Detection and molecular typing of Cryptosporidium in South African wastewater plants

Anton de Jong

Master Degree Project in Infection Biology, 30 credits. Spring 2017
Department: Medical Biochemistry and Microbiology
Supervisor: Dr. Karin Troell, National Veterinary Institute, Sweden
Co-Supervisor: Prof. Thor-Axel Stenström, Durban University of Technology, South Africa
Abstract

Cryptosporidium is a protozoan parasite infecting the intestines of its hosts, leading to acute diarrheal disease. Out of 26 recognized species, 14 are known to infect humans. Of most importance, from a human perspective are Cryptosporidium parvum and Cryptosporidium hominis, of which the former is known to have zoonotic potential. Globally, cryptosporidiosis affect people with lowered immune status particularly hard; among children under five it is the most important parasitic cause of gastroenteritis. In the region of KwaZulu-Natal, on the east coast of South Africa, Cryptosporidium is considered endemic. Drinking water is frequently collected from river systems and as Cryptosporidium spp. can be transmitted via contaminated water, this may be one source of infection. Research on the species distribution is important for outbreak investigations and prevention efforts. In water and wastewater such speciation is commonly performed using immunomagnetic separation, an antibody dependent method. There is however a suspicion that these antibodies have less affinity to some species and hence contorts the detected species distribution. An alternative approach is therefore of interest.

In the present study, Cryptosporidium diversity in wastewater collected from four different wastewater treatment plants in KwaZulu-Natal, is evaluated with an optimized antibody-free workflow and a single cell platform. It was shown that the workflow is suitable for complex samples, such as wastewater. Furthermore, diversity was assessed with amplicon sequencing, revealing four different species and genotypes. Further modifications of the methods used could benefit the field of Cryptosporidium research, along with improving global health and preventing disease outbreaks.
Popular scientific summary

A small parasite of global significance

The single-cell parasite *Cryptosporidium* induces acute diarrheal disease referred to as cryptosporidiosis. The 26 species that are known infect mammals, fish, reptiles and birds and more than half of them have potential to cause disease in humans. The parasite is transmitted through small (5 micrometer) spore-like oocysts that can survive long periods of time in the environment and it is enough with just a few to contract disease. They can also resist disinfectants like chlorine, something that explain why outbreaks connected to swimming pools and drinking water are not uncommon. In most people the symptoms only last up to two weeks, but in those with impaired immune defense, such as an HIV positive individual or young child, the disease can be more severe and sometimes life threatening. *Cryptosporidium* can then also spread to other parts of the body apart from the small intestine where it is usually located. In children under five years of age, it has recently been listed as one of the most important parasitic causes of diarrhea.

Studying *Cryptosporidium* in South African wastewater

In this study, wastewater samples were collected in Durban, South Africa, with the intent of identifying the present species. In this region, cryptosporidiosis is common and a large part of the population are HIV positive (40% of pregnant women in 2013). One popular approach to detect parasites in a water sample is to use a method based on antibodies. The antibodies are attached to a magnetic bead that is used to "fish" out the oocysts. The problem however is that previous observations indicate that this method might not be equally good for all species. Therefore an antibody free method has been used in this study.

Methods used

Samples from four different wastewater plants were collected and processed in South Africa and then transported back to Sweden for analysis. In a study published last year it was shown that it is possible to separate individual oocysts with a cell sorter and examine their DNA. In a complex sample like wastewater, where a great number of different Cryptosporidium variants are expected, such a technique would give a much more detailed picture of what the sample contains.

Another way is to use so called deep sequencing. By utilizing a “biological copy-machine”, called PCR, it is possible to select and amplify a certain region of the parasite’s DNA. This material can later be analysed in detail and different DNA variants can be detected in the same analysis. Both methods were used in the study.

Results and discussion

*Cryptosporidium* oocysts were found in all examined samples. The number of parasites were relatively high compared to other studies, something that possibly reflects the high burden of the disease in this region. The method optimization showed that it was possible to purify and concentrate the parasite without the use of antibodies. It was further possible to sort oocysts from the samples fairly good, with around half of the ones targeted identified as the common species *Cryptosporidium parvum*. With deep sequencing, this species was also found, along with three other variants. Based on the results in this study it seems as if the antibody free workflow works well and a high oocyst recovery is achieved. With further developments, the method could possibly be useful when studying complex samples such as wastewater.
Table of contents

Key words ............................................................................................................................... 1

Introduction ............................................................................................................................ 2
1.1 A parasite of importance with zoonotic potential ......................................................... 2
1.2 Biology and life cycle ........................................................................................................ 2
1.3 Clinical picture and pathogenesis .................................................................................. 2
1.4 Morphology, detection and diagnostics ........................................................................ 3
1.5 Taxonomy and conventional typing methodology ......................................................... 4
1.6 Epidemiology, prevention and underreporting ............................................................... 4
1.7 Cryptosporidiosis in South Africa and KwaZulu-Natal ................................................... 5
Aim ......................................................................................................................................... 7

Materials and method ............................................................................................................ 7
3.1 Purification method evaluation ....................................................................................... 7
3.2 Surface disinfection with sodium hypochlorite ............................................................. 8
3.3 Adapting the protocol for larger volumes ...................................................................... 8
3.4 Sample collection and study design ............................................................................. 9
3.5 Dissecting Cryptosporidium diversity of wastewater samples ....................................... 10
3.5.1 [SiCell workflow] Sorting of oocysts through FACS ................................................. 10
3.5.2 [SiCell workflow] Alkaline heat lysis and MDA of sorted oocysts ............................. 11
3.5.3 [SiCell workflow] Amplification of MDA product with 18S nested PCR .................... 11
3.5.4 [SiCell workflow] Purification of PCR product through ExoSap and species determination using Sanger sequencing ........................................................................... 12
3.5.5 [Amplicon sequencing] DNA extraction of pooled samples ..................................... 12
3.5.6 [Amplicon sequencing] Amplification of extracted DNA through nested 18S PCR ................................................................. 12
3.5.7 [Amplicon sequencing] Index PCR and cleanup ....................................................... 13
3.5.8 [Amplicon sequencing] Library quantification, pooling and MiSeq loading and bioinformatic analysis .......................................................... 13

Results ................................................................................................................................... 13
4.1 Flotation with saturated NaCl-glucose method gives high oocyst recovery, compared to other methods ................................................................. 13
4.2 Sodium hypochlorite solution of 10% and above efficiently kill bacteria in the faecal samples ........................................................................................................... 14
4.3 The oocyst recovery rate decreases with increased sample volume ................................ 14
4.4 Oocyst enumeration of SA samples .............................................................................. 15
4.5 [SiCell workflow] Almost half of sorted and screened single cells could be identified as Cryptosporidium ................................................................. 15

Discussion ............................................................................................................................. 18
5.1 The antibody-free workflow and its applications ........................................................... 18
5.1.1 Flotation with saturated NaCl-Glucose as opposed to other flotation fluids ............ 19
5.1.2 High oocyst recovery with saturated- NaCl-glucose flotation fluid .......................... 20
5.2 Elevated turbidity, a possible explanation for lower recovery rates with large sample volume ............................................................................................................. 21
5.3 Utilization of sampling capsules would reduce hands-on time required by the method .............................................................................................................. 21
5.4 The 46% success rate of the single cell workflow indicates potential for further developments ........................................................................................................ 21
5.5 Amplicon sequencing reveal a species composition that hints towards animal influx ....... 22
5.6 Conclusion ....................................................................................................................... 22
<table>
<thead>
<tr>
<th><strong>Key words</strong></th>
<th>Define</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSA</td>
<td>Bovine Serum Albumin</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic Local Alignment Search Tool</td>
</tr>
<tr>
<td>DNA</td>
<td>DeoxyriboNucleic Acid</td>
</tr>
<tr>
<td>DUT</td>
<td>Durban University of Technology</td>
</tr>
<tr>
<td>dNTP</td>
<td>DeoxyNucleotide TriPhosphate</td>
</tr>
<tr>
<td>ECDC</td>
<td>European Centre for Disease prevention and Control</td>
</tr>
<tr>
<td>EPA</td>
<td>United States Environmental Protection Agency</td>
</tr>
<tr>
<td>FACS</td>
<td>Fluorescence-Activated Cell Sorting</td>
</tr>
<tr>
<td>FITC</td>
<td>Fluorescein IsoThioCyanate</td>
</tr>
<tr>
<td>FSC</td>
<td>Forward SCatter</td>
</tr>
<tr>
<td>g</td>
<td>Relative centrifugal force</td>
</tr>
<tr>
<td>gp60</td>
<td>GlycoProtein 60</td>
</tr>
<tr>
<td>HIV</td>
<td>Human Immunodeficiency Virus</td>
</tr>
<tr>
<td>IMS</td>
<td>ImmunoMagnetic Separation</td>
</tr>
<tr>
<td>mAbs</td>
<td>Monoclonal AntiBodieS</td>
</tr>
<tr>
<td>MDA</td>
<td>Multiple Displacement Amplification</td>
</tr>
<tr>
<td>MM</td>
<td>Master Mix</td>
</tr>
<tr>
<td>mZN</td>
<td>Modified Ziehl-Neelson (Staining)</td>
</tr>
<tr>
<td>NaClO</td>
<td>Sodium hypochlorite</td>
</tr>
<tr>
<td>NCBI</td>
<td>National Center for Biotechnology Information</td>
</tr>
<tr>
<td>NTC</td>
<td>No Template Control</td>
</tr>
<tr>
<td>NTU</td>
<td>Nephelometric Turbidity Unit</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate Buffered Saline</td>
</tr>
<tr>
<td>PEAR</td>
<td>Paired-End reAd mergeR</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase Chain Reaction</td>
</tr>
<tr>
<td>psi</td>
<td>Pound per Square Inch</td>
</tr>
<tr>
<td>RFLP</td>
<td>Restriction Fragment Length Polymorphism</td>
</tr>
<tr>
<td>RPM</td>
<td>Revolutions Per Minute</td>
</tr>
<tr>
<td>rRNA</td>
<td>Ribosomal RiboNucleic Acid</td>
</tr>
<tr>
<td>SA</td>
<td>South Africa</td>
</tr>
<tr>
<td>SiCell</td>
<td>Single Cell</td>
</tr>
<tr>
<td>SSC</td>
<td>Side SCatter</td>
</tr>
<tr>
<td>SSU</td>
<td>Small SubUnit (18S rRNA gene)</td>
</tr>
<tr>
<td>SVA</td>
<td>Statens Veterinärmedicinska Anstalt (National Veterinary Institute)</td>
</tr>
<tr>
<td>WGA</td>
<td>Whole Genome Amplification</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
<tr>
<td>WWTP</td>
<td>WasteWater Treatment Plant</td>
</tr>
</tbody>
</table>
Introduction

1.1 A parasite of importance with zoonotic potential
The apicomplexan genus Cryptosporidium currently represents 26 recognized species of which many have a significant importance to global health (Ryan et al., 2014). Of these, 16 species and 3 genotypes have been reported to infect humans, more or less frequently (Ryan et al., 2014). Just as related genera Eimeria, and many other important protozoan parasites such as Giardia, Cryptosporidium is spread through the fecal-oral route. It was first recognized as a human pathogen in 1976 (Meisel et al., 1976) and since then its status has sailed up from a mere curiosity, to a pathogen of worldwide importance.

Of all reported global waterborne outbreaks between 2004 and 2010, Cryptosporidium was recognized as the cause in 60.3% of the cases (Baldurson & Karanis, 2011). Furthermore, when considering diarrheal disease in children below 11 months of age, where many deaths are attributed to diarrhea, Cryptosporidium has recently been established as the second most important infectious agent after rotavirus, and the single most important parasitic cause of gastroenteritis (Kotloff et al., 2013). Among children under five, diarrhea is responsible for 16% of deaths globally and Cryptosporidium is the protozoan parasite most frequently isolated from patients (WHO/UNICEF, 2009).

The only external form of the parasite, the oocyst, is hardy and remain infectious for several months in harsh environments including seawater and soil (Fayer et al., 1998: Kuczynska & Shelton, 1999). It has also been shown that many otherwise useful water treatment processes such as chlorination have little or no effect on the oocyst viability (Robertson et al., 1992). Apart from sewage contamination of drinking water, farm animals in general and young calves in particular, are thought to have a major impact on the spread of cryptosporidiosis (Kuczynska & Shelton, 1999).

1.2 Biology and life cycle
The entire developmental life cycle of Cryptosporidium takes place within one host and can be accomplished in two days (Tzipori et al., 1981). An individual is infected after ingesting the oocyst, transmitted through contaminated water, food or by direct contact (Ryan et al., 2014). Even a mechanical vector, such as flies has been implicated as a source in some cases (Conn et al., 2007). Less than ten oocysts has been shown to suffice for establishment of an infection (ID50: 10-83 oocysts for C. hominis) (Chappell et al., 2006).

After entering the gastrointestinal tract, excystation is triggered by the exposure to bile salts, pancreatic enzymes and increased temperature, resulting in a suture opening up in the oocyst wall (Fayer & Leek, 1984: Reduker et al., 1985). Each oocyst contain four sporozoites which attaches to the microvilli of the epithelial cells, leading to the development of a vacuole and attachment organelles (O'Donoghue, 1995). They are now known as trophozoites and initiate two rounds of asexual replication, merogony, with the first one producing eight merozoites that invade nearby cells (O'Donoghue, 1995).

The second round give rise to four merozoites that develop into the micro (male) and macrogametes (female), together giving rise to a zygote after sexual reproduction (O'Donoghue, 1995). Out of the sporulated oocysts that are formed as a result, around 20% have a thin and fragile wall that breaks and facilitate autoinfection of the host, while the remaining exits with the faeces (Plorde, 2004).

1.3 Clinical picture and pathogenesis
If symptomatic, the characteristic clinical signs are excessive, watery diarrhea resulting in weight loss, sometimes also accompanied by abdominal pain and vomiting. (Current & Garcia, 1991). Symptoms manifest themselves between one and two weeks after initial infection and rarely lasts longer than two
weeks (Plorde, 2004). However, both among humans and other animals, immunosuppression increases susceptibility, severity of disease and duration, turning it into a potentially life threatening affliction (Ferreira, 2000: Slapeta, 2013).

In immunocompetent hosts, the production of both thin and thick-walled oocysts eventually ceases, but when a sufficient immune response is lacking, the constant proliferation of thin-walled oocysts maintain a persistent infection despite lack of external re-introduction of oocysts (Plorde, 2004). In immunodeficient individuals it is also common that the Cryptosporidium infection disseminate from the jejunum to the biliary tract, liver and respiratory system, in addition to the continuous diarrhea leading to wasting and severe malnourishment (Leitch & He, 2011).

The pathogenesis resulting in diarrhea is not completely understood. Studies have shown a loss of absorptive surface and villus atrophy, leading to osmotic diarrhea (Tzipori & Ward, 2002: Argenzio, et al., 1990). Other contributing factors that has been seen are chloride secretion and a reduced barrier function, leading to increased permeability (Leitch & He, 2011). This appear to be a common feature in most Cryptosporidium infections, although the histopathology is often more complex in immunodeficient patients due to frequent co-infection with other pathogens (Leitch & He, 2011).

After invading the host cell, the trophozoite resides between the cell membrane and cytoplasm in its parasitophorous vacuole, where it is protected from the host immune defense while still being able to access nutrients from the cytoplasm (Tzipori & Ward, 2002).

1.4 Morphology, detection and diagnostics

Diagnosis is frequently done through microscopic detection of oocysts in faecal material and the presence of them often align with onset of clinical symptoms (O'Donoghue, 1995). In order to differentiate them from the background, two types of staining are primarily used.

The acid-fast surface of oocysts enables the use of staining technique similar to that of mycobacteria. Additionally, the fact that acid-fast organisms are rarely found in feces facilitate easy differentiation with this modified Ziehl-Neelson (mZN) method (Plorde, 2004). It is a cost efficient method, but it requires experience to identify the oocysts in the microscope (Leitch & He, 2011).

A good, albeit more expensive alternative is fluorescein isothiocyanate (FITC) -labeled monoclonal antibodies (mAbs) targeting the oocyst surface (Leitch & He, 2011). Examining the sample with epifluorescence microscopy enables rapid detection of the oocysts that are around 5 μm in size, spherical and appear bright green when excited at the 490 nm wavelength (Leitch & He, 2011: Cellabs, 2012). When compared to acid-fast staining, it has previously been shown to have higher sensitivity, particularly in samples with a low concentration of oocysts (Quilez et al., 1996: Mtambo, et al., 1992).

It can often be difficult to detect the oocysts if the faecal sample is thick and more liquid stool improves the limit of detection (Leitch & He, 2011). To make the screening easier and to concentrate the oocysts, various flotation methods are often used. A heavy fluid such as saturated sodium chloride or glucose is used to separate the larger particles from the lighter oocysts (O'Donoghue, 1995). The efficiency of different methods has been evaluated over the years and the result differs (Anderson, 1983: Mtambo et al., 1992: Kuczynska et al., 1999: McNabb et al., 1985: Kvac et al., 2003: Kar et al., 2011), probably reflecting differences in sample material, fat content being one possible factor (Weber et al., 1992). After removal of larger particles, it is usually necessary to wash the sample, both to further remove unwanted debris, but also to remove salts to enable longer storage of the oocysts (O'Donoghue, 1995). In order to pellet the oocysts, a centrifugation is necessary for an extended period of time, preferably for at least 10 minutes at 500 g (O'Donogue, 1995).
Since the different species of Cryptosporidium are almost identical in terms of morphology, it is necessary to use molecular methods for species identification (Thompson et al., 2016). The most common target is the 18S small subunit (SSU) rRNA gene that exist in several copies and have both semi conserved and hyper-variable regions (Xiao, 2010). Another popular method used for genotyping is Restriction Fragment Length Polymorphism (RFLP). Amplified fragments of the SSU is used. Differences in the DNA sequence is detected by digesting the DNA molecule with restriction enzymes, revealing a unique pattern when separating the fragments through gel electrophoresis (Xiao, 1999).

For subtyping within some of the species, analysis of the 60 kDa glycoprotein (gp60) DNA is common. The sequence is highly polymorphic with tandem repeats of serine-coding nucleotides (TCG, TCT and TCA) at the 5’ end of the gene and the variations of the sequence enable differentiation between some of the subtypes, such as the common C. hominis and C. parvum (Xiao, 2010). In addition to its high sequence variability, the protein that it codes for is a main target of host antibody attachment and it can therefore be connected to the biological properties of the specific subtype family (Xiao, 2010).

1.5 Taxonomy and conventional typing methodology

The taxonomy of Cryptosporidium is constantly being revised and debated. From previously being based on mainly host specificity (it was often elucidated that many species was not very specific and the name would therefore sometimes be considered inaccurate), it is now determined by molecular means (Thompson et al., 2016). Most of the names reflect the host that they were originally found in (Tzipori & Ward, 2002).

Slapeta specify in his taxonomic review from 2013 that the number of validated species and genotypes are 30 (Slapeta, 2013), while Ryan and colleagues consider the number to be 26, of which at least 19 infect humans (Ryan et al., 2014).

The nomenclature of the gp60 subtypes start with specifying whether its C. hominis (I) or C. parvum (II) and the designated letter, e.g. Ia indicate C. hominis subtype family A and IIb stands for C. parvum subtype family B (Sulaiman, 2005). The amounts of TCA and TCG repeats are thereafter specified, for instance IIdA20G1 indicates that it is C. parvum of allele family IId with 20 copies of the TCA repeat and one copy of the TCG repeat (Sulaiman, 2005).

1.6 Epidemiology, prevention and underreporting

Cryptosporidium affect a large number of vertebrates, such as mammals, birds, reptiles and fish with the common denominator that young individuals are often more heavily affected (Plorde, 2004). Its zoonotic potential is often seen in association with infected livestock which is considered a main reservoir but other domesticated animals such as cats and dogs are also presumed as potentially important sources (Thompson et al., 2016).

In the developing world, poor sanitation practices concerning waste disposal and water supply, along with close proximity to infected animals account for a large part of the transmission (Baldursson & Karanis, 2011b). Inadequate hygiene among family members provide an additional source for direct transmission between humans (Ryan et al., 2014). In the developed world on the other hand, contaminated drinking and recreational water and food are more common sources; for instance fruits and vegetables can be exposed through irrigation water or the handling of agricultural workers (Ryan et al., 2014).

The largest known outbreak of cryptosporidiosis occurred in April 1993, in Milwaukee, resulting in 403 000 estimated symptomatic infections (Mac Kenzie et al., 1994). It was suspected to have been caused by inadequately treated water entering Milwaukee Treatment Works from Lake Michigan (Mac Kenzie et al.,
1994). The treatment works served a population of 800 000 residents within Milwaukee county, in addition to the many visitors that could have consumed the indicated water during the time period (Hoxie et al., 1997). The estimated total number of affected citizens was based on telephone interviews covering 613 representative households, of which 30% of household members reported typical cryptosporidiosis symptoms within the relevant time period (Mac Kenzie et al., 1994). Only 285 cases of laboratory-confirmed cryptosporidiosis cases were reported, highlighting the difficulties associated with getting an accurate picture of Cryptosporidium prevalence, as people are reluctant to report and leave faecal samples (Mac Kenzie et al., 1994). Within a time period of 2 years after the outbreak, at least 54 persons had cryptosporidiosis listed as a cause of death, of which 85% were immunosuppressed (Hoxie et al., 1997).

In a review of 89 reported outbreaks of waterborne disease in England and Wales between 1992 and 2003, Cryptosporidium was connected to 69% of these, of which 39% were connected to swimming pools (Smith et al., 2006). Public swimming pools were also indicated as the source of a large outbreak in Australia in 2009, leading to cryptosporidiosis in 1141 persons (Waldron et al., 2011).

Much of the epidemiological knowledge regarding cryptosporidiosis stems from outbreak investigations and few of the people contracting disease every year are expected to seek medical care, Vijgen and colleagues estimated that only 7% do so, and in developing countries the number might be lower (Vijgen et al., 2007). In general, information about the proportion of the affected population that truly gets diagnosed is limited (Socialstyrelsen, 2014). When screening for Cryptosporidium in faecal samples, microscopic evaluation of mZN-stained material is often the method of choice, but it requires experience to detect the oocysts and hence the number of false negatives could be high (Socialstyrelsen, 2014). This problem was illustrated in a survey covering 21 laboratories in Sweden that performed Cryptosporidium diagnostics in 2014 and 2015 (Harvala et al., 2016). The three laboratories that analysed all (diarrheal or parasite-request) samples for Cryptosporidium reported a significantly higher incidence than the rest that only did so if specifically asked to. The authors of the study concluded that if all samples sent in to the other laboratories had been properly tested for Cryptosporidium, four times as many infections would have been reported on a national level. And this would still only represent a fraction of actual cases, as most are never sampled (Harvala et al., 2016).

Also within and between other European countries there is a lack of harmonization in terms of sampling practices, diagnostic methods and submission of data to the European Centre for Disease Control (ECDC) (Caccio & Chalmers, 2016). Several EU countries do not report cases routinely and Cryptosporidium is often only tested for if viral or bacterial pathogens have been excluded earlier (Caccio & Chalmers, 2016). Only seven of the countries reported 50 or more cases in 2014 and the true burden of cryptosporidiosis in Europe is therefore largely unknown (ECDC, 2014: Caccio & Chalmers, 2016). In 2004 WHO included Cryptosporidium in the Neglected Diseases Initiative, a group of diseases of global burden preventing populations in developing countries to rise from poverty (Savioli et al., 2006).

As disease is usually self-limiting in immunocompetent individuals, and availability of efficient anti-parasitic options are limited, supportive therapy is often the only treatment given if diarrhea is excessive (Leitch & He, 2011). In cases with severe and secretory diarrhea the synthetic analogue to somatostatin, Octreotide can be administered (Moroni et al., 1993). This far, the only drug shown to be more efficient than placebo and licensed for use in the United States, is Nitazoxanide (Rossignol, 2010). It is however not useful in patients with impaired immune system, where anti-retroviral therapy instead is the best option, even though cryptosporidiosis relapses may occur (Rossignol, 2010: Leitch & He, 2010).

1.7 Cryptosporidiosis in South Africa and KwaZulu-Natal
South Africa is a country with limited rainfall and increasing water pollution, at the same time as the wastewater infrastructure in many areas is unable to efficiently secure the quality of the effluent water (van
Vuuren, 2009). The alarming decrease in freshwater availability risk not only to stagnate economic growth but to put further strain on the already challenged public health sector (Adewumi et al., 2010). A total of 14% of the national water resources stems from treated sewage water, and inadequate purification is highly accountable for freshwater pollution in the country (van Vuuren, 2009). There is an interest in increasing the reuse of wastewater in many South African communities, requiring further research of wastewater composition (Adewumi et al., 2010).

Just like in many other parts of Africa, Cryptosporidium spp. is frequently isolated in South Africa from surface waters that are used for both human and animal consumption (Aldeyarbi et al., 2016). The situation is further complicated by the fact that the country accommodate a fifth of the world’s HIV positive population, with prevalence rates among the highest in the world (Karim et al., 2010).

In terms of diversity studies, *C. hominis* consistently appear to be the most frequently isolated species in humans in South Africa, followed by *C. parvum* and *C. meleagridis* (Squire & Ryan, 2017). A study of stool samples from young children with diarrhea identified 76% of *Cryptosporidium* positive samples as *C. hominis* and 20% *C. parvum* (Samra et al., 2013). Another study, conducted in the Limpopo province on samples from school children and hospital patients demonstrated 82% to be *C. hominis* and 18% *C. parvum* (Samie et al., 2006). And finally, among children living at the Kruger National Park area, 75% of *Cryptosporidium*-positive samples were *C. hominis* (Samra et al., 2016).

A majority of published articles on *Cryptosporidium* diversity concerns faecal samples obtained directly from patients or volunteers. It is important to establish how the disease burden look and it is to some extent possible to infer the source of infection based on the species found. However, considering the known importance of contaminated water for *Cryptosporidium* transmission, and that half of South African wastewater treatment plants (WWTPs) fail to meet national and international water quality standards, diversity assessment of wastewater could be important (Kuczynska & Shelton, 1999: Mthembu et al., 2013).
When studying *Cryptosporidium* in wastewater, in South Africa and elsewhere, it is often standard to use the method for oocyst recovery described in the United States Environmental Protection Agency (EPA) protocol from 1999 (Method 1623) (Sturbaum *et al.*, 2002; EPA, 1999). This method relies on extraction of the oocysts from the sample material through Immunomagnetic Separation (IMS), utilizing small magnetic beads with *Cryptosporidium*-specific mAbs attached (EPA, 1999).

There are however indications that the affinity of the antibody used might differ depending on the *Cryptosporidium* species/subtype (Personal Communications, Karin Troell, SVA). Such a difference in affinity would induce a bias when dissecting diversity of samples. In order to circumvent this problem, an antibody free workflow has previously been developed and will be used in the present study. As the wastewater collected is expected to contain several different *Cryptosporidium* populations, a single-cell approach, previously used by Troell and colleagues (Troell *et al.*, 2016), will be utilized.

**Aim**

To study *Cryptosporidium* diversity in wastewater collected from four different treatment plants in KwaZulu-Natal, South Africa and to assess the suitability of the workflow when processing large volumes of sample material.

- Collect and process material from the wastewater plants of Isipingo, Shallcross, Marianridge and Kingsburgh that receives water from mainly domestic settings but also to some extent from storm water, where an animal influx of *Cryptosporidium* could be present.
- Evaluate methods of oocyst concentration and purification to suit large amounts of complex wastewater samples with a primarily antibody-free workflow.
- Assess the diversity through a single-cell workflow, amplicon sequencing using NGS and conventional molecular typing methodology.

**Materials and method**

### 3.1 Purification method evaluation

*Cryptosporidium* positive bovine faecal material, obtained from the parasitology diagnostics department at SVA, Sweden, originating from farms around Sweden were used for optimization and evaluation. The samples were diluted with water and filtered through gauze and stored at 4°C prior to use. To prepare 300 mL of the NaCl-glucose flotation fluid, 50 g of glucose was mixed with 100 mL saturated NaCl and 150 mL purified water at room temperature, until homogenous and transparent.

Three different methods of oocyst purification through flotation were assessed in terms of oocyst recovery and suitability: flotation with saturated NaCl, saturated NaCl-glucose and a novel method referred to as Babcock sedimentation. Prior to purification, 10 μL of the thoroughly suspended faecal material was added to Teflon printed 12 mm, 3 well slide (Immuno-Cell, Mechelen, Germany) after a tenfold dilution, dried on the well and fixated by submerging the slide in acetone. The material was thereafter stained with FITC-labeled anti-*Cryptosporidium/Giardia* mAbs (Cellabs, Sydney, Australia) for 30 minutes in moisture chamber, before rinsing the antibody-solution off with PBS. For enumeration, brightly stained oocysts with typical morphology were counted in 10 viewing fields at 250 x magnification and then the average amount of oocysts per viewing field was multiplied with 177 (for counting strictly within the borders of the 12 mm well, the amount of viewing fields at 250 x magnification is 177) and divided by 10 to obtain the number of oocysts per μL. See Appendix A for full calculations.
The saturated NaCl flotation method has previously been described as highly efficient at recovering oocysts from faecal samples (Kuczynska & Shelton, 1999). For the flotation, 1 mL of faecal suspension was mixed with 3 mL of purified water in a 50 mL centrifuge tube and 4 mL saturated NaCl (room tempered, specific gravity 1.21 g/mL) was added. Following agitation by vortexing and shaking, the sample was centrifuged for 1 minute at 1540 x g. The supernatant was transferred to a new 50 mL centrifuge tube and filled up with purified water. After an additional 10 minute centrifugation at 1540 x g, around 45 mL of the supernatant was decanted, the pellet disturbed and 45 mL purified water added. This washing step was repeated an additional time and the 5 mL pellet transferred to a 15 mL tube for an additional wash, resulting in 2 mL material, of which 10 μL to be added to a Teflon printed diagnostic slide for enumeration, as previously described.

Flotation with saturated NaCl-glucose (specific gravity 1.07 g/mL) was performed in an almost identical manner as described by Maddox-Hytte and colleagues (Maddox-Hytte et al., 2006). In short, 1 mL of the suspended sample was mixed with 7 mL PBS-0.01% Tween 20 in a 15 mL centrifuge tube and vortexed intensively to emulsify the lipids. After carefully inserting a glass Pasteur pipette into the bottom of the tube, 3.5 mL of flotation fluid was added through it, to prevent oocysts from being trapped under the heavy solution. Ensuing a short centrifugation for 1 min at 80 x g, the sample was washed in the same manner as described for saturated NaCl flotation and 10 μL of the remaining 2 mL sample was processed for oocyst enumeration as described above.

For Babcock sedimentation, 20 mL of faecal suspension was added to 700 mL of purified water and 7 mL 2% H2SO4. After mixing for 5 minutes on a magnetic stirrer, it was left to settle for one hour. The clear upper layer (400 mL) was aspirated, added to 50 mL centrifuge tubes and spinned down for 10 minutes at 1540 x g. The pooled and condensed material (2 mL) was then examined as described previously.

### 3.2 Surface disinfection with sodium hypochlorite

In order to remove contaminants such as bacteria from the oocyst surface, NaClO had to be used at a concentration strong enough to disinfect, while still not damaging the oocysts. A bovine faecal sample known to have a high concentration of Cryptosporidium oocysts (enumerated prior to the evaluation) was used to assess four different concentrations of NaClO. Aliquots of 2 mL sample material was mixed with 500 μL of 100%, 50%, 10% and 1% NaClO solution (4.2g/L NaClO diluted with PBS) respectively, in a 15 mL centrifuge tube. The tube was subsequently filled up with sterile PBS-0.01% Tween20. The tubes were then immediately centrifuged at 1540 x g for 10 minutes, after which the supernatant was decanted and the pellet washed with sterile PBS (10 min centrifugation, 1540 x g) until the chlorine scent could no longer be felt, a minimum of four washes was required.

To evaluate the disinfecting effect of the different concentrations, each sample was diluted 1:20. Of the diluted sample, 200 μL was spread onto two sets of blue agar and bovine blood agar, where one plate was incubated at 20 °C and the other one at 37 °C. Furthermore each sample was plated onto Sabouraud agar, incubated at 27 °C. As a positive control, untreated sample material was spread in the same manner as NaClO-treated samples. All plates were left to incubate for four days before being examined.

### 3.3 Adapting the protocol for larger volumes

As the flotation with saturated NaCl-glucose was deemed most suitable for the present study (see section 4.1), but had only been assessed using small volumes of sample material, a modification of the protocol was necessary to suit the large amounts of material generated when sampling wastewater.

The same bovine sample material used for previous evaluation of flotation methods were used; 4 mL sample material (28 mL PBS-0.01% Tween20 and 14 mL flotation fluid), 8 mL (26 mL PBS-0.01%
Tween20 and 12 mL flotation fluid) and 12 mL (24 mL PBS-0.01% Tween20 and 10 mL flotation fluid) were allocated into 50 mL centrifuge tubes, as opposed to the 15 mL tubes previously used. The oocyst recovery between the methods was compared. For the up-scaled protocol, 1.7 mL of 10% NaClO was used for surface disinfection.

3.4 Sample collection and study design

During the first two weeks of March 2017, wastewater samples were collected from four different treatment plants, within the eThekwini municipality (KwaZulu-Natal, South Africa). The work was performed together with researchers at the Institute for Water and Wastewater Technology at Durban University of Technology (DUT). Permission to collect samples from the facilities had been granted beforehand.

Two of the four different WWTPs visited, Isipingo and Kingsburgh are located in the southern parts of the city of Durban. Both plants collect from domestic sources, i.e. mainly sewage waste from the households around the area but also storm runoff to some extent as parts of the sewage collecting network is not enclosed (Mzulwini, 2004). Due to the elevation of the surrounding area, the Isipingo WWTP collects through gravitation and handles large volumes of wastewater, 17 million liter per day, compared to the 3 million liters processed each day by Kingsburgh WWTP (Mzulwini, 2004).

The other two sites, Mariannridge and Shallcross WWTPs are parallel plants with different collection sources of wastewater, but a combined effluent outlet in the Umhlatuzana River (together the plants are referred to as the Umhlatuzana works). Mariannridge WWTP receive wastewater from both industrial outlets (30%) and domestic (70%), handling a total volume of 8 million liters per day (Mhlanga et al., 2009). Shallcross on the other hand only source from domestic sewage (2 million liters/day) (Mhlanga et al., 2009). See Figure 1 for the geographical location of the WWTPs within the municipality.

Due to differences in construction of the plants in regards to their treatment processes, it was not possible to access sludge from all of them. Sample collection was done with a metallic sampling stick, rinsed in the sampling material prior to collection. The same procedure was done with the plastic jugs used for transportation of the material, to prevent contamination from previous samplings. See Table 1 for information regarding the samples collected.

**Table 1.** Samples were collected between 2/3 and 13/3 2017 from mainly influent wastewater source but also from untreated sludge deposits and final effluent wastewater

<table>
<thead>
<tr>
<th>Treatment plant</th>
<th>Sample type</th>
<th>Volume collected</th>
<th>Date collected</th>
<th>Designated name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isipingo</td>
<td>Influent</td>
<td>2.5 liter</td>
<td>2017-03-02</td>
<td>Isi. Inf. 1</td>
</tr>
<tr>
<td>Isipingo</td>
<td>Influent</td>
<td>2.5 liter</td>
<td>2017-03-02</td>
<td>Isi. Inf. 2</td>
</tr>
<tr>
<td>Isipingo</td>
<td>Sludge</td>
<td>0.25 liter</td>
<td>2017-03-13</td>
<td>Isi. Slu. 1</td>
</tr>
<tr>
<td>Isipingo</td>
<td>Sludge</td>
<td>0.25 liter</td>
<td>2017-03-13</td>
<td>Isi. Slu. 2</td>
</tr>
<tr>
<td>Isipingo</td>
<td>Sludge</td>
<td>0.25 liter</td>
<td>2017-03-13</td>
<td>Isi. Slu. 3</td>
</tr>
<tr>
<td>Isipingo</td>
<td>Sludge</td>
<td>0.25 liter</td>
<td>2017-03-13</td>
<td>Isi Slu. 4</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Influent</td>
<td>2.5 liter</td>
<td>2017-03-07</td>
<td>King. Inf. 1</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Influent</td>
<td>2.5 liter</td>
<td>2017-03-07</td>
<td>King. Inf. 2</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Influent</td>
<td>2.5 liter</td>
<td>2017-03-07</td>
<td>King. Inf. 3</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Influent</td>
<td>2.5 liter</td>
<td>2017-03-07</td>
<td>King. Inf. 4</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Final effluent</td>
<td>5 liter</td>
<td>2017-03-07</td>
<td>King. Eff. 1</td>
</tr>
<tr>
<td>Location</td>
<td>Type</td>
<td>Volume</td>
<td>Date</td>
<td>Code</td>
</tr>
<tr>
<td>------------</td>
<td>--------------</td>
<td>--------</td>
<td>------------</td>
<td>------------</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Final effluent</td>
<td>5 liter</td>
<td>2017-03-07</td>
<td>King. Eff. 2</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Sludge</td>
<td>0.25 liter</td>
<td>2017-03-07</td>
<td>King. Slu. 1</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Sludge</td>
<td>0.25 liter</td>
<td>2017-03-07</td>
<td>King. Slu. 2</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Sludge</td>
<td>0.25 liter</td>
<td>2017-03-07</td>
<td>King. Slu 3</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Sludge</td>
<td>0.25 liter</td>
<td>2017-03-07</td>
<td>King. Slu 4</td>
</tr>
<tr>
<td>Mariannridge</td>
<td>Influent</td>
<td>2.5 liter</td>
<td>2017-03-13</td>
<td>Mar. Inf. 1</td>
</tr>
<tr>
<td>Shallcross</td>
<td>Influent</td>
<td>2.5 liter</td>
<td>2017-03-13</td>
<td>Shall. Inf. 1</td>
</tr>
</tbody>
</table>

All collected samples were processed to rid the material of large particles and to get a suitable consistency for downstream processing. In brief, the influent wastewater material was filtered through a 20 μm metallic sieve and condensed through centrifugations at 3000 x g for 5 minutes (4°C) to obtain a semi-solid material suitable for subsequent flotation. For the sludge samples, a 25 μm metallic sieve was used, after it had been diluted with purified water, due to its thicker consistency.

For the effluent water samples, Envirocheck Sampling Capsules (PALL Corporation, New York, United States) was used according to the producer’s specification. It is a filter cartridge specifically designed for the recovery of Cryptosporidium and Giardia. When water is passed through it, the oocysts are trapped on the 1 μm pore size membrane and several liters of water can be processed per capsule. Afterwards, an elution buffer consisting of mainly PBS and an anti-foaming agent is used to rinse the filter.

During the entire process, care was taken to only use purified water. Equipment used were treated with UV light over-night, after thorough cleaning and soaking in 70% ethanol.

For the flotation, each 50 mL centrifuge tube had 10 mL material added to it, 10 mL saturated NaCl-glucose flotation fluid and 30 mL PBS-0.01% Tween20, according to the methodology previously described (section 3.3). All of the material annotated for each sample in Table 1 was used for flotation. Following three washes, the samples were treated with 10% NaClO solution (described in section 3.2).

The NaClO–treated samples were diluted with sterile PBS and filtered through a nylon net filter (Merck Millipore, Billerica, United States) with a pore size of 20 μm, followed by an additional filtration through a 2 μm filter piece. This to first remove everything larger than the 4-6 μm oocysts and then trap them on the smaller filter, while small particles are filtered through.

Pooled samples were enumerated according to previously described procedure, with the difference that the entire well of the Teflon printed diagnostic slide was scanned for Cryptosporidium at 100 x magnification and discovered oocysts confirmed at 400 x magnification.

The samples were heat inactivated at 65°C for 5 minutes, prior to transportation back to Sweden.

3.5 Dissecting Cryptosporidium diversity of wastewater samples

Two different methods were used to answer this question, namely typing of individual oocysts sorted through the single cell (SiCell) platform at Science for Life Laboratory, Uppsala and amplicon sequencing of extracted bulk DNA.

3.5.1 [SiCell workflow] Sorting of oocysts through FACS

Sample King. Inf. 2 was selected for sorting with MoFlo Astrios cell sorter (Beckman Coulter, Brea, United States). Sample selection was based on low turbidity and comparatively high oocyst count (established before surface disinfection, calculated to 2600 oocysts per mL flotated material). The sample was diluted
1:10 (30 μL material in PBS) prior to sorting. Sorting region was based on side scatter (SSC) signal. As no distinct populations was detected due to a large amount of background debris, it was decided that further purification attempts were needed along with inclusion of FITC-labeled mAbs (Cellabs, Sydney, Australia).

Another sample from the same wastewater treatment plant, King. Slu. 4 was selected with an estimated concentration of 1400 oocysts per mL (flotated material). The sample was filtered an additional time through a 2 μm nylon net filter (Merck Millipore, Billerica, United States) in order to minimize small particles that would overlap with the oocysts in the scattergram.

After a dilution of the sample 1:4 with PBS and incubation with 5 μL FITC-antibody in 4 mL sample, a tube sort of the material was initiated with enrich sorting mode and 1-2 drop envelope. Tube sort is performed to collect many particles, within the given parameters, in an Eppendorf tube. This sample can then be further processed as a bulk sample or sorted again, the second time with individual particles into individual wells. Sorting selection was based on two predefined regions in the scatter plot suspected to represent the FITC fluorescence and size of Cryptosporidium (measured by side scatter). The aforementioned MoFlo Astrios EQ sorter was used, with a 70 μm nozzle and a 488 nm laser. For sheath fluid, 1% filtered NaCl was used with the sheath pressure of 60 psi.

Following the tube-sort, a re-sort of the enriched material was done with single cell sorting mode and 0.5 drop envelope. Individual oocysts were sorted into 44 wells of a 96-well plate. In two of the wells, 5 cells each were collected and 9 wells were selected to be NTC.

3.5.2 [SiCell workflow] Alkaline heat lysis and MDA of sorted oocysts

Lysis buffer D2 (x1, REPLI-g Mini Kit) was prepared by adding 0.25 μL DTT 1M to 2.75 μL Buffer DLB (Qiagen, Hilden, Germany), containing potassium hydroxide and nuclease free water. After adding 1 μL of the prepared alkaline buffer D2 to each of the sorted oocysts, the plate was incubated at 95°C for 45 seconds and then left on ice for 10 minutes, after which 1 μL of the neutralization buffer, REPLI-g Stop Solution (Qiagen, Hilden, Germany) was added.

For the whole genome amplification, MDA, RepliPHI Phi29 Reagent set (1.0μg/μl) (Epicentre, Madison, United States) was used. In brief, the reaction was set up at 30 °C for 16 hours in 15 μl reactions with a final concentration of 1x reaction buffer, 1.6 mM dNTPs, 50 μM Phi29 random hexamer primers with phosphorothioate bonds at the 3’end (Integrated DNA Technologies, Coralville, United States), 10 μM DTT, 5% DMSO, 40U Phi 29 enzyme, 0.5 μM SYTO13 (Life Technologies, Carlsbad, United States) and water. Lysis reagents, stop solution and MDA mix (excluding SYTO13) had previously been decontaminated with UV at 2 J in a Stratalinker. The reaction ended in a heat-inactivation step lasting for 3 minutes and the amplified DNA was thereafter stored at -20 °C until further use.

3.5.3 [SiCell workflow] Amplification of MDA product with 18S nested PCR

Fourteen of the wells, including a negative control, were selected for evaluation and screening. From each well, 2 μL material was diluted 20-fold in sterile water and subjected to Cryptosporidium specific 18S nested PCR.

The first amplification reaction contained 1x AmpliTaq Gold Buffer (Thermo Fisher Scientific, Waltham, United States), 3 mM MgCl2, 0.2 mM dNTP mix, 0.16 mg/mL bovine serum albumin (BSA), 200 μM of forward and reverse primer (5’-TTCTAGAGCTAATACATCCG-3’ and 5’-CCCATTCTCTTGAAACA GGA-3’), 1 unit of AmpliTaq Gold Polymerase (Thermo Fisher Scientific, Waltham, United States) and 2 μL of diluted MDA product in a total volume of 25 μL. Subsequent to an initial denaturation for 5 minutes
at 95°C, 35 cycles ensued, comprised of 94°C for 45 sec, 56°C for 45 sec and 72°C for 1 minute, followed by a final extension at 72°C for 7 minutes.

For the second amplification reaction, 2 μL from the first reaction was used as DNA template and added to a reaction mix identical to the first with the exception that BSA had been omitted and secondary primers were used. The reaction conditions were also the same as in the first reaction, with the exception that it consisted of 40 cycles with an annealing temperature of 58°C. Amplified products were detected by UV transillumination after electrophoretic size separation on a 1% agarose gel.

3.5.4 [SiCell workflow] Purification of PCR product through ExoSap and species determination using Sanger sequencing
Out of the PCR products showing positive bands on the gel, 5 μL was mixed with 0.5 μL Exonuclease I (Exo I, 20 u/μL) and 1 μL SAP (Fermentas FastAP Thermosensitive Alkaline Phosphatase 1u/μL) and incubated at 37°C for 15 minutes, followed by 85°C for an additional 15 min.

The purified PCR-product (2 μL) was added to a sequencing mix consisting of 1.0 μL BD stock, 0.5 μL 5X buffer and 1.5 μL internal primers (5pmol/μL), with a total reaction volume of 5 μL. The product was amplified in 9800 Fast Thermal Cycler for 25 cycles of 96°C for 10 seconds, 50°C for 10 seconds and 60°C for 30 seconds. It was subsequently cleaned with Montage Seq96 Sequence Reaction Cleanup kit (Merck Millipore, Billerica, United States) according to the user guide (Merck Millipore, 2016) before sequencing in both directions with 3130xl genetic analyzer (Applied Biosystems, Foster City, United States).

BioEdit Sequence Alignment Editor was used to assemble consensus sequences that were compared with sequences available in GenBank (NCBI) using Basic Local Alignment Search Tool (BLAST).

3.5.5 [Amplicon sequencing] DNA extraction of pooled samples
In order to get representative material from each of the four wastewater plants, every sample stemming from the same location was pooled, resulting in four tubes. The material was condensed through centrifugation and DNA was extracted using MoBio Powersoil DNA Isolation Kit (QIAGEN, Hilden, Germany) according to the manufacturer’s instructions.

Briefly, approximately 250 μg of sample material was homogenized and lysed using the supplied PowerBead Tubes and a buffer dissolving humic acids along with solution C1, facilitating complete cell lysis. Bead beating was performed using MP FastPrep-24 5G (MP Biomedicals, Solon, United States) for 2 x 60 seconds at 2400 RPM. After a brief centrifugation, the supernatant was collected and mixed with inhibitor removal solution C2 that precipitates non-DNA substances. Following an additional wash with a second inhibitor removal, the collected supernatant was mixed with a high concentration salt solution (C4). This to enable binding of DNA to the silica membrane used when washing the DNA with ethanol. Finally, after all residual ethanol had been removed through centrifugation, the DNA was released from the silica membrane with the sterile elution buffer included in the kit.

3.5.6 [Amplicon sequencing] Amplification of extracted DNA through nested 18S PCR
The first amplification was performed as described in section 3.6.2, using the same primers, reaction conditions and components. For the second amplification, custom designed primers were used, ordered from Integrated DNA Technologies (Coralville, United States). These primers have 3’-end targeting Cryptosporidium 18S and an Illumina adaptor in the 5’-end, designed according to the specifications of Illumina (Illumina, 2016).
Forward sequence used was 5’- TCG TCG GCA GCG TCA GAT GTG TAT AAG AGA CAG TAT CAA TTG GAG GGC AAG TC -3’ and reverse 5’- GTC TCG TGG GCT CGG AGA TGT GTA TAA GAG ACA GTC CCC TAA CTT TCG TTC TTG -3’. The second amplification reaction contained 2 μL amplified material from the first reaction, 1x AmpliTaq Gold Buffer (Thermo Fisher Scientific, Waltham, United States), 3 mM MgCl₂, 0.2 mM dNTP mix, 200 μM of forward and reverse primer, 1 unit of AmpliTaq Gold Polymerase (Thermo Fisher Scientific, Waltham, United States). The conditions were an initial denaturation for 5 minutes at 95°C, 35 cycles of 94°C for 45 sec, 58°C for 45 sec and 72°C for 1 minute, followed by a final extension at 72°C for 7 minutes. The amplified products were visualized through agarose gel (1%) electrophoresis and UV transillumination.

3.5.7 [Amplicon sequencing] Index PCR and cleanup
Dual indices and Illumina sequencing adapters were added using the Nextera XT Index Kit, according to the procedure described in the Illumina Library Prep Guide (Illumina, 2016).

In short, a PCR reaction was set up containing 1x AmpliTaq DNA Polymerase Buffer (Thermo Fisher Scientific, Waltham, United States), 0.2 mM dNTP mix, Nextera XT Index 1 Primers (N7XX, 5 μL per sample), Nextera XT Index 2 Primers (S5XX, 5 μL per sample) and 1 unit of AmpliTaq DNA Polymerase (Thermo Fisher Scientific, Waltham, United States) in a reaction volume of 50 μL in total.

The PCR reaction conditions consisted of an initial denaturation of 95 °C for 3 minutes, 8 cycles of 95 °C for 30 sec, 55 °C for 30 sec, 72 °C for 30 sec and a final elongation at 72 °C for 5 minutes. The product was then cleaned using the AMPure XP beads, according to the guide (Illumina, 2016), however 25 μL of beads were added to each well instead of the stated 56 μL. The final library was validated on a Bioanalyzer DNA 1000 chip.

3.5.8 [Amplicon sequencing] Library quantification, pooling and MiSeq loading and bioinformatic analysis
The samples were diluted to a final concentration of 12 nM, denatured using NaOH, diluted to 4 pM and finally pooled in equal amounts before cluster generation and sequencing on Illumina miSeq Benchtop Sequencer (Illumina, San Diego, United States).

The reads generated (representing the targeted 18S SSU rRNA gene region, 2 x 250 bp) were paired using PEAR (Paired-End reAd mergeR) and 18S fragments were clustered using the rapid sequence analysis tool Usearch. Resulting clusters were compared to the NCBI nt database. To ensure an accurate and contemporary identification, sequences were aligned with reference sequences using ClustalW Multiple alignment tool in BioEdit software version 7.2.6. Upon discrepancies between the query and reference sequences, NCBI BLAST algorithm was used to find other potential identities.

Results

4.1 Flotation with saturated NaCl-glucose method gives high oocyst recovery, compared to other methods
The three methods of oocyst purification were compared as described in section 3.1. All methods were sufficient to recover Cryptosporidium oocysts but the recovery efficiency varied greatly between the methods (Table 2) It was apparent that flotation with saturated NaCl and glucose resulted in the highest recovery of the tested methods, based on these data. It was therefore selected as the method of choice for the subsequent wastewater sample purification in the present study.
Table 2. Comparison of flotation methods in terms of oocyst recovery. The same faecal sample was used to compare the three different methods. The values obtained from the saturated NaCl method and NaCl-glucose was an average of 4 and 2 replicates respectively, while the oocyst count for Babcock sedimentation was based on only one replicate. See Appendix A for full calculations.

<table>
<thead>
<tr>
<th>Method</th>
<th>Oocysts in</th>
<th>Oocysts out</th>
<th>Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Babcock Sed.</td>
<td>93.6 x 10⁶</td>
<td>11.2 x 10⁶</td>
<td>12.0 %</td>
</tr>
<tr>
<td>Sat. NaCl</td>
<td>4.7 x 10⁶</td>
<td>1.0 x 10⁶</td>
<td>21.6 %</td>
</tr>
<tr>
<td>Sat. NaCl-Glu</td>
<td>4.7 x 10⁶</td>
<td>4.1 x 10⁶</td>
<td>87.8 %</td>
</tr>
</tbody>
</table>

4.2 Sodium hypochlorite solution of 10% and above efficiently kill bacteria in the faecal samples

Cryptosporidium-positive sample aliquots (2 mL) was treated with 500 μL of 100%, 50%, 10% and 1% NaClO solution (4.2g/L NaClO diluted with PBS) respectively. A control consisting of unchlorinated (untreated) sample was also included. This, with the purpose of evaluating the disinfecting properties of the different NaClO concentrations. See section 3.2.

After four days of incubation, growth was observed on the bovine blood and blue agar plates treated with 1% NaClO solution (final concentration 0.2% NaClO), incubated at 37 °C (See Appendix B, Figure 1). Bacterial growth could also be seen on all agar plates spread with unchlorinated (referred to as untreated) flotated sample (Appendix B, Figure 2), except for the Sabouraud agar where there was no growth detected.

No growth was observed at any of the other plates, leading to the conclusion that all NaClO solutions with a concentration of 10% or above, are sufficient to kill bacteria in flotated faecal samples. Observation of the oocysts through epifluorescence microscopy did not indicate any oocyst damage due to the NaClO exposure. However, considering that oocysts found in wastewater might be in poor condition, it was decided to use the lowest concentration proven to disinfect, i.e. 10% NaClO, during the continuation of the present study.

4.3 The oocyst recovery rate decreases with increased sample volume

As described in section 3.3, three different volumes of sample material were evaluated and a marked decrease in oocyst recovery rate could be seen for the largest volume used, 12 mL (in 50 mL centrifugation tubes), compared to when 1 mL was used in the original protocol (See Table 3). For full calculations, refer to Appendix A, section A3. More than half of the oocysts were however recovered in all tested sample volumes and the possibility to process larger volumes of material (a prerequisite for the present study) was considered to restitute for this decrease.

It was observed that when 12 mL of material was used it was difficult to distinguish between the supernatant to be extracted and the underlying layer of flotation fluid. For practical reasons, the processing of wastewater sample material collected in South Africa, was decided to volumes of 10 mL. Volumes of PBS-0.01% Tween 20 and flotation fluid (Saturated NaCl-Glucose) was adjusted accordingly, see section 3.4.
Table 3. Oocyst recovery rate when three different volumes of material is flotated with saturated NaCl-glucose protocol (adapted for larger volumes). Sample used for the evaluation had 4.7 x 10^6 oocysts per mL.

<table>
<thead>
<tr>
<th>Volume used</th>
<th>Oocysts in</th>
<th>Oocysts out</th>
<th>Recovery</th>
</tr>
</thead>
<tbody>
<tr>
<td>4 mL</td>
<td>18.7 x 10^6</td>
<td>15.6 x 10^6</td>
<td>83.1 %</td>
</tr>
<tr>
<td>8 mL</td>
<td>37.4 x 10^6</td>
<td>28.2 x 10^6</td>
<td>75.3 %</td>
</tr>
<tr>
<td>12 mL</td>
<td>56.2 x 10^6</td>
<td>35.3 x 10^6</td>
<td>62.8 %</td>
</tr>
</tbody>
</table>

4.4 Oocyst enumeration of SA samples

Due to a number of logistic and practical circumstances it was only possible to make a few attempts at oocysts enumeration of the samples collected in SA. The result after screening the entire well and extrapolating the count to the amount of oocysts per liter wastewater can be seen in Table 4.

Table 4. Estimated concentration of oocysts in collected wastewater, based on enumeration of oocysts. The sample number, if multiple samples exists, is indicated together with sample type. Flot = Flotation. NaClO = Surface disinfection with 10% NaClO solution. Filt = Filtration with 20 μm and 2μm pore size filter.

<table>
<thead>
<tr>
<th>WWTP</th>
<th>Type of sample</th>
<th>Treatment</th>
<th>Oocysts per L</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isipingo</td>
<td>Influent (#1+2)</td>
<td>Flotation</td>
<td>3 167</td>
</tr>
<tr>
<td>Isipingo</td>
<td>Influent (#1+2)</td>
<td>Flot+NaClO</td>
<td>1000</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Sludge (#1+2)</td>
<td>Flotation</td>
<td>84 000</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Influent (#1)</td>
<td>Flotation</td>
<td>38 280</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Influent (#2)</td>
<td>Flotation</td>
<td>15 600</td>
</tr>
<tr>
<td>Kingsburgh</td>
<td>Sludge (#4)</td>
<td>Flot+NaClO+Filt.</td>
<td>11 200</td>
</tr>
<tr>
<td>Isipingo</td>
<td>Sludge (#1)</td>
<td>Flot+NaClO+Filt.</td>
<td>8 800</td>
</tr>
<tr>
<td>Marrianridge</td>
<td>Influent</td>
<td>Flot+NaClO+Filt.</td>
<td>240</td>
</tr>
<tr>
<td>Shallcross</td>
<td>Influent</td>
<td>Flot+NaClO+Filt.</td>
<td>400</td>
</tr>
</tbody>
</table>

4.5 [SiCell workflow] Almost half of sorted and screened single cells could be identified as Cryptosporidium

Roughly half of the 96-well sorting plate (55 wells) was allocated for the cell sorting, of which 44 were single-cell wells and two contained five cells. In addition, nine wells were selected to be negative controls. The Ct statistics for these three groups after genome amplification with MDA can be seen in Table 5. For Cryptosporidium specific 18S amplification, 14 wells were selected, including the two 5-cell wells, eight single-cell wells and one negative control.

Out of these 14, six were considered positive based on bands of correct size on the agarose gel. These were all single-cell wells and had an average Ct value of 20. This amplified material was further processed and typed with Sanger sequencing. All sequenced cells could be identified as C. parvum based on the 18S rRNA gene sequence. In regards to these 13 sorted oocysts, the success rate of the full single workflow (sorting, lysis, WGA and screening) was 46%.
Table 5. Ct values for the three different categories of wells containing either five, one or zero cells (oocysts), after MDA.

<table>
<thead>
<tr>
<th>Description</th>
<th>Number of wells</th>
<th>Avg C(t)</th>
<th>Max C(t)</th>
<th>Min C(t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 cells</td>
<td>2</td>
<td>15.95</td>
<td>19.94</td>
<td>11.97</td>
</tr>
<tr>
<td>1 cell</td>
<td>44</td>
<td>20.16</td>
<td>27.37</td>
<td>13.05</td>
</tr>
<tr>
<td>MDA neg</td>
<td>9</td>
<td>22.84</td>
<td>33.43</td>
<td>17.35</td>
</tr>
</tbody>
</table>

4.6 [Amplicon sequencing] A total of four different species and genotypes could be detected in the samples

The amplicon sequencing generated between 18 396 and 60 259 reads per sample, translating into 67-340 unique sequence identities per sample. The fraction of Cryptosporidium reads detected compared to other genera varied between 46% and 0.7%, with C. parvum making up 83.8% of all Cryptosporidium -reads. Other genera frequently picked up by the sequencing included Blastocystis, Candida, Polytoma, Trichosporon and Cryptococcus.

The pooled material from Isipingo WWTP generated a total of 60 259 reads, of which 27 956 (46%) corresponded to Cryptosporidium spp (See Table 6). As described in section 3.7.3, reads were compared to NCBI GenBank nucleotide database and assigned an identity. Presented in table 5 are the names given in GenBank. However, these might be misleading if not recently deposited since the Cryptosporidium species are constantly being revised. The time for the project did not allow this to be thoroughly edited for all sequences.

A majority of the crypto reads from this WWTP were associated with GenBank submission ID AF093490, annotated as ‘Cryptosporidium parvum, strain Bovine C. parvum genotype’. As stated above, some submissions in GenBank are old, and C. parvum bovine genotype is no longer used. Therefore the acquired sequence was compared with trusted reference sequences from other species/genotypes and was confirmed to align with C. parvum IOWA AF164102, a well-known C. parvum isolate.

A smaller proportion of the reads belonged to what was determined to be C. hominis and deer mouse genotype III. In regards to the former, the query sequence was almost identical to C. hominis reference sequence EF570922 with the exception of an A to T nucleotide substitution in the codon following the AAT TTA TAT stretch in the middle of the sequence. This type of minor sequence heterogeneity is often found in the various NCBI submissions of the same species/genotype. It could be useful to keep the submissions apart, should the taxonomy be modified in the future.

Deer mouse genotype III was previously known as genotype W1 as it was found in storm water collected from watersheds in New York and the host was unknown (Jiang et al., 2005; Jiang et al., 2006) but finally established to be rodents of the genus deer mouse (Peromyscus sp.) and Eastern gray squirrel (Sciurus carolinensis) (Feng et al., 2007). These three studies were all conducted on samples collected in connection to watersheds used by New York City water supply. The reads generated in the present study matched the sequence submitted by Feng and colleagues (KT027464) with the characteristic TCA CAA TAA stretch that differ it from C. hominis, parvum, wrairi and other species/genotypes. The reference sequence does however contain a silent substitution 10 codons upstream of this stretch (TCA replaced with TCG) but other than that no deviations of relevance.

Table 6. The NCBI database annotations of the corresponding sequences obtained after amplicon sequencing on miSeq (Illumina) platform and contemporary species/genotype determination for Isipingo pooled samples. Identity confirmed with the help of the following, matching reference sequences: C. parvum AF164102, Deer mouse genotype III KT027464 and C. hominis EF570922.
As seen in Table 7, the material from Kingsburgh WWTP generated the largest number of *Cryptosporidium* species identities. Out of 27 535 reads, 11 505 (41%) was confirmed to belong to *Cryptosporidium* spp., most prominent being *C. parvum* and mink genotype, but also *C. hominis* and deer mouse genotype III was found. As with the results from Ispingo some matching sequences were from older submissions. Identity was confirmed with the help of the aforementioned reference sequences: *C. parvum* AF164102, deer mouse genotype III KT027464, *C. hominis* EF570922 and also mink genotype JX471002, found for the first time in the study by Feng and colleagues (Feng et al., 2007). There were some degrees of sequence heterogeneity for many of the clustered reads, as mentioned in the table legend.

### Table 7. The NCBI database annotations of the corresponding sequences obtained after amplicon sequencing on miSeq (Illumina) platform and species/genotype determination for Kingsburgh pooled samples.

<table>
<thead>
<tr>
<th>NCBI Accession ID</th>
<th>Number of reads</th>
<th>Species/Genotype Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF093490</td>
<td>27 861</td>
<td><em>Cryptosporidium parvum</em></td>
</tr>
<tr>
<td>EF641014</td>
<td>77</td>
<td><em>Cryptosporidium sp. deer mouse genotype III</em></td>
</tr>
<tr>
<td>gi</td>
<td>937959809</td>
<td>18</td>
</tr>
<tr>
<td>gi</td>
<td>930425855</td>
<td>4149</td>
</tr>
<tr>
<td>EF641015</td>
<td>4147</td>
<td><em>Cryptosporidium sp. mink genotype</em></td>
</tr>
<tr>
<td>AF093490</td>
<td>2221</td>
<td><em>Cryptosporidium parvum</em></td>
</tr>
<tr>
<td>gi</td>
<td>636633525]</td>
<td>499</td>
</tr>
<tr>
<td>gi</td>
<td>63109354]</td>
<td>499</td>
</tr>
<tr>
<td>gi</td>
<td>805308682]</td>
<td>443</td>
</tr>
<tr>
<td>gi</td>
<td>701218921]</td>
<td>23</td>
</tr>
<tr>
<td>gi</td>
<td>937959809]</td>
<td>9</td>
</tr>
<tr>
<td>EF641014</td>
<td>9</td>
<td><em>Cryptosporidium sp. deer mouse genotype III</em></td>
</tr>
<tr>
<td>AF093491</td>
<td>5</td>
<td><em>Cryptosporidium hominis</em>(^6)</td>
</tr>
</tbody>
</table>

Deer mouse genotype III was also identified in the pooled sample from Mariannridge (Table 8), where it was the only *Cryptosporidium* species found. Out of the 18 396 reads generated, *Cryptosporidium*
accounted for 11%. Reads from the yeast *Trichosporon montevideense* made up 82% of the total reads from this sample.

Table 8. The NCBI database annotations of the corresponding sequences obtained after amplicon sequencing on miSeq (Illumina) platform and species/genotype determination for *Mariannridge pooled samples*. Identity confirmed with reference sequence for deer mouse genotype III KT027464.

<table>
<thead>
<tr>
<th>NCBI Accession ID</th>
<th>Number of reads</th>
<th>Species/Genotype Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>EF641014</td>
<td>2046</td>
<td><em>Cryptosporidium</em> sp. deer mouse genotype III</td>
</tr>
</tbody>
</table>

From the 30 351 reads obtained out of the Shallcross pooled material (Table 9), an absolute majority were identified as the chlorophyte *Polytoma uvella* and only 0.7% as *Cryptosporidium*, namely *C. parvum*.

Table 9. The NCBI database annotations of the corresponding sequences obtained after amplicon sequencing on miSeq (Illumina) platform and species/genotype determination for *Shallcross pooled samples*. Identity confirmed with *C. parvum* Iowa reference sequence AF164102.

<table>
<thead>
<tr>
<th>NCBI Accession ID</th>
<th>Number of reads</th>
<th>Species/Genotype Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF093490</td>
<td>221</td>
<td><em>Cryptosporidium parvum</em></td>
</tr>
</tbody>
</table>

Discussion

5.1 The antibody-free workflow and its applications

It has been shown in the present study that the optimized antibody-free workflow is a viable way to study *Cryptosporidium* diversity in wastewater samples. Compared to the commonly used antibody dependent IMS method, this workflow is more time consuming. However, when compared to parallel runs of the same samples (collected and processed by the Institute of Waste and Wastewater Technology, Durban University) where oocysts had been extracted through IMS, it appears as if the antibody free workflow yields higher recovery of oocysts. One example is Kingsburgh WWTP, where six previous samplings of influent wastewater during 2015-2016 yielded an average of 189 oocysts per liter (median 136 oocysts/L) (Adeyemo, F.E., Unpublished data), while the present study showed 38 280 and 15 600 oocysts per liter for the two influent samples enumerated. Results from the same study also included 6 samplings of Isipingo WWTP (359 oocysts/L, median 226), Mariannridge (133 oocysts/L, median 124) and Shallcross (163 oocysts/L, median 144) as compared to 3167, 240 and 400 oocysts/L respectively in the present study. It should be noted that for the Mariannridge and Shallcross enumeration, the samples had been processed with the entire workflow (including chlorination and filtration) before enumeration, and a significant loss of oocysts can be expected to have occurred, compared with merely floated samples. Hence, an even more pronounced difference can be expected between the two methods.

In this study we obtained unexpectedly high parasite counts in all samples. Since the results are based on simply one sampling occasion and only few samples, the results should be interpreted with caution. For instance, the aforementioned oocyst count of almost 28 000 oocysts per liter can be compared with the 27 000/liter peak value measured in influent wastewater during the large cryptosporidiosis outbreak in Östersund, Sweden (Smittskyddsintitutet, 2011). During this outbreak a large percentage of the city of
Östersund was ill with acute cryptosporidiosis and can be expected to have excreted high numbers of oocysts. Although Cryptosporidium is considered endemic in the Durban area, these high values are probably not representative for the overall oocyst concentration in the wastewater. With that being said, it would require a 99% reduction of the lowest Kingsburgh (flotated) sample concentration obtained in the present study, to reach the average concentration found in previous assessments based on IMS, 189 oocysts/L (Adeyemo, F.E., Unpublished data).

In samplings performed by Umgeni Water at 83 other locations around Kwazulu-Natal during 2016, of which 48 included raw influent wastewater, the highest concentration obtained was 19 oocysts/L (Unpublished data, Ian Bailey, Umgeni Water). Ten liters of water was collected and concentrated with Envirocheck sampling capsules at each of the sites and at most locations Cryptosporidium was not detected at all. In a study during spring 2008 in Guateng province, the influent wastewater of Zeekoevat, Baviaanspoort, Rayton and Refilwe WWTPs contained average oocyst concentrations of 37, 26, 21 and 43 per liter of sampled water respectively (Dungeni & Momba, 2010). On a weekly basis, 2 L of wastewater was collected, filtered and oocysts were captured through IMS, yielding values between 0 and 150 oocysts per liter (Dungeni & Momba, 2010).

The workflow of the present study is slightly more time consuming than extracting oocysts with IMS, however it is significantly cheaper. In addition, the suspicion that there are differences in antibody affinity, which would affect what is detected and not, in for example diversity and outbreak studies, the antibody-free workflow can be a useful replacement. The price of purchasing an IMS kit varies depending on location and company, but as an example the antibody manufacturer Virusys (Taneytown, United States) provides a kit containing reagents for 40 samples for 1320 USD (as per 2017-08-08), the equivalent of 33 USD per sample. Using the entire antibody-free workflow (flotation, NaClO treatment and filtration) would not cost much more than 3 USD per sample, with the main cost associated to it being the nylon net filter pieces (Merck Millipore, Billerica, United States, prices obtained 2017-08-08) (excluding the costs for the centrifuge tubes that would also be required when preparing the sample for IMS). Hence the cost differs depending on how many filters that are used for each sample (determined by turbidity). If the aim is not to use flow cytometry for a single cell workflow, the flotation alone is sufficient and could be useful in countries or institutions with limited funding.

Upon evaluation, IMS frequently yields a high rate of oocyst recovery, albeit large variations are often seen. In a comparison between two of the more popular IMS kits from Dynal (Lake Success, United States) and Clearwater Diagnostics (Maine, United States), the recovery rate varied between 34% and 100% (Rochelle et al., 1999). Feng and colleagues deduced that depending on the medium that the oocysts are suspended in (tap water, reclaimed water, microfiltrated water and reservoir water), the recovery rate differed between 14% and 95% (Feng et al., 2003). It was postulated that size and concentration of other particles suspended in the water accounted for this discrepancy (Feng et al., 2003). Yakub and Stadterman-Knauer showed that recovery is strongly affected by the concentration of dissolved iron (Yakub & Stadterman-Knauer, 2000). Untreated sewage is a very complex medium, and it is not unreasonable that its constituents has an effect on the recovery efficacy of the magnetic beads. Furthermore, Koompapong et al. found that flotation with sucrose was slightly more efficient at recovering oocysts at various turbidities, compared to IMS (Koompapong et al., 2009). These factors combined with the possibility of antibody bias towards certain species, encourages further comparisons of IMS contra flotation.

5.1.1 Flotation with saturated NaCl-Glucose as opposed to other flotation fluids

Over the years, many different concentration techniques such as the use of various flotation fluids have been evaluated (Kuczynska et al., 1999: Kar et al., 2011: Anderson, 1983: Peeters & Villacorta, 1995). Two of the studies concluded flotation with saturated NaCl to be superior (Kuczynska et al., 1999: Kar et
A majority of the methods were originally developed for larger helminth eggs and cysts and it is an important factor to consider when interpreting the results of such comparisons, as Cryptosporidium oocysts are light (specific gravity estimated to around 1.05) (Current & Garcia, 1991: Kuczynska et al., 1999). Many of them were also developed with bright field or phase-contrast microscopy in mind, for instance when using Sheater’s sugar solution the oocysts appear slightly pink or bright (when viewed through a phase-contrast microscope), as opposed to for instance yeast cells (Current & Garcia, 1991). This type of consideration is not important when staining the oocysts with FITC-labeled mAbs and enumerating through epifluorescence microscopy.

Choice of concentration method should also be based on the properties of the sample material. The lipid content of diarrheal samples is often high and physico-chemical concentration methods such as flotation with diethyl ether has previously been recommended (Peeters & Villacorta, 1995). The faecal samples used in the present study had semi-high fat content, but the requirement to handle diethyl ether in a fume hood, made this method more cumbersome and no apparent advantage was seen when using it on some of the samples.

In the present study, PBS-0.01% Tween 20 was used together with mechanical dispersion through vortexing and vigorous shaking. This process, detaching oocysts from faecal particles, seem in general to be one of the key aspects of a successful flotation. Tween is a widely used non-ionic surfactant and Tween 20 and 80 has both been shown to efficiently prevent protein aggregation even at low concentrations (Chou, 2005: Johnson, 2013). Kuczynska and colleagues speculate that the success of flotation with NaCl in their study was due to the abilities of Na⁺ ions to disperse soil particles (Kuczynska et al., 1999). They also noted a significant difference in recovery when a dispersing agent was used, as opposed to purified water or PBS; the highest recovery was observed when using Triss-0.5% Tween 80 (Kuczynska et al., 1999).

Two attributes that is often desired of a flotation fluid are high specific gravity and low viscosity (Kuczynska et al., 1999). The gravity will determine where the oocysts will be located in the gradient after centrifugation. Due to the addition of purified water (1:1) to the combined saturated NaCl and glucose, the specific gravity of it is fairly low, 1.07 (107g/100 mL). For a slightly different formulation of the flotation fluid commonly used (400 g NaCl and 500g sugar, mixed with 1000 mL water), the specific gravity is 1.28 (The RVC/FAO Guide to Veterinary Diagnostic Parasitology, 2004) and therefore more viscous. Examples of the density of other flotation fluids are 1.21 for Saturated NaCl, 1.18 for cold sucrose and 1.18 for Sheater’s sugar solution (Kuczynska et al., Peeters & Villacorta, 1995). In that regard, Percoll (density centrifugation media) is more similar with a specific gravity of 1.09 (Peeters & Villacorta, 1995). Through the method of discontinuous Percoll gradient centrifugation, a gradient band can be seen with the upper parts of the band having a specific gravity of 1.05 and the lower 1.09 and oocysts (estimated specific gravity: 1.05) are extracted from this section (Peeters & Villacorta, 1995). With the saturated-NaCl-glucose formulation used in the present study, the oocysts can therefore be expected to be found slightly above the flotation fluid after the first centrifugation.

5.1.2 High oocyst recovery with saturated- NaCl-glucose flotation fluid
As no previous data of oocyst recovery through this method was available, further and more systematic assessments of it had to be made to confirm the high recovery (87.8%) achieved. Neither Babcock sedimentation has been evaluated previously. However, data concerning recovery through the use of saturated NaCl does exist. When using bovine faecal samples, Kuczynska and colleagues recovered 17% of oocysts, which is comparable with the 21.6% recovered in the present study when using saturated NaCl (Kuczynska et al., 1999). Kar et al. on the other hand achieved a recovery rate of 71.2% when using this method (Kar et al., 2011).
Factors that could cause discrepancies apart from sample material properties (such as lipid content), when comparing recovery rates, are viability of oocysts (non-viable oocysts are potentially less likely to be recovered) and oocyst concentration (a higher concentration appear to yield higher proportional recovery) (Smith & Grimason, 2003; Kvac et al., 2003).

5.2 Elevated turbidity, a possible explanation for lower recovery rates with large sample volume
When utilizing the saturated-NaCl-glucose flotation method, oocysts are expected to be found in the supernatant above the flotation fluid, that after the initial brief centrifugation also contains heavier debris. If a larger amount (than 1g/1 mL) of sample is used, the author of the present study has noticed that it becomes increasingly difficult to differentiate between the flotation fluid and the supernatant above. It is then necessary to estimate where the border between them is, and together with inability to efficiently separate oocysts from heavier debris, this is thought to account for the loss in oocyst recovery seen when higher volumes are used (Table 3).

5.3 Utilization of sampling capsules would reduce hands-on time required by the method
It was believed that the use of influent wastewater would quickly clog the membrane filter of the Envirocheck sampling capsules (Personal Communication, Folasade Adeyemo, Institute for Water and Wastewater Technology, DUT) and that it was only suitable for the concentration of oocysts in effluent wastewater. A test-run with pre-sieved (20 μm) influent water of fairly high turbidity did however show that at least 2 liters can be run through it without heavy clogging. It is a quick and easy method, albeit expensive (a case of 25 units cost between 2800 and 4400 USD, depending on supplier), compared to concentration through centrifugation.

The filter capsules have been popular within the field for almost two decades, with Matheson and colleagues being one of the first groups demonstrating its efficacy in regards to oocyst recovery (70% recovery rate) (Matheson et al., 1998). Since then it has been frequently used, and is a part of the U.S Environmental Protection Agency (EPA) Method 1623 for Cryptosporidium and Giardia recovery from source waters, since 2005 (EPA, 2005). It has also been used by Yang et al., 2013 (Detection of C. hominis and C. parvum in environmental samples, 44-48% recovery) and Jellison et al., 2009 (Cryptosporidium diversity study in effluent WWTP water). Furthermore, Umgeni Water (Pietermaritzburg, KwaZulu Natal), one of the largest water management boards in South Africa routinely uses Envirocheck sampling capsules in their water quality assessments (Personal Communication, Ian Bailey, Section Head Biological Sciences at Umgeni Water)

While aimed for lower-turbidity source waters such as effluent flows or surface waters, Lee and colleagues showed that even at turbidities as high as 50 NTU (corresponding to the highest acceptable turbidity level of source water among some WWTPs, according to the authors) it was possible to filter up to 12 liters of water (Lee et al., 2004). The mean recovery rate was then 23.6%, as compared with samples having a turbidity of 10 NTU where the mean recovery rate was 38% (Lee et al., 2004). Water parameters, including turbidity, for the material used in the present study was not measured. However, for future samplings, the use of filter cartridges could enable the processing of larger volumes and hence a more reliable result.

5.4 The 46% success rate of the single cell workflow indicates potential for further developments
Last year, Troell and colleagues reported an 80% success rate when sorting from flotated faecal samples (Troell et al., 2016) with estimated 2.5 x 10^6 oocysts per gram. It has been observed in the present study that chlorination and filtration reduces the concentration of oocysts by 80-90%. A low ratio of oocysts to debris in the sample makes it challenging to sort individual oocysts and the purified sample used (see section 3.5.1) was estimated to contain less than 700 oocysts in 4 mL material. In light of that, a 46% success rate is promising, and further improvements of the technique could most likely increase the success
rate. However, it should be emphasized that a comparatively low number of sorted oocysts were sequenced in the present study. In addition, the most difficult step in the workflow is a successful flow cytometry. If this can be improved it is likely that the success rate from oocyst to sequence can be improved as well.

5.5 Amplicon sequencing reveal a species composition that hints towards animal influx

Considered the high prevalence of *C. hominis* within the population as assessed by previous studies (Samra *et al.*, 2013; Samie *et al.*, 2006; Samra *et al.*, 2016), a more prominent role of this species could have been expected. Out of 41,728 crypto reads, only 498 (1.1%) was identified as *C. hominis*. The high proportion of *C. parvum* reads (84%) does however not necessarily mean that there is an animal (such as dairy cattle) influx on the wastewater. If the main source is from humans, new light may be shed on the relative impact of the species. That the individual oocysts sorted at the SiCell platform also were identified as *C. parvum* is consistent with the dominance of this species within the Kingsburgh sample (used at the SiCell platform), based on the proportion of reads (60%).

*Cryptosporidium* sp. mink genotype and deer mouse genotype III have both been found previously in storm waters and are known to reside in proximity to watersheds comprised of this water (Jiang *et al.*, 2005: Jiang *et al.*, 2006: Feng *et al.*, 2007). Apart from these findings in the United States, deer mouse genotype III has also recently been detected in Canada (Thomas *et al.*, 2016) but no human infection is known to date. In contrast, *Cryptosporidium* sp. mink genotype, first discovered in China (Wang *et al.*, 2008), has been isolated from patients in Australia (Ebner *et al.*, 2015: Ng-Hublin *et al.*, 2016). Its presence in the wastewater of Kingsburgh WWTP therefore has zoonotic potential. Deer mouse genotype III was found in both Kingsburgh and Mariannridge WWTP and as human hosts are not known, it indicates a moderate animal influence on the source water.

Based on the published literature, it appears to be the first time *Cryptosporidium* sp. deer mouse genotype III and mink genotype are found on the African continent.

5.6 Conclusion

The novel antibody-free workflow, designed to deal with potential bias introduced by antibody affinity, was optimized in the present study to suit large volumes of wastewater. A flotation fluid consisting of saturated NaCl and glucose was found to yield the highest rates of recovery, when compared with two other flotation fluids. This adapted method of oocyst concentration was evaluated for the first time on material from four different WWTPs. Compared to the standard method of IMS, the antibody free workflow performed well, resulting in comparatively high concentrations of oocysts. It is also considerably cheaper, albeit slightly more time consuming. To dissect the species composition of these samples a single cell (SiCell) workflow and amplicon sequencing on the miSeq platform (Illumina) was utilized. With the former, a small subset of single oocysts from Kingsburgh WWTP were sorted and sequenced individually with Sanger sequencing technology. The success rate was 46% and they were all determined to be *C. parvum*. With amplicon sequencing, four different species/genotypes were found, with a majority of reads belonging to *C. parvum*, followed by *C. hominis*. Additionally, *Cryptosporidium* sp. deer mouse genotype and mink genotype were found, possibly for the first time on the African continent.
Acknowledgements

I would like to extend my feelings of gratitude to my supervisor Karin Troell, for always, regardless of the circumstances having an optimistic input on hurdles reached during the project and picking apart the different obstacles with a contagious enthusiasm. She was always available for questions and valuable feedback and made the whole project a very exciting one.

I am also very grateful for having had the opportunity to perform this project at SVA with guidance, support and helpful insights from Harri Ahola, Anna-Maria Divne (at SciLife lab), Tobias Lilja, Cecilia Alsmark and many others, including the friendly staff at the parasitology department.

The project would have been impossible without the help and logistics provided by my co-supervisor, the very charismatic Thor-Axel Stenström at DUT. I received a very warm welcome by him and his research group. Thanks to them I enjoyed an amazing and indeed very interesting time at their facilities and the city of Durban.

An especially sincere thank you to Gulshan Singh and Folasade Adeyemo for all their support and friendly company and to the entire group at the Institute for water and wastewater technology, DUT.
References


Merck Millipore (2016) Montage SEQ96 Sequencing Reaction Cleanup Kit User Guide


Mhlanga, FT. Brouckaert, CJ. Foxon, KM. Fennemore, C. Mzulwini, D. Buckley, CA (2009) Simulation of a wastewater treatment plant receiving industrial effluents. Water SA. 35 (4), ISSN 1816-7950


Samra, NA. Jori, F. Caccio, SM. Frean, J. Poonsamy, B. Thompson, PN (2016) Cryptosporidium genotypes in children and calves living at the wildlife or livestock interface of the Kruger National Park, South Africa. *Onderstepoort j. vet. res.* 83 (1), ISSN 2219-0635


Appendix A

A1. Calculating the number of viewing fields per 12 mm well at a Teflon Printed Diagnostic Slide, Article number 61.100.03 (Immuno-Cell, Mechelen, Germany)

Considering that each well is circular, the amount of viewing fields per well can be calculated with the formula:

\[
\left( \frac{\text{viewing fields across the entire well}}{2} \right)^2 \times \pi
\]

When enumerating oocysts from slides it is sometimes observed that they can be found on the edge or slightly outside of the well (on top of the Teflon). Depending on if these are counted or not, the amount of viewing fields screened per well naturally differs.

- Viewing fields across the well if the edge is also included, at 250x magnification: \(16.5\)
  - Viewing fields per well is hence: \(\left( \frac{16.5}{2} \right)^2 \times \pi = 213.8\)
- Viewing fields across the well if the edge is also included, at 250x magnification: \(15.0\)
  - Viewing fields per well is therefore: \(\left( \frac{15.0}{2} \right)^2 \times \pi = 176.7\)

A2. Evaluation of oocyst recovery in three different flotation methods

Bovine fecal sample highly positive for Cryptosporidium, obtained from the diagnostic department at SVA, was used. It was referred to as #1275.

Enumeration of pre-floatated sample was done by diluting it 1:10 with PBS and adding 10 μL to a Teflon printed diagnostic slide, as described in section 3.1. A total of 30 viewing fields were counted, distributed over 3 wells.

Well I: 20 oocysts/viewing field (based on average of 10 fields of view)
Well II: 32.1 oocysts/viewing field (based on average of 10 fields of view)
Well III: 27.4 oocysts/viewing field (based on average of 10 fields of view)

Average amount of oocysts/viewing field: 26.5

- To compensate for the 1:10 dilution: 26.5 oocysts x 10
- Amount of oocysts in the entire well: 265 oocysts x 176.7 viewing fields (in the 12 mm circular well)
- Since 10 μL had been added to the slide: 46825 oocysts/10 μL

\[4682.5 \text{ oocysts/μL in the original sample material} = 4.68 \times 10^6/\text{mL}\]

Enumeration of Babcock sedimentation:

31.8 oocysts/viewing field (based on average of 10 fields of view)

- To compensate for the 1:10 dilution: 31.8 oocysts x 10
- Amount of oocysts in the entire well: 318 oocysts x 176.7 viewing fields (in the 12 mm circular well)
- Since 10 μL had been added to the slide: 56190 oocysts/10 μL

5619 oocysts/μL x 2000 μL material after purification (the volume of material obtained is 2 mL):

\[11.2 \times 10^6/\text{mL}\]

Enumeration of saturated NaCl flotation:

Well I: 22.2 oocysts/viewing field (based on average of 10 fields of view)
Well II: 16.8 oocysts/viewing field (based on average of 10 fields of view)
Well III: 45.2 oocysts/viewing field (based on average of 10 fields of view)
Well IV: 30.8 oocysts/viewing field (based on average of 10 fields of view)
Average amount of oocysts/viewing field: 28.75
  - Amount of oocysts in the entire well: 28.75 oocysts x 176.7 viewing fields (in the 12 mm circular well)
  - Since 10 µL had been added to the slide: 5080 oocysts/10 µL

508 oocysts/µL x 2000 µL material after purification (the volume of material obtained is 2 mL):

1.01 x 10⁶/mL

Enumeration of saturated NaCl-Glucose flotation:
Well I: 120.4 oocysts/viewing field (based on average of 10 fields of view)
Well II: 112.4 oocysts/viewing field (based on average of 10 fields of view)
Average amount of oocysts/viewing field: 116.4
  - Amount of oocysts in the entire well: 116.4 oocysts x 176.7 viewing fields (in the 12 mm circular well)
  - Since 10 µL had been added to the slide: 20567.9 oocysts/10 µL

2056.8 oocysts/µL x 2000 µL material after purification (the volume of material obtained is 2 mL):

4.11 x 10⁶/mL

A3. Evaluation of oocyst recovery in adapted protocol
The same sample as in previous evaluations (#1275) was also used here with a concentration of 4682.5 oocysts/µL. (See section A2). Three different amounts of starting material, 4 mL, 8 mL and 12 mL were used and oocyst recovery calculated. The flotated material was subjected to a tenfold dilution before 10 µL of it was streaked on a Teflon printed diagnostic slide, according to aforementioned method.

Enumeration of adapted protocol with 4 mL starting material:
17.6 oocysts/viewing field (based on average of 10 fields of view)
  - To compensate for the 1:10 dilution: 17.6 oocysts x 10
  - Amount of oocysts in the entire well: 176 oocysts x 176.7 viewing fields (in the 12 mm circular well)
  - Since 10 µL had been added to the slide: 31099 oocysts/10 µL

3109 oocysts/µL x 5000 µL material after purification (the volume of material obtained is 5 mL):

15.55 x 10⁶/mL

Enumeration of adapted protocol with 8 mL starting material:
31.0 oocysts/viewing field (based on average of 10 fields of view)
  - To compensate for the 1:10 dilution: 31.9 oocysts x 10
  - Amount of oocysts in the entire well: 319 oocysts x 176.7 viewing fields (in the 12 mm circular well)
  - Since 10 µL had been added to the slide: 56367 oocysts/10 µL

5636 oocysts/µL x 5000 µL material after purification (the volume of material obtained is 5 mL):
28.18 x 10^6/mL

Enumeration of adapted protocol with 12 mL starting material:

39.9 oocysts/viewing field (based on average of 10 fields of view)

- To compensate for the 1:10 dilution: 39.9 oocysts x 10
- Amount of oocysts in the entire well: 399 oocysts x 176.7 viewing fields (in the 12 mm circular well)
- Since 10 μL had been added to the slide: 70503 oocysts/10 μL

7050 oocysts/μL x 5000 μL material after purification (the volume of material obtained is 5 mL):

35.25 x 10^6/mL

A4. Enumeration of oocysts in South African wastewater samples

Due to lack of time and limited availability to epifluorescence microscopy, only a few samples could be enumerated. The entire well of each prepped sample (10 μL) was scanned through epifluorescence microscopy, first at 100x and then at 400x for confirmation. The software Axiovision v.4.7.1 (Carl Zeiss, Oberkochen, Germany)

**Isipingo influent [flotated], combined material from Isi.Inf.1 and Isi.Inf.2**

Well I: 9 oocysts in total
Well II: 6 oocysts in total
Well III: 10 oocysts in total

- 25 oocysts/30 μL (3 slides): 0.83 oocysts/μL
- The total volume of flotated sample from Isipingo influent is 19 000 μL (well mixed and homogenous) x 0.83 oocysts/μL
- 15833 oocysts stemming from 5 liters of influent wastewater
- **3166 oocysts/liter**

**Isipingo influent [treated with NaClO], combined material from Isi.Inf.1 and Isi.Inf.2**

Well I: 3 oocysts in total
Well II: 2 oocysts in total

- 5 oocysts/20 μL: 0.25 oocysts/μL
- The total volume of chlorinated sample from Isipingo influent, is 20 000 μL (well mixed and homogenous) x 0.25 oocysts/μL
- 5000 oocysts stemming from 5 liters of influent wastewater
- 1000 oocysts/liter

**Kingsburg sludge [flotated], combined material from King. Slu. 1 and King. Slu. 2**

Well I: 21 oocysts in total

- 21 oocysts/10 μL: 2.1 oocysts/μL
- The total volume of flotated sample from Kingsburgh sludge is 40 000 μL (well mixed and homogenous) x 2.1 oocysts/μL
- 84 000 oocysts stemming from 1 liters of sludge
- **84 000 oocysts/liter sludge**

**Kingsburg influent [flotated], material from King. Inf. 1**

Well I: 66 oocysts in total

- 66 oocysts/10 μL: 6.6 oocysts/μL
• The total volume of flotated sample from Kingsburgh influent 1 is 14 500 μL (well mixed and homogenous) x 6.6 oocysts/μL.
• 95700 oocysts stemming from 2.5 liters of influent
• **38 280 oocysts/liter sludge**

**Kingsburg influent [flotated], material from King. Inf. 2**

Well I: 26 oocysts in total

• 26 oocysts/10 μL: 2.6 oocysts/μL.
• The total volume of flotated sample from Kingsburgh influent 2 is 15 000 μL (well mixed and homogenous) x 2.6 oocysts/μL.
• 39 000 oocysts stemming from 2.5 liters of influent
• **15 600 oocysts/liter sludge**

**Kingsburg sludge [flot+NaClO+filt], material from King. Slu.4 (Enumerated in Sweden at 250x, 2017-03-28)**

Well I: 14 oocysts in total

• 14 oocysts/10 μL: 1.4 oocysts/μL.
• The total volume of available material from Kingsburgh sludge 4 is 2000 μL.
• 2000 μL x 1.4 oocysts/μL: 2800 oocysts in total
• 2800 oocysts stemming from 0.25 liters of sludge
• **11 200 oocysts/liter sludge**

**Isipingo sludge [flot+NaClO+filt], material from Isi. Slu.1 (Enumerated in Sweden at 250x, 2017-03-31)**

Well I: 11 oocysts in total

• 11 oocysts/10 μL: 1.1 oocysts/μL.
• The total volume of available material from Isipingo sludge 1 is 2000 μL.
• 2000 μL x 1.1 oocysts/μL: 2200 oocysts in total
• 2200 oocysts stemming from 0.25 liters of sludge
• **8800 oocysts/liter sludge**

**Marianridge influent [flot+NaClO+filt], material from Mar. Inf (Enumerated in Sweden at 250x, 2017-03-31)**

Well I: 3 oocysts in total

• 3 oocysts/10 μL: 0.3 oocysts/μL.
• The total volume of available material from Marianridge influent is 2000 μL.
• 2000 μL x 0.3 oocysts/μL: 600 oocysts in total
• 600 oocysts stemming from 2.5 liters of influent
• **240 oocysts/liter influent**

**Shallcross influent [flot+NaClO+filt], material from Shall. Inf (Enumerated in Sweden at 250x, 2017-03-31)**

Well I: 5 oocysts in total

• 5 oocysts/10 μL: 0.5 oocysts/μL.
• The total volume of available material from Marianridge influent is 2000 μL.
• 2000 μL x 0.5 oocysts/μL: 1000 oocysts in total
• 1000 oocysts stemming from 2.5 liters of influent
• **400 oocysts/liter influent**
Appendix B

Appendix B. Figure 1. Flotated faecal sample treated with 1% NaClO solution and streaked onto agar. A. Blue agar, incubated at 37 °C. Bacterial colony is marked with a red circle. B. Bovine blood agar, incubated at 37 °C. Bacterial colony is marked with a blue circle.

Appendix B. Figure 2. Flotated faecal that had not been disinfected with NaClO, streaked onto agar. A. Bovine blood agar incubated at 37 °C. B. Blue agar incubated at 20 °C. C. Blue agar incubated at 37 °C. D. Bovine blood agar incubated at 20 °C.