2007) also proposed that library-independent markers such as HF183 and CF128 were not geographically limited. Notwithstanding certain relationships having been observed between studies in tropical regions, Nshimyimana *et al.* (2017) asserted the necessity of assay validation due to recent findings related to intestinal microbial community composition and variability thereof in different countries.

Once validations have been performed, adjustments may be made to assays in an attempt to increase performance or efficiency. Assay modifications had been made by several authors in Table 2.3-1 including: modification in DNA extraction methods (Gourmelon *et al.*, 2007); adaptation of PCR cycling conditions (Nshimyimana *et al.*, 2014); altered primer or reverse primer sequences (Van De Werfhorst *et al.,* 2011) as well as the use of other quencher sequences in qPCR assays (Shanks *et al.*, 2010). McLain *et al.* (2009) advocated for modification of assays to increase stringency. Whilst assay performance may be improved by adaptations in protocols, Ebentier *et al.* (2013) voiced concerns regarding the knowledge and proficiency required of technicians in order to apply modifications that would be both efficient and successful.

Ultimately standardisation of protocols is necessary in order to make efficient use of these assays for microbial source tracking (Mayer *et al.*, 2016). Standardisation of samples to a specific benchmark threshold value such as making use of a specific concentration of DNA for all samples is a basic method that can be applied to obtain relatable data (Shanks *et al.*, 2010).

Interpretation of data comparing studies with many variables is much more difficult than making use of standard protocols (Ervin *et al.*, 2013). Green *et al.* (2014) and Ebentier *et al.* (2013) noted that standardisation increased the overall reproducibility of *Bacteroides* related assays. Subsequently, increased reproducibility of a standardised protocol rendered it suitable for commercial use (Harwood *et al.*, 2009). Layton *et al.* (2013) focused the attention on the fact that making use of standardised protocols was not a guarantee that the same results would be obtained as those in previous studies, or by other laboratories performing the same assay. Sauer *et al.* (2011) recommend relating standardised data to older methods in order to ensure the most efficient assay has been developed, as well as to assist in troubleshooting when results differ from expected.

Even when protocols are standardised, researchers are encouraged to apply these assays in toolbox strategies for most effective source tracking (Harwood *et al.*, 2009). Six studies in Table 2.3-1 made use of toolbox approaches. *Bacteroides* assays could be combined with assays for traditional indicators such as *E. coli*, faecal streptococci and pathogens such as *Salmonella* species, as compared by Shahryari *et al.* (2014). Gourmelon *et al.* (2007) suggested making use of multiple *Bacteroides* markers for determination of animal sources, but deemed HF183 sufficient for the identification of sources of human faecal pollution. Although there were no direct links between the occurrence and abundance of traditional FIB with *Bacteroides* markers (Ahmed *et al.*, 2009[A]; Ahmed *et al.*, 2009[B]; Mayer *et al.*, 2016), the general trend was established that high FIB levels and simultaneous presence of *Bacteroides* markers in a sample pointed to recent contamination by faecal sources (Mika *et al.*, 2014).

Fourteen studies in Table 2.3-1 confirmed the successful use of *Bacteroides* host specific assays in specific countries in their research. Shanks *et al.* (2010) stated that *Bacteroides* assays were suitable for use across the United States as long as they were validated regionally to compensate for abundance of specific markers in host sources, and which types of animal hosts were present within the geographic area. Developing countries did not always have the resources necessary to perform complicated or costly assays. Despite relatively high start-up costs, Nshimyimana *et al.* (2017) proposed investment in molecular based methodologies by governments and municipalities in developing countries such as Singapore. Mitigation costs could thus be reduced and faecal contamination in urban environments be better controlled, paying back the investment into molecular methodologies in the long run (Nshimyimana *et al.*, 2017). At this stage Nshimyimana *et al.* (2017) still recommended that assay validations be carried out prior to application on a specific geographic or environmental sample set.

Kirs *et al.* (2016) were aware that larger sample volumes or improved concentration techniques could benefit their study, but reported that human-associated markers can be detected successfully in tropical environments such as Hawaii. Similarly Toledo-Hernandez *et al.* (2013) found that end-point PCR methods were stable for use in Puerto Rico, but noted that further studies were required to fully understand the survival of *Bacteroides* in tropical regions. Once this has been achieved development of quantitative methods of detection tailored to tropical environments could take place (Toledo-Hernandez *et al.*, 2013). Although a sensitivity of only 65% was recorded by Jenkins *et al.* (2009), the claim was made that the assay was robust and reliable enough to apply for the detection of faecal pollution in water sources in Africa. Limited resources for on-site processing of samples was available in rural areas, and samples therefore had to be shipped away for testing (Jenkins *et al.*, 2009). Whilst making use of Bac32 assays are a step in the right direction, making use of this universal marker alone did not prove to be suitable for microbial source tracking in Iran (Shahryari *et al.*, 2014). Odagiri *et al.* (2015) set forth results indicating the possibility of the universal application of HF183SYBR. They claimed that Genbac assays were best suited for the detection of faecal pollution in water sources across India. Interestingly, lower numbers of *Bacteroides* markers were reported with illness and associated gastrointestinal symptoms such as diarrhoea than found in healthy patients' faeces, but further research into this topic was required (Odagiri *et al.*, 2015).

2.4 Conclusions

As demonstrated by the findings in Table 2.3-1, one can see that microbiological occurrence, abundance, and detection are influenced by a plethora of factors, environmental and otherwise. Through careful examination of the selected literature the following inferences can be made: A multitude of commercial DNA extraction kits exist and the type of samples being assessed need to be considered before selecting a specific type of kit. Various methodologies for detection and quantification of host-specific *Bacteroides* markers have been developed. Developments in PCR methods render host specific marker techniques faster than traditional culture based methods. Molecular methodologies can be optimised to increase accuracy and efficiency of assays. *Bacteroides* species make up the largest proportion of mammalian intestinal bacterial communities. Quantitative *Bacteroides* data may thus be applied in such a manner that the extent of pollution from a specific source can be estimated. Although there is no distinction between viable and dead cells when using molecular techniques, relative concentrations of *Bacteroides* DNA are linked to recent pollution due to their anaerobic nature. Subsequently low numbers or no detection of *Bacteroides* markers should indicate that no recent contamination has taken place. Live/dead PCR distinction technology by means of PMA

and other chemicals is in development. *Bacteroides* associated assays are in most cases used comparatively with traditional faecal indicators or in combination with various chemical, physical and molecular assays in a toolbox approach to increase the accuracy of data related to faecal contamination and the sources thereof in environmental water.

Researchers in Table 2.3-1 faced several challenges with regards to the application of *Bacteroides* host specific molecular assays. Correct sample collection, storage and preparation is important in regardless of sample and assay types. Sample preparation and DNA extraction may influence molecular applications downstream. A need exists for more efficient filtration methods with regards to extremely turbid samples as well as large volumes of water to increase the detection efficiency and concentration and quality of DNA which may potentially be extracted from filtered samples Sample weight or volume as well as the number of samples in a sample set need to be clearly defined in relation to the purpose of the data being collected. Fresh samples are required for effective (quantitative) comparisons as the abundance of *Bacteroides* markers present within the samples will not have been subjected to degradation due to exposure to environmental factors. Survival and persistence of *Bacteroides* and the subsequent detection of markers from environmental samples may be influenced by factors such as temperature and dissolved oxygen in the immediate environment as well as more widespread geographical factors such as the underlying geology of a specific region. It is generally accepted that *Bacteroides* do not survive for extended periods of time in the environment, but certain strains show minor adaptations for survival with regards to oxygen and temperature tolerance. Seasonal variations within the environment, including general traffic or human/animal activity should be taken into account for effective data interpretation. Surface runoff, particularly after storm events show increased detection of *Bacteroides* markers. Care should be taken that recent rainfall events and the resulting non-point source contamination due to increased volumes of runoff water into receiving water bodies is not falsely interpreted. The nature of samples should be considered carefully before selection of DNA extraction and purification kits in order to maximise the quality and quantity of DNA which can be extracted from each sample. Inhibition due to the presence of humic acids or other organic or chemical compounds must be considered. Even when making use of a DNA extraction kit designed to minimise the effects of inhibition, it is recommended to perform inhibition controls on environmental samples prior to or alongside intended assays for data collection. Horizontal transfer of *Bacteroides* between hosts may lead to cross amplification of markers in samples not associated with the expected host. Faecal samples from animal hosts that may influence the study area directly must be included in the study of the area in order to exclude possible false positive results. Specificity of markers may be further explored in order to develop more sensitive or specific markers. On the other hand, creating markers which

have an extremely narrow range of detection may lead to false negative results. Depending on the desired data, analysis methods should be chosen carefully. Presence/absence data is best obtained from end-point PCR where quantitative data is best collected by qPCR. Assays may be adapted to include touchdown, nested or digital PCR techniques in order to increase efficiency, repeatability, detection and specificity. Data interpretation with regards to shortcoming in the techniques should be considered before comparisons can be drawn between different methods or techniques that make use of different technologies. Examples of baseline threshold values which may influence data interpretation includes whether DNQ results will be regarded as positive or negative or whether DNA has been standardised to a specific test volume for assays. Adequate knowledge of *Bacteroides* species, host specific markers and molecular methodologies including factors that influence assay efficiency is thus necessary for effective application of *Bacteroides* PCR techniques.

CHAPTER 3 – CONCLUSIONS AND RECOMMENDATIONS

3.1 Water pollution challenges and the need for alternative indicators

Worldwide, water quality and water pollution is of concern, especially with regard to faecal contamination. Assessing the microbial quality of water for potable and recreational use aims to reduce health risks to humans. Monitoring standard faecal indicator bacteria (FIB), though relatively easy and cost effective, have some disadvantages, such as a relatively long incubation periods, lack of specificity, and inaccuracy regarding detection of microorganisms which are slow growing or viable but difficult to culture organisms (non-culturable). Current methods also do not exclude false positives or false negatives due to genetic diversity of FIBs, and do not assist in pinpointing sources of contamination, thus not assisting authorities to address any problems. Alternative techniques making use of alternative indicator organisms are therefore regarded as necessary.

3.2 Advantages of using *Bacteroides* related markers alongside or instead of traditional FIB

Bacteroides related PCR and qPCR methods are advantageous for use as they are relatively sensitive with regards to detection and quantification of host-associated markers. These methods seem most effective when applied in conjunction with other assays in a toolbox approach. Making use of a toolbox approach may aid in exclusion of false positive and false negative data and lessen the risk of over or underestimation of the extent of pollution in water from a specific area of concern. Affordability and robustness of these methods determine whether they may be applied in developing countries. Application in developing countries may contribute to improvement of overall water quality and effective management of water management systems throughout. Mitigation and maintenance cost may be drastically reduced by incorporation of *Bacteroides* assays into municipal management regimes.

3.3 Shortcomings of *Bacteroides* molecular markers and associated assays

Although *Bacteroides* assays are relatively sensitive, various factors may influence the efficiency and success of these applications. Inhibition, lack of standardisation and limited knowledge of technicians on the subject of *Bacteroides* as well as molecular methodologies in general may influence their application. Validation and optimisation of assays are required

for every location or geographic region where application is desired. Developing countries under financial strain may not have the resources available for purchasing of equipment and reagents needed for molecular assays. Due to the lack in standardisation, careful interpretation of data is necessary as different baseline threshold values in comparative studies may lead to over or underestimation of the extent of contamination or associated risk.

3.4 Recommendations on the application of *Bacteroides* related molecular assays for detection and quantification of faecal pollution in environmental water sources in developing countries

There is a definitive need for standardisation of *Bacteroides* assays globally. Recommendations for effective future application include:

- Assessment of a wider range of animal hosts in an attempt to reduce the occurrence of false positive results due to low marker specificity. Further research is necessary into the genetics behind the primers and their binding sites in the host genome in order to increase sensitivity and specificity of markers.
- Understanding has to be developed on how host diet and intestinal physiology, as well as environmental factors such as temperature or differences in geographic regions influence the presence and abundance of *Bacteroides* species. These factors must be taken into account when processing collected data in order to deliver comparable results between different hosts or samples from different environments and regions.
- Any seasonal or similar variations may have to be studied over longer periods of time in order to establish trends on whether contamination in an area of concern in periodical or chronic. If pollution trends indicate chronic contamination further source tracking steps may be performed. Further steps include assessment of the area for potential contamination sources such as storm water outfalls, leaking sewer pipes, cattle feedlots etc. Validation of assays intended for use on samples from the study area. Sectional sampling may take place in order to disregard sources which deliver negative results. False positives from wildlife and horizontal transfer between hosts in the area must then be excluded. Samples taken after these assessments potentially indicate the true source of contamination within the environment.
- Clear definitions of baseline threshold values which are acceptable for use with regards to *Bacteroides* assays should be made available to aid in standardisation of these methods. Care should be taken that laboratories testing for *Bacteroides* markers make

use of the correct equipment and employ staff proficient in molecular methodology who bear knowledge of *Bacteroides* assays and factors which may influence these techniques. Standardisation of assay protocols is imperative for the collection of comparable data, particularly when comparing samples from different environments or geographic regions.

- Application of *Bacteroides* assays in developing countries are dependent on available funding, equipment and other resources. Validation of assays applied on samples from any new environment or geographic region have to take place. Robustness and efficiency of assays which can be applied with minimal resources, must be assessed for effective application in developing countries. Developing fast, low cost assays may be advantageous for use in Africa, India and elsewhere where resources are limited. Alternative methods such as ddPCR, may even be considered in such instances as equipment and reagents for DNA extraction, filtration etc. are not essential.
- Once standardisation of assays has taken place, host specific *Bacteroides* PCR and/or qPCR techniques may be incorporated into governmental and municipal water management programmes such as the Blue and Green drop system used in South Africa. This in turn may aid in more effective mitigation of faecal contamination issues seen in water in distribution systems such as the case in Sannieshof and the environment with reference to the current water quality issues in the Vaal River.

REFERENCES

Ahmed, W., Goonetilleke, A., Powell, D., Chauhan, K. and Gardner, T. (2009)[A] Comparison of molecular markers to detect fresh sewage in environmental waters. *Water Research*, 43(19), pp. 4908–4917.

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

Ahmed, W., Gyawali, P., Sidhu, J. P. S. and Toze, S. (2014) Relative inactivation of faecal indicator bacteria and sewage markers in freshwater and seawater microcosms. *Letters in Applied Microbiology*, 59(3), pp. 348–354.

Ahmed, W., Masters, N. and Toze, S. (2012) Consistency in the host specificity and host sensitivity of the *Bacteroides* HF183 marker for sewage pollution tracking. *Letters in Applied Microbiology*, 55(4), pp. 283–289.

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

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

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

Anonymous.1 (2013) Save the Vaal Environment. www.save.org.za (date of access 31/10/2018)

Anonymous 2. (2018) Human Rights Commission to hold inquiry into toxic Vaal River. Timeslive. 4 September 2018. www.timeslive.co.za/news/south-africa/2018-09-04-human-rights-commision-to-hold-inquiry-into-toxic-vaal-river/ (date of access 20/11/2018)

Anonymous 3. (2017) Municipality stricter on illegal water and electricity connections- Make use of amnesty period. DA. 1 September 2017. https://www.da.org.za/government/undefined/ 2017/09/municipality-stricter-illegal-water-electricity-connections-make-use-amnesty-period (date of access 30/10/2018).

Anonymous 4. (2018) Resident blows the whistle on illegal water connections. Northcliff Melville Times. 5 February 2018. https://northcliffmelvilletimes.co.za/267158/resident-blows-whistle-illegal-water-connections/ (date of access 30/10/2018).

Aslan, A., Kinzelman, J., Dreelin, E., Anan'eve, T., Lavender, J., Aw, T. G. and Wisnieski, L. C. (2015) A guidance document for testing recreational waters using USEPA qPCR method C, Update for method C. www.cws.msu.edu/documents/qPCRmanual_final.pdf (date of access 31/10/2018)

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

Ballesté, E. and Blanch, A. R. (2010) Persistence of *Bacteroides* species populations in a river as measured by molecular and culture techniques. *Applied and Environmental Microbiology*, 76(22), pp. 7608–7616.

Ballesté, E., Bonjoch, X., Belanche, L. A. and Blanch, A. R. (2010) Molecular indicators used in the development of predictive models for microbial source tracking. *Applied and Environmental Microbiology*, 76(6), pp. 1789–1795.

Bell, A., Layton, A. C., Mckay, L., Williams, D., Gentry, R. andSayler, G.S. (2009) Factors influencing the persistence of fecal *Bacteroides* in stream water. *Journal of Environmental Quality, 38, pp. 1224-1232.*

Bernard, A.E. and Field, K. G. (2000) A PCR assay to discriminate human and ruminant feces on the basis of host differences in *Bacteroides-Prevotella* genes encoding 16S rRNA. *Water Research*, 66(10), pp. 565–567.

Boehm, A. B., Soller, J. A. and Shanks, O. C. (2015) Human-associated fecal quantitative polymerase chain reaction measurements and simulated risk of gastrointestinal illness in recreational waters contaminated with raw sewage. *Environmental Science and Technology Letters*, 2(10), pp. 270–275.

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

Bonkosky, M., Hernández-Delgado, E. A., Sandoz, B., Robledo, I. E., Norat-Ramírez, J. and Mattei, H. (2009) Detection of spatial fluctuations of non-point source fecal pollution in coral reef surrounding waters in southwestern Puerto Rico using PCR-based assays. *Marine Pollution Bulletin*, 58(1), pp. 45–54.

Bornman, J. (2018) 2 Waterfall City developers arrested for alleged illegal water connections. News24. 30 May 2018. https://www.news24.com/SouthAfrica/News/2-waterfall-city-developers-arrested-for-alleged-illegal-water-connections-20180530 (date of access 30/10/2018).

Bower, P. A., Scopel, C. O., Jensen, E. T., Depas, M. M. and Mclellan, S. L. (2005) Detection of genetic markers of fecal indicator bacteria in Lake Michigan and determination of their relationship to *Escherichia coli* densities using standard microbiological methods. *Applied and Environmental Microbiology*, 71(12), pp. 8305–8313.

Burges, J. (No Date) South African green drop certification for excellence in wastewater treatment plant operation. International Water Association, www.iwa-network.org/filemanager-uploads/WQ_Compendium/Cases/South%20African%20Green.pdf. (date of access 31/10/2018)

Cao, Y., Raith, M.R. And Griffith, J.F. (2015) Droplet digital PCR for simultaneous quantification of general and human-associated fecal indicators for water quality assessment. *Water Research*, 70, pp. 337-349.

Chase, E., Hunting, J., Staley, C. and Harwood, V. J. (2012) Microbial source tracking to identify human and ruminant sources of faecal pollution in an ephemeral Florida river. *Journal of Applied Microbiology*, 113(6), pp. 1396–1406.

Conroy, R.M., Meegan, N.E., Joyce, T., McGuigan, K., and Barnes, J. (2001) Solar disinfection of drinking water protects against Cholera in children under six years of age. *Archives of Disease in Childhood*, 85, pp. 293–295.

Dick, L. K., Bernhard, A. E., Brodeur, T. J., Santo Domingo, J. W., Simpson, J. M., Walters, S. P. and Field, K. G. (2005) Host distributions of uncultivated fecal Bacteroidales bacteria reveal genetic markers for fecal source identification. *Applied and Environmental Microbiology*, 71(6), pp. 3184–3191.

Dick, L. and Field, K. (2004) Rapid estimation of numbers of fecal Bacteroidetes by use of a quantitative PCR assay for 16S rRNA genes. *Applied and Environmental Microbiology*, 70(9), pp. 5695–5697.

D'Inverno, G., CArosi, L., Roman, G and Guerrini, A. (2018) Water pollution in wastewater treatment plants: An efficiency analysis with undesirable output, *European Journal of Operational Research*, 269, pp. 24-34.

Dorai-Raj, S., Grady, J. O. and Colleran, E. (2009) Specificity and sensitivity evaluation of novel and existing Bacteroidales and Bifidobacteria-specific PCR assays on feces and sewage samples and their application for microbial source tracking in Ireland. *Water Research*, 43(19), pp. 4980–4988.

Ebdon, J., Muniesa, M. and Taylor, H. 2007) The application of a recently isolated strain of *Bacteroides* (GB-124) to identify human sources of faecal pollution in a temperate river catchment. *Water Research*, 41(16), pp. 3683–3690.

Ebentier, D. L., Hanley, K. T., Cao, Y., Badgley, B. D., Boehm, A. B., Ervin, J. S., Goodwin, K. D., Gourmelon, M., Griffith, J. F., Holden, P. A., Kelty, C. A., Lozach, S., McGee, C., Peed, L. A., Raith, M., Ryu, H., Sadowsky, M. J., Scott, E. A., Domingo, J. S., Schriewer, A., Sinigalliano, C. D., Shanks, O. C., Van De Werfhorst, L. C., Wang, D., Wuertz, S. and Jay, J. A. (2013) Evaluation of the repeatability and reproducibility of a suite of qPCR-based microbial source tracking methods. *Water Research*, 47(18), pp. 6839–6848.

Edokpayi, J. N., Rogawski, E. T., Kahler, D. M., Hill, C .L., Reynolds, C., Nyathi, E., Smith, J. A., Odiyo, J. O., Samie, A.,Bessong, P., and Dillingham, R. (2018) Challenges to sustainable safe drinking water: A case study of water quality and use across seasons in rural communities in Limpopo Province, South Africa. *Water*, 10(159).

Ervin, J. S., Russell, T. L., Layton, B. A., Yamahara, K. M., Wang, D., Sassoubre, L. M., Cao, Y., Kelty, C. A., Sivaganesan, M., Boehm, A. B., Holden, P. A., Weisberg, S. B. and Shanks, O. C. (2013) Characterization of fecal concentrations in human and other animal sources by physical, culture-based, and quantitative real-time PCR methods. *Water Research*, 47(18), pp. 6873–6882.

Figueredo-Fernández, M., Gutiérrez-Alfaro, S., Acevedo-Merino, A and Manzano, M. A. (2017) Estimating lethal dose of solar radiation for enterococcus inactivation through radiation reaching the water layer. Application to solar water disinfection (SODIS). *Solar Energy,* 158, pp. 303-310.

Gawler, A. H., Beecher, J. E., Brandão, J., Carroll, N. M., Falcão, L., Gourmelon, M., Masterson, B., Nunes, B., Porter, J., Rincé, A., Rodrigues, R., Thorp, M., Martin Walters, J. and Meijer, W. G. (2007) Validation of host-specific Bacteriodales 16S rRNA genes as markers to determine the origin of faecal pollution in Atlantic Rim countries of the European Union, *Water Research*, 41(16), pp. 3780–3784.

GDR. (2011) Green drop report. Chapter 9: North West. Department of Water Affairs, South Africa, pp 293-323. http://www.dwa.gov.za/Documents/GD/GDIntro.pdf (date of access 30/10/2018)

Chalchisa, D., Megersa, M. and Beyene, A. (2017) Assessment of the quality of drinking water in storage tanks and its implication on the safety of urban water supply in developing countries. *Envrionemtnal Systems Research*, 6(12).

Getachew, A., Tadie, A., Chercos, D.H and Guadu, T. (2018) Level of Faecal Coliform Contamination of Drinking Water Sources and Its Associated Risk Factors in Rural Settings of North Gondar Zone, Ethiopia: A Cross-Sectional Community Based Study. *Ethiopian Journal of Science*, 28(2), pp. 227-234.

González-Gómez, F., García-Rubio, M. A., Guardiola, J. (2011) Why is non-revenue water so high in so m any cities? *Water Resources Development*, 27(2), pp. 345-360.

Gordon, K. V., Brownell, M., Wang, S. Y., Lepo, J. E., Mott, J., Nathaniel, R., Kilgen, M., Hellein, K. N., Kennedy, E. and Harwood, V. J. (2013) Relationship of human-associated microbial source tracking markers with Enterococci in Gulf of Mexico waters. *Water Research*. 47(3), pp. 996–1004.

Gorham, T. J., Yoo, J., Garabed, R., Mouhaman, A. and Lee, J. (2017) Water access, sanitation, and hygiene conditions and health outcomes among two settlement types in rural far North Cameroon. *International Journal of Environmental Research and Public Health*, 14(4), pp. 441.

Gourmelon, M., Caprais, M. P., Ségura, R., Le Mennec, C., Lozach, S., Piriou, J. Y., Rincé, A., Ahmed, W., Stewart, J., Powell, D., Gardner, T., Gourmelon, M., Caprais, M. P., Ségura, R., Le Mennec, C., Lozach, S., Piriou, J. Y., Rincé, A., Savichtcheva, O., Okayama, N. and Okabe, S. (2007) Evaluation of two library-independent microbial source tracking methods to identify sources of fecal contamination in French estuaries. *Applied and Environmental Microbiology*, 73(15), pp. 4857–4866.

Gouws, C., Moeketsi, I., Motloung, S., Tempelhoff, J., van Greuning, G. and van Zyl, L. (2010) SIBU and the crisis of water service delivery in Sannieshof, North West Province, *The Journal for Transdisciplinary Research in Southern Africa*, 6(1), pp. 25–56.

Green, H. C. and Field, K. G. (2012) Sensitive detection of sample interference in environmental qPCR. *Water Research*, 46(10), pp. 3251–3260.

Green, H. C., Haugland, R. A., Varma, M., Millen, H. T., Borchardt, M. A., Field, K. G., Walters, W. A., Knight, R., Sivaganesan, M., Kelty, C. A. and Shanks, O. C. (2014) Improved HF183 quantitative real-time PCR assay for characterization of human fecal pollution in ambient surface water samples. *Applied and Environmental Microbiology*, 80(10), pp. 3086–3094.

Green, H. C., Shanks, O. C., Sivaganesan, M., Haugland, R. A. and Field, K. G. (2011) Differential decay of human faecal *Bacteroides* in marine and freshwater. *Environmental Microbiology*, 13(12), pp. 3235–3249.

Gundry, S., Wright, J., and Conroy, R. (2004) A systematic review of the health outcomes related to household water quality in developing countries. *Journal of Water and Health*, 2, pp. 1–14.

Haack, S. K., Fogarty, L. R., Stelzer, E. A., Fuller, L. M., Brennan, A. K., Isaacs, N. M. and Johnson, H. E. (2013) Geographic setting influences Great Lakes beach microbiological water quality. *Environmental Science and Technology*, 47(21), pp. 12054–12063.

Hall, V., Collins, M. D., Hutson, R. A., Lawson, P. A., Falsen, E. and Duerden, B. I. (2003) *Corynebacterium atypicum* sp. nov., from a human clinical source, does not contain corynomycolic acids. *International Journal of Systematic and Evolutionary Microbiology*, 53(4), pp. 1065–1068.

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

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

Heaney, C. D., Exum, N. G., Dufour, A. P., Brenner, K. P., Haugland, R. A., Chern, E., Schwab, K. J., Love, D. C., Serre, M. L., Noble, R. and Wade, T. J. (2014) Water quality, weather and environmental factors associated with fecal indicator organism density in beach sand at two recreational marine beaches. *Science of the Total Environment*. 497–498, pp. 440–447.

Hernández-Rodríguez, P. and Ramirez, A. G. (2012) Polymerase chain reaction : types , utilities and limitations. *INTECH Open Access Publisher*, pp. 157–172.

Hill, J.E., Seipp, R.P., Betts, M., Hawkins, L., Van Kessel, A.G., Crosby, W.L., and Hemmingson, S.M. (2002) Extensive profiling of a complex microbial community by high-throughput sequencing. *Applied and Environmental Microbiology*, 68(6), pp. 30553066.

Hudson, W.D. (1964) Reduction of unaccounted-for water. *Journal of the American Water Works Association*, 56(2), pp. 143–148.

Human, J. (2017). Councillor allegedly connects illegal pipes directly to reservoir amidst #watercrisis. *OFM News*. 2 December 2017. https://www.ofm.co.za/article/localnews/254466/photo-gallery-councillor-allegedly-connects-illegal-pipes-directly-to-reservoir-amidstwatercrisis (date of access 30/10/2018).

Hurtgen, P.J. (1931) Unaccounted for water. *Journal of the American Water Works Association*, 24(4), pp. 505–514.

Jadhav, A. S., Chembe, D. K., Strauss, J. M. and van Niekerk, J. L. (2017) Status of solar technology implementation in the Southern African developing community (SADC) region. *Renewable and Sustainable Energy Reviews,* 73, pp. 622-631.

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

Jensen, P.K., Ensink, J.H.J., Jayasinghe, G., Van der Hoek, W., Cairncross, S., and Dalsgaard, A. (2002) Domestic transmission routes of pathogens: the problem of in-house contamination of drinking water during storage in developing countries. *Tropical Medicine and International Health*, 7(7), pp. 604–609.

Jofre, J., Blanch, A. R., Lucena, F. and Muniesa, M. (2014) Bacteriophages infecting *Bacteroides* as a marker for microbial source tracking. *Water Research*, 55, pp. 1–11.

Jang, D. and Choi, G. (2017) Estimation of non-revenue water ratio using MRA and ANN in water distribution networks. *Water*, 10(2).

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

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

Kirs, M., Harwood, V. J., Fidler, A. E., Gillespie, P. A., Fyfe, W. R., Blackwood, A. D. and Cornelisen, C. D. (2011) Source tracking faecal contamination in an urbanised and a rural waterway in the Nelson-Tasman region, New Zealand. *New Zealand Journal of Marine and Freshwater Research*, 45(1), pp. 43–58.

Kreader, C. A. (1998) Persistence of PCR-Detectable *Bacteroides distasonis* from Human Feces in River Water. *Applied and Environmental Microbiology,* 64(10).

Kumar, S. (2010) Unaccounted for water and the performance of water utilities: an empirical analysis from India. *Water Policy*, 12, pp. 707-721

Kurokawa, K., Takehiko, I., Tomomi, K., Kenshiro, O., Hidehiro, T., Atsushi, T., Hideto, T., Hidetoshi, M., Vineet, K. S., Tulika, P. S., Todd, D. T., Hideki, N., Hiroshi, M., Yoshitoshi, O., Dusko, S. E.,Kikuji, I., Toshihisa, T., Yoshiyuki, S., Tetsuya, H. and Masahira, H. (2007) Comparative metagenomics revealed commonly enriched gene sets in human gut microbiomes. *DNA Research*, 14, pp. 168161.

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

Layton, A., McKay, L., Williams, D., Garrett, V., Gentry, R. and 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), pp. 4214–4224.

Lee, C. S., Marion, J. W. and Lee, J. (2011) A novel genetic marker for the rapid detection of *Bacteroides fragilis* in recreational water as a human-specific faecal indicator. *Journal of Water and Health*, 9(2), pp. 253–264.

Lee, Y. J., Molina, M., Santo Domingo, J. W., Willis, J. D., Cyterski, M., Endale, D. M. and Shanks, O. C. (2008) Temporal assessment of the impact of exposure to cow feces in two watersheds by multiple host-specific PCR assays. *Applied and Environmental Microbiology*, 74(22), pp. 6839–6847.

Lee, E. J. and Schwab, K. J. (2005) Deficiencies in drinking water distribution systems in developing countries. *Journal of Water and Health*, 3(2), pp. 109–127.

Lee, D. Y., Weir, S. C., Lee, H. and Trevors, J. T. (2010) Quantitative identification of fecal water pollution sources by TaqMan real-time PCR assays using *Bacteroidales* 16S rRNA genetic markers. *Applied Microbiology and Biotechnology*, 88(6), pp. 1373–1383.

Liang, Z., He, Z., Zhou, X., Powell, C. A., Yang, Y., He, L. M. and Stoffella, P. J. (2013) Impact of mixed land-use practices on the microbial water quality in a subtropical coastal watershed. *Science of the Total Environment*, 449, pp. 426–433.

Mackintosh, G. and Colvin, C. (2003) Failure of rural schemes in South Africa to provide potable water. *Environmental Geology*, 44(1), pp. 101–105.

Mac Mahon, J. and Gill, L. W. (2018) Sustainability of novel water tteatment technologies in developing countries: Lessons learned from research trials on a pilot continuous flow solar water disinfection system in rural Kenya. *Development Engineering,* 3, pp.47-59.

Mayer, R. E., Bofill-Mas, S., Egle, L., Reischer, G. H., Schade, M., Fernandez-Cassi, X., Fuchs, W., Mach, R. L., Lindner, G., Kirschner, A., Gaisbauer, M., Piringer, H., Blaschke, A. P., Girones, R., Zessner, M., Sommer, R. and Farnleitner, A. H. (2016) Occurrence of human-associated *Bacteroidetes* genetic source tracking markers in raw and treated wastewater of municipal and domestic origin and comparison to standard and alternative indicators of faecal pollution. *Water Research*, 90, pp. 265–276.

Mckenzie, R., Siqalaba, Z. N. and Wegelin, W. A. (2012) The state of non-revenue water in south africa. WRC Report No. TT 522/12. *Water Research Commission*, Pretoria.

McLain, J. E. T., Ryu, H., Kabiri-Badr, L., Rock, C. M. and Abbaszadegan, M. (2009) Lack of specificity for PCR assays targeting human *Bacteroides* 16S rRNA gene: Cross-amplification with fish feces. *FEMS Microbiology Letters*, 299(1), pp. 38–43.

McLellan, S. L. and Eren, A. M. (2014) Discovering new indicators of fecal pollution. *Trends in Microbiology*, 22(12), pp. 697–706.

McQuaig, S., Griffith, J. and 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), pp. 6423–6432.

Mika, K. B., Ginsburg, D. W., Lee, C. M., Thulsiraj, V. and Jay, J. A. (2014) Fecal indicator bacteria levels do not correspond with incidence of human-associated HF183 *Bacteroides* 16S rRNA genetic marker in two urban southern California watersheds. *Water, Air, and Soil Pollution*, 225, 1960.

Mochoari, R. (2012). Varke kry onwettig water [web log post]. 7 September 2012. https://www.facebook.com/search/str/Vrystaat+Nuus%2FNews+Refilwe+Mochoari/keywords search. (date of access 30/10/2018).

Nakamura, J., Kubota, Y., Miyaoka, M., Saitoh, T., Mizuno, F. and Benno, Y. (2002) Comparison of four microbial enzymes in *Clostridia* and *Bacteroides* isolated from human feces. *Microbiology and Immunology*, 46(7), pp. 487–490.

Napier, M. D., Haugland, R., Poole, C., Dufour, A. P., Stewart, J. R., Weber, D. J., Varma, M., Lavender, J. S. and Wade, T. J. (2017) Exposure to human-associated fecal indicators and self-reported illness among swimmers at recreational beaches: A cohort study. *Environmental Health: A Global Access Science Source*, 16(1), pp. 1–16.

Newton, R. J., VandeWalle, J. L., Borchardt, M. A., Gorelick, M. H. and 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), pp. 6972–6981.

Nnane, D. E., Ebdon, J. E. and Taylor, H. D. (2011) Integrated analysis of water quality parameters for cost-effective faecal pollution management in river catchments, *Water Research*. 45(6), pp. 2235–2246.

Nshimyimana, J. P., Cruz, M. C., Thompson, R. J. and Wuertz, S. (2017) Bacteroidales markers for microbial source tracking in Southeast Asia. *Water Research*. 118, pp. 239–248.

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

NWRS. (2013) National Water Resource Strategy, Chapter 2. South Africa's water situation and strategies to balance supply and demand. Department of Water Affairs and Forestry, South Africa. pp. 17-54. http://www.dwa.gov.za/nwrs/NWRS2013. aspx (date of access 30/10/2018)

Obiri-Danso, K., Okore-Hanson, A. and Jones, K. (2003) The microbiological quality of drinking water sold on the streets in Kumasi, Ghana. *Letters in Applied Microbiology*, 37(4), pp. 334–339.

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

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

Onyango, A. E., Okoth, M. W., Kunyanga, C. N. and Aliwa, B. O. (2018) Microbiological quality and contamination level of water sources in Isiolo County in Kenya. *Journal of Envrionmental and Public Health*.

Oshiro, R.K. (2002) Method 1604: Total coliforms and *Escherichia coli* in water by membrane filtration using a simultaneous detection technique (MI Medium). US-EPA. https://nepis.epa.gov/Exe/ZyPDF.cgi/P1002D57.PDF?Dockey=P1002D57.PDF (date of access 31/10/2018)

Pawson, M., and Boshoff, C. (2017) Afriforum Branch project report for the Blue- and Green drop project for 2017. www.afriforum.co.za/wp-content/uploads/Druppel-Verslag-eng/1.pdf (date of access 30/10/2018)

Phakgadi, P. (2018) Vaal water crisis: “Why are investigations only beginning now?” News24. 19 November 2018. https://m.news24.com/SouthAfrica/News/vaal-water-crisis-why-are-investigations-only-beginning-now-20181119 (date of access 20/11/2018)

Pijoos, I. (2018) Raw sewage to keep flowing into Vaal as intervention will take a year. BusinessDay, 12 November 2018. www.businesslive.co.za/bd/national/2018-11-12-raw-sewage-to-keep-flowing-into-vaal-as-intervention-will-take-a-year/ (date of access 20/11/2018)

Pollet, R. M., D’Agostino, E. H., Walton, W. G., Xu, Y., Little, M. S., Biernat, K. A., Pellock, S. J., Patterson, L. M., Creekmore, B. C., Isenberg, H. N., Bahethi, R. R., Bhatt, A. P., Liu, J., Gharaibeh, R. Z. and Redinbo, M. R. (2017) An atlas of β-Glucuronidases in the human intestinal microbiome. *Structure*, 25(7), p. 967–977.

Potgieter, S., Pinto, A., Sigudu, M., du Preez, H., Ncube, E. and Venter, S. (2018) Long-term spatial and temporal microbial community dynamics in a large-scale drinking water distribution system with multiple disinfectant regimes. *Water Research*, 139, pp. 406-419

Quinto, G. (1962) Nutrition of five *Bacteroides* strains. *Journal of Bacteriology*, 84(3), pp. 559-562.

Raith, M. R., Kelty, C. A., Griffith, J. F., Schriewer, A., Wuertz, S., Mieszkin, S., Gourmelon, M., Reischer, G. H., Farnleitner, A. H., Ervin, J. S., Holden, P. A., Ebentier, D. L., Jay, J. A., Wang, D., Boehm, A. B., Aw, T. G., Rose, J. B., Balleste, E., Meijer, W. G., Sivaganesan, M. and Shanks, O. C. (2013) Comparison of PCR and quantitative real-time PCR methods for the characterization of ruminant and cattle fecal pollution sources. *Water Research*, 47(18), pp. 6921–6928.

Ridley, C. M., Jamieson, R. C., Truelstrup Hansen, L., Yost, C. K. and Bezanson, G. S. (2014) Baseline and storm event monitoring of Bacteroidales marker concentrations and enteric pathogen presence in a rural Canadian watershed. *Water Research*, 60, pp. 278–288.

Riedel, T. E., Thulsiraj, V., Zimmer-Faust, A. G., Dagit, R., Krug, J., Hanley, K. T., Adamek, K., Ebentier, D. L., Torres, R., Cobian, U., Peterson, S. and Jay, J. A. (2015) Long-term monitoring of molecular markers can distinguish different seasonal patterns of fecal indicating bacteria sources. *Water Research*, 71, pp. 227–243.

Riedel, T. E., Zimmer-Faust, A. G., Thulsiraj, V., Madi, T., Hanley, K. T., Ebentier, D. L., Byappanahalli, M., Layton, B., Raith, M., Boehm, A. B., Griffith, J. F., Holden, P. A., Shanks, O. C., Weisberg, S. B. and Jay, J. A. (2014) Detection limits and cost comparisons of human- and gull-associated conventional and quantitative PCR assays in artificial and environmental waters. *Journal of Environmental Management*, 136, pp. 112–120.

Roeger, A. and Tavares, A.F. (2018) Water safety plans by utilities: A review of research on implementation. *Utilities Policy*, 53, pp. 15-24.

Rojko, C. (2003) Solar disinfection of drinking water. (M.Sc. Thesis, Worchester Polytechnic Institute). www.dl.mozh.org/up/rojko.pdf (date of access 30/10/2018)

Rompré, A., Servais, P., Baudart, J., De-Roubin, M.-R. and Patrick Laurent, P. (2002) Detection and enumeration of coliforms in drinking water: current methods and emergin approaches. *Journal of Microbiological Methods*, 49(2002), pp. 31–54.

Rouault, M. and Richard, Y. (2003) Intensity and spatial extension of drought in South Africa at different time scales. *Water SA*, 29(4), pp. 489–500.

Rusca, M., Boakye-Anash, A. S., Loftus, A., Ferrero, G. and van der Zaag, P. (2017) An interdisciplinary poliutical ecology of drinking water quiality. Exploring sociao-ecological inequalities in Lilongwe's water supply network. *Geoforum*, 84, pp. 138-146.

Samir, N., Kansoh, R., Elbarki, W. and Fleifle, A. (2017) Pressure control for minimizing leakage in water ditribution systems. *Alexandria Engineering Journal,* 56, pp. 601-612.

SANS 241 (2015) *South African National Standard: Drinking water.* SABS, Standards Division. 240-1: 2015.

Santoro, A. E. and Boehm, A. B. (2007) Frequent occurrence of the human-specific *Bacteroides* fecal marker at an open coast marine beach: Relationship to waves, tides and traditional indicators. *Environmental Microbiology*, 9(8), pp. 2038–2049.

Sauer, E. P., VandeWalle, J. L., Bootsma, M. J. and 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), pp. 4081–4091.

Savichtcheva, O. and Okabe, S. (2006) Alternative indicators of fecal pollution: Relations with pathogens and conventional indicators, current methodologies for direct pathogen monitoring and future application perspectives. *Water Research*, 40(13), pp. 2463–2476.

Schlosser, C.A., Strzepek, K., Gao, X., Fant, C., Blanc, E., Paltsev, S., Jacoby, H., Reilly, J. and Gueneau, A. (2014) The future of global water stress: An integrated assessment. *AGU Earth's Future*, 2, pp. 341–361,

Schriewer, A., Miller, W. A., Byrne, B. A., Miller, M. A., Oates, S., Conrad, P. A., Hardin, D., Yang, H. H., Chouicha, N., Melli, A., Jessup, D., Dominik, C. and Wuertz, S. (2010) Presence of Bacteroidales as a predictor of pathogens in surface waters of the central California coast. *Applied and Environmental Microbiology*, 76(17), pp. 5802–5814.

Seidel, L., Strathmann, M. and Nocker, A. (2017) The feasibility of improved live-dead distinction in qPCR-based microbial source tracking. *Journal of Microbiological Methods*, 140, pp. 23–31.

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

Shah, V. G., Dunstan, R. H., Geary, P. M., Coombes, P., Roberts, T. K. and Rothkirch, T. (2007) Comparisons of water quality parameters from diverse catchments during dry periods and following rain events. *Water Research*, 41(16), pp. 3655–3666.

Shahryari, A., Nikaeen, M., Khiadani (Hajian), M., Nabavi, F., Hatamzadeh, M. and Hassanzadeh, A. (2014) Applicability of universal Bacteroidales genetic marker for microbial monitoring of drinking water sources in comparison to conventional indicators. *Environmental Monitoring and Assessment*, 186(11), pp. 7055–7062.

Shanks, O. C., White, K., Kelty, C. a, Sivaganesan, M., Blannon, J., Meckes, M. and Varma, M. (2010) Performance of PCR based assays targeting Bacteriodales genetic markers of human fecal pollution in sewage and fecal samples. *Environmental Science and Technology*, 44(16), pp. 6281–6288.

Sidhu, J. P. S., Ahmed, W., Gernjak, W., Aryal, R., McCarthy, D., Palmer, A., Kolotelo, P. and Toze, S. (2013) Sewage pollution in urban stormwater runoff as evident from the widespread presence of multiple microbial and chemical source tracking markers. *Science of the Total Environment*, 463–464, pp. 488–496.

Sidhu, J. P. S., Hodgers, L., Ahmed, W., Chong, M. N. and Toze, S. (2012) Prevalence of human pathogens and indicators in stormwater runoff in Brisbane, Australia. *Water Research*, 46(20), pp. 6652–6660.

Sinton, L. W., Finlay, R. K. and Hannah, D. J. (1998) Distinguishing human from animal faecal contamination in water: A review. *New Zealand Journal of Marine and Freshwater Research*, 32(2), pp. 323–348.

Sowah, R. A., Habteselassie, M. Y., Radcliffe, D. E., Bauske, E. and Risse, M. (2017) Isolating the impact of septic systems on fecal pollution in streams of suburban watersheds in Georgia, United States. *Water Research*, 108, pp. 330–338.

Staley, C., Gordon, K. V., Schoen, M. E. and Harwood, V. J. (2012)[A] Performance of two quantitative PCR methods for microbial source tracking of human sewage and implications for microbial risk assessment in recreational waters. *Applied and Environmental Microbiology*, 78(20), pp. 7317–7326.

Staley, C., Reckhow, K. H., Lukasik, J. and Harwood, V. J. (2012)[B] Assessment of sources of human pathogens and fecal contamination in a Florida freshwater lake. *Water Research*, 46(17), pp. 5799–5812.

Stea, E. C., Truelstrup Hansen, L., Jamieson, R. C. and Yost, C. K. (2015) Fecal contamination in the surface waters of a rural- and an urban-source watershed. *Journal of Environment Quality*, 44(5), p. 1556.

Stewart, J.R., Boehm, A.B., Dubinsky, E.A., Fong, T., Goodwin, K.D., Griffith, J.F., Noble, R.T., Shanks, O.C., Vijayavel, K., and Weisberg, S.B. (2013) Recommendations following a multi-laboratory comparison of microbial source tracking methods. *Water Research*, 47, pp. 6829–6838.

Stringer, A. M., Gibson, R. J., Logan, R. M., Bowen, J. M., Yeoh, A. S. J. and Keefe, D. M. K. (2008) Faecal microflora and β-glucuronidase expression are altered in an irinotecan-induced diarrhoea model in rats. *Cancer Biology and Therapy*, 7(12), pp. 1919–1925.

Toledo-Hernandez, C., Ryu, H., Gonzalez-Nieves, J., Huertas, E., Toranzos, G. A. and Santo Domingo, J. W. (2013) Tracking the primary sources of fecal pollution in a tropical watershed in a one-year study. *Applied and Environmental Microbiology*, 79(5), pp. 1689–1696.

Tumwine, J. K., Thompson, J., Katua-Katua, M., Mujwajuzi, M., Johnstone, N. and Porras, I. (2002) Diarrhoea and effects of different water sources, sanitation and hygiene behaviour in East Africa. *Tropical Medicine & International Health*, 7(9), pp. 750–756.

Van De Werfhorst, L. C., Sercu, B. and Holden, P. A. (2011) Comparison of the host specificities of two Bacteroidales quantitative PCR assays used for tracking human fecal contamination. *Applied and Environmental Microbiology*, 77(17), pp. 6258–6260.

Venter. A (2000) Quality of domestic water supplies volume 2 : sampling guide. *Water Research Commission*, 2 (Report No. TT 117), p. 99.

Wada, Y., Bierkens, M.F.P., de Roo, A., Dirmeyer, P.A., Famiglietti, J.S., Hanasaki, N., Konar, M., Liu, J., Shmied, H. M., Oki, T., Pokhrel, Y., Sivapalan, M., Troy, T. J., van Dijk, A. I. J M., van Emmerik, T., Van Huigevoort, M. H. J., Van Lanen, H. A. J., Vörösmarty, C. J., Wanders, N. and Wheater, H. (2017) Human–water interface in hydrological modelling: current status and future directions *Hydrology and Earth System Sciences*, 21, pp. 4169–4193

Wade, T. J., Sams, E., Brenner, K. P., Haugland, R., Chern, E., Beach, M., Wymer, L., Rankin, C. C., Love, D., Li, Q., Noble, R. and Dufour, A. P. (2010) Rapidly measured indicators of recreational water quality and swimming-associated illness at marine beaches: A prospective cohort study. *Environmental Health: A Global Access Science Source*, 9(1), pp. 1–15.

Walters, S. P. and Field, K. G. (2009) Survival and persistence of human and ruminant-specific faecal Bacteroidales in freshwater microcosms. *Environmental Microbiology*, 11(6), pp. 1410–1421.

Wexler, H.M. (2007) *Bacteroides*: the good, the bad, and the nitty-gritty. *Clinical Microbiology Reviews,* 20(4), pp. 593–621.

Wang, X., van Dam, K. H., Triantafyllidis, C., Koppelaar, R. H. E. M., andShah, N. (2017) Water and energy systems in sustainable city development: A case of Sub-Saharan Africa. *Procedia Engineering,* 198, pp. 948-957.

Wiegner, T. N., Edens, C. J., Abaya, L. M., Carlson, K. M., Lyon-Colbert, A. and Molloy, S. L. (2017) Spatial and temporal microbial pollution patterns in a tropical estuary during high and low river flow conditions. *Marine Pollution Bulletin,* 114(2), pp. 952–961.

Zheng, Y., Kelly, W. R., Panno, S. V and Liu, W. (2013) Identification of sources of fecal pollution of karst waters. *Contract Report 2013-02, Illinois State Water Survey*, pp. 1-18.