Quality Assurance Project Plan for

Cryptosporidium Source Tracking to Enhance Source Water Protection Implementation in the Potomac River Watershed

Lihua Xiao, Principal Investigator
Division of Parasitic Diseases, Centers for Disease Control and Prevention, Atlanta, GA 30341

Plato P.T. Chen, Co-Principal Investigator
Washington Suburban Sanitary Commission, Laurel, MD 20707

_____
Lihua Xiao
Principal Investigator
Centers for Disease Control and Prevention

_____
Plato P.T. Chen
Co-Principal Investigator
Washington Suburban Sanitary Commission

_____
Eric N. Villegas
EPA Project Officer
National Exposure Research Laboratory, U.S. Environmental Protection Agency

_____
Margie Bassett
Quality Assurance Manager
National Exposure Research Laboratory, U.S. Environmental Protection Agency

Table of Contents

Distribution List 3
Project Organization 3
Problem Definition/ Background 4
Project Quality Objectives 6
Experimental Design/Techniques 8
Quality Control 12
Peer Review/Audits 15
Data Management 15
Work Plan, Deliverables and Timeline 16
References 17
Appendix 18

Distribution List

Lihua Xiao, CDC
Plato Chen, WSSC
Wenli Yang, CDC
Ronald Landy, EPA
Charles Kanetsky, EPA
Eric Villegas, EPA
Margie Bassett, EPA

Project Organization

The project is a joint effort between the Potomac River Basin Drinking Water Source Protection Partnership (DWSPP), CDC, and EPA, with most of the sampling and microscopic analysis done through the DWSPP utilities and most of the molecular analyses done at CDC. EPA provides most of the funds needed for the project, with additional in-kind support from DWSPP and CDC. EPA personnel provide supervision and oversight over the project. The following personnel are associated with the project:

Lihua Xiao is the principal investigator of the project, supervises all aspects of the molecular analysis, and performs the data analyses.

Plato Chen is the co-principal investigator of the project, responsible for the initiation of the project and coordination among participating organizations, and organizes and supervises the sampling and microscopic testing of water samples.

Eric Villegas is the EPA Office of Research and Development Project Officer, and supervises the execution and performance of the project and the interagency agreement between EPA and CDC.

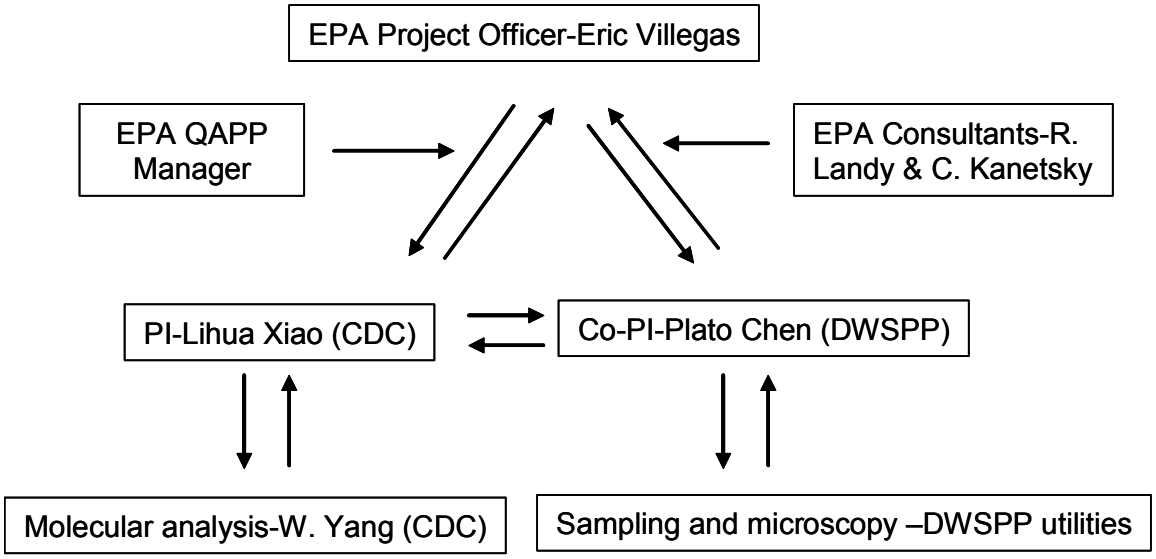
Wenli Yang is a postdoctoral associate under the supervision of Lihua Xiao, and performs all molecular analyses.

Ronald Landy serves as the on-site EPA Region III supervisor for the project.

Charles Kanetsky serves as an EPA expert advisor on surface and drinking water issues.

Margie Bassett is the QA Manager for the National Exposure Research Laboratory in Cincinnati, OH.

The project organization is shown in the diagram below:



Problem Definition/ Background

The Potomac River is a water supply critical to many communities in the Mid-Atlantic. The population of the Potomac River Basin is approximately 5.35 million (2000 Census) and is growing rapidly. The Washington Metropolitan Area has approximately 3.7 million residents or almost three-quarters of the basin's population, and the non-tidal Potomac River is the main water supply for these people. The non-tidal Potomac watershed is an inter-jurisdictional, multi-state watershed encompassing over 11,500 square miles (including areas of Maryland, Virginia, Pennsylvania, and West Virginia) with thousands of potential sources of contamination. In April 2004, the Potomac River Basin Drinking Water Source Protection Partnership (DWSPP) was formalized and mobilized to help coordinate and initiate efforts by the local drinking water utilities and government agencies to protect this highly valued watershed for drinking water purposes. The DWSPP members currently include the three major water utilities serving the Washington DC metropolitan area, six other water utilities in the watershed, and nine federal and state government agencies (including the U.S. EPA Region III).

One of the common contaminants of concern that has been identified in all of the Source Water Assessments for the various Potomac River water treatment plants is *Cryptosporidium*. *Cryptosporidium* spp. are protozoan parasites that are excreted by infected animals and humans and have been found in most drinking water sources. The most significant human-infectious species are *Cryptosporidium parvum* and *Cryptosporidium hominis*. The former infects both humans and ruminants, whereas the latter is almost exclusively a human pathogen. Many other *Cryptosporidium* species and genotypes are also present in animals and frequently also contaminate surface water, causing problems in assessing the human infection potential of *Cryptosporidium* oocysts (the infective stage) found in water. *Cryptosporidium* oocysts are resistant to disinfection by chlorination, and cause significant gastrointestinal illness in infected individuals. They have been the cause of numerous outbreaks of diarrheal illness associated with drinking and recreational water. Therefore, the recently implemented Long Term 2 Enhanced Drinking Water Treatment Rule (LT2ESWTR) by EPA requires the regular monitoring of *Cryptosporidium* oocysts in source water (USEPA 2006).

Currently, only limited data of varying degrees of quality are available on the occurrence of *Cryptosporidium* oocysts in Potomac watershed and there are no specific data indicating the likely source of contamination and public health significance of oocysts found in water. Over the past decade, a total of 239 samples were collected from five water utilities on the Potomac. The mean values from all of these data for upstream Potomac and Washington DC metropolitan area locations were 0.013 and 0.021 oocysts/liter, respectively, with maximum levels of 0.2 and 1.0 oocysts/liter, respectively. However, these data represent results of a number of different generations of methods for measuring *Cryptosporidium*, including an ASTM method, the ICR method, and EPA method 1622/1623, and all of these methods have a significant deficiency with respect to recovery. Furthermore, the sampling for these data was not intentionally matched with different hydrologic conditions. Results of a study conducted by the Maryland Department of the Environment (MDE) and Johns Hopkins University with samples from the Potomac Basin between 1998 and 2002 suggest that the levels of *Cryptosporidium* may be significantly higher than those previously observed in sampling by the local utilities.

The greatest problem with most of the existing data is that they do not identify human infection potential and the specific sources of *Cryptosporidium* in water; they simply enumerate the quantity of oocysts in the raw water. A method of *Cryptosporidium* "source-tracking" needs to be applied in conjunction with these techniques to enable more specific identification of sources (*e.g.*, human, bovine, swine, avian, rodent, deer, and various wild animals) so that source protection efforts can be focused and effective. The suspected sources of *Cryptosporidium* oocysts in the Potomac River watershed include agricultural activities/animal operations, combined sewer overflows or wastewater treatment discharges, wild animals, and storm water runoff.

Recently, molecular diagnostic tools have been developed to assess the human infection potential of *Cryptosporidium* oocysts in water and to track the sources of contamination. One of the most widely used molecular detection tools is a nested PCR-based genotyping method developed at the Centers for Disease Control and Prevention (CDC). This method has been used by many research groups in differentiating *Cryptosporidium* in humans and animals, and has been used effectively in studies in assessing the public health significance and source of *Cryptosporidium* contamination in watersheds in New York and Milwaukee in the United States and Alberta and Ontario in Canada (Xiao et al. 2000; Xiao et al. 2001; Zhou et al. 2003; Jiang et al. 2005b; Ruecker et al. 2005; Ruecker et al. 2006). Because different species/groups of animals are infected with different host-adapted *Cryptosporidium* species or genotypes, the finding of a particular species or genotype in water indicates a likely source of animal contamination for oocysts detected. It is also very sensitive compared to the conventional EPA Method 1622/1623 (Xiao et al. 2000; Jiang et al. 2005a; Jiang et al. 2005b).

Recently, the DWSPP water utility members and EPA Region III have proposed to work with the EPA Office of Research and Development (ORD) to develop and implement a monitoring program in the Potomac River Watershed for *Cryptosporidium* source tracking in order to identify the most significant sources of *Cryptosporidium* oocysts within the watershed. Once the significant sources are known, appropriate source protection efforts can be made to eliminate or minimize the contributions of these sources. The Partnership proposes to take advantage of the recent developments in *Cryptosporidium* source tracking to identify the likely source of *Cryptosporidium* oocyst contamination in the Potomac watershed. If successful, this approach could be applied to other major watersheds to generate data needed for source water protection and watershed management.

Project Quality Objectives

The objective of this project is to track the sources of *Cryptosporidium* contamination in the Potomac River Basin using genotyping and subtyping tools. This project will involve a collaborative effort between DWSPP, CDC, and EPA scientists. Because PCR methods will be used in the detection of *Cryptosporidium* oocysts in water, the QA/QC procedures outlined in the recent EPA publication (815-B-04-001) "Quality Assurance/Quality Control Guidance for Laboratories Performing PCR Analyses on Environmental Samples" will be followed. Sampling, microscopy detection and enumeration of *Cryptosporidium* oocysts will use the standard EPA method 1622 (USEPA 2001), which has established QA/QC procedures in the standard operational procedures (SOP). The QC contained within this document only pertains to the data obtained by PCR detection of *Cryptosporidium* oocysts. The objective of quality control in this project is to provide scientific data that will be repeatable.

a) Basic QA Objective:

The important issues for QA in this project are reduction of specimen contamination, poor oocysts recovery in Method 1622, accurate PCR amplification and sequencing, and correct interpretation and reporting of PCR results and sequence data. Although PCR and sequencing errors are easy to detect, they nevertheless prevent the acquisition of accurate PCR products and nucleotide sequences, which are the bases to molecular analysis. Thus, the major effort of QA in this project will be the elimination of sample and PCR contamination, poor PCR amplification, and sequencing errors. Because the standardized EPA Method 1622 will be used in most of the sample processing and microscopic detection steps in EPA certified laboratories, the QA work of the project concentrates on the sampling and molecular detection steps.

b) Type and Quality of Data Collected:

The project involves the collection of water samples by participating DWSPP water utilities, filtration of the water samples using procedures specified in Method 1622, isolation and detection of *Cryptosporidium* oocysts by microscopy in EPA certified laboratories using Method 1622, and tracking the source of *Cryptosporidium* contamination using molecular methods. The type of data generated by procedures and instruments used in the molecular work will be mostly qualitative data. The sensitivity of the method is the ability to detect low levels of *Cryptosporidium* oocysts, whereas the specificity is the ability to detect only *Cryptosporidium* oocysts or DNA in the presence of other organisms. The precision of detection is whether the method correctly detects the presence or absence of *Cryptosporidium* oocysts, i.e., whether there is a good agreement between results of microscopy and PCR.

Below are types of data collected, the required quality, and evaluation methods:

Oocyst levels. Results of microscopic detection of *Cryptosporidium* will be expressed as the number of oocysts per liter of water. The accuracy of results will be evaluated by the certified testing laboratories using matrix spikes and method controls as part of the required routine QA/QC practices implemented in Method 1622.

PCR positivity. The accuracy goal for PCR is to detect *Cryptosporidium* oocysts with high sensitivity and specificity. For sensitivity, the ultimate goal is to detect a single oocyst in

water in the presence of interference materials, and false negative is not acceptable. Samples seeded with known number of oocysts will be used to estimate the sensitivity of the test. For specificity, the goal is to detect only *Cryptosporidium* DNA. Thus, false positive is also not acceptable. Because only qualitative data will be generated, data from all PCR runs will be evaluated for the presence of false positive and false negative result, using positive and negative controls in each PCR run. In addition, all PCR products will be sequenced to confirm results of the PCR-RFLP diagnosis. Only data from PCR runs without false positive and false negative results will be acceptable.

RFLP patterns. Restriction fragment length polymorphism (RFLP) analysis of PCR products will be used to determine the *Cryptosporidium* species/genotypes present in water samples and the possible occurrence of mixed species or genotypes. The accuracy of RFLP analysis will be assessed with the inclusion of the RFLP analysis of positive control with each sample run. Because the results are interpreted against known profiles for various *Cryptosporidium* species and genotypes, any deviation in RFLP patterns will be obvious to trained personnel. All positive PCR products will be sequenced to verify the diagnosis by RFLP analysis.

DNA sequences. The accuracy goal for DNA sequencing is to generate DNA sequences free of errors. All PCR products from each sample will be sequenced in both the 5' and 3' directions, and the sequences obtained will be compared with each other and those downloaded from the public sequence database GenBank. Any sequence differences will be verified through manual inspection of the electropherogram (track file from the automatic DNA sequencer), and if needed by sequencing a second PCR product from the same sample or DNA.

c) Error Sources or Roadblocks:

The major error sources and roadblocks in the effort to track the source of *Cryptosporidium* oocysts are PCR contamination, poor PCR amplification, and PCR inhibition. The former results in false positive and is common to all PCR techniques, especially nested PCR techniques. The latter two are more common to the analysis of environmental samples. Poor PCR amplification is usually a result of un-optimized PCR primers and conditions. There are many standardized approaches to address PCR contaminations and poor PCR amplifications, which will be used to minimize the occurrence of the two problems in this project (see the Quality Control section below). Thus, we do not expect them to be major roadblocks in this project. What is more problematic and common to all PCR methods for the analysis of environmental samples is the inherent richness of PCR inhibitors in water samples. The co-extraction of PCR inhibitors during DNA isolation can greatly reduce the sensitivity of PCR detection, leading to false negative results. In this project, we have chosen to purify oocysts using immunomagnetic separation established in Method 1622. Because only purified oocysts will be used in DNA extraction, it is expected DNA of high purity will be produced. We will also use a high concentration of non-acetylated bovine serum albumin in PCR to neutralize residual inhibitors. This approach has been used by us in effective PCR detection of *Cryptosporidium* oocysts in various types of water (Xiao et al. 2000; Xiao et al. 2001; Jiang et al. 2005a; Jiang et al. 2005b).

Another potential roadblock for this project is the potentially low percentage of *Cryptosporidium* detections due to the limitations of the existing 1622 method and the relatively low concentration of oocysts in natural environments. Existing data from two of the major

utilities in the watershed reveal a relatively low percentage of detections by the 1622 method. To address this roadblock, the sampling program will utilize monthly (at a minimum) baseline sampling as well as targeted storm sampling (minimum of 6 storm events) to provide an adequate dataset on which a statistical analysis can be performed and significant quantitative and qualitative trends can be identified. Furthermore, the use of the nested-PCR method is likely to result in more positive detections than the 1622 method because PCR is much more sensitive than the 1622 method. A second year of monitoring is also planned to fortify the database and the statistical significance of the study.

Experimental Design/Techniques

The Partnership is proposing to conduct a *Cryptosporidium* source tracking study using the genotyping method developed at CDC. During the pilot study period, raw water 20-L samples will be collected monthly from two water treatment plant intake sites and three sub-watershed locations by Partnership members during base-flow. Six water samples will also be collected from each site during the year after storm events. The sampling will follow the Method 1622 and the LT2ESWTR specifications. The monitoring is intended to reflect seasonal and hydrologic variations in *Cryptosporidium* occurrence and source contributions. Each water sample will consist of two 10-L aliquots, and each aliquot will be filtered through Envirochek filters. Where possible, the entire 10-L aliquot will be filtered through one Envirochek filter; for high turbidity waters where this is not possible, either 10-L will be filtered through two Envirochek filters or as much volume as possible will be filtered through two filters if the filters clog before filtering of 10-L. The filter(s) from one aliquot will be processed for *Cryptosporidium* oocyst detection and enumeration by EPA Method 1622 by certified commercial laboratories contracted by the Partnership members (USEPA 2001). The filter(s) from the other aliquot will be processed by CDC for *Cryptosporidium* detection and genotyping by molecular diagnostic methods. Filtration will be performed either in the field (with subsequent shipping of the filters to the respective labs) or at the respective laboratories (after receipt of the 10-L bulk sample).

Upon the arrival of filters in the laboratory, CDC researchers will elute the concentrated particles off the filters and isolate *Cryptosporidium* oocysts from water concentrates by immunomagnetic separation, following procedures specified in Method 1622 (USEPA 2001). DNA will then be extracted from the IMS-oocysts mixture using the Qiagen tissue DNA extraction kits as previously described (Xiao et al. 2004) (See Appendix 2). *Cryptosporidium* oocysts presented will be genotyped using an established genotyping method (Xiao et al. 2004). This method amplifies a fragment of the small subunit ribosomal RNA (SSU rRNA) gene of ~830 bp by nested PCR, and differentiates specific *Cryptosporidium* species or genotypes based on results of restriction fragment length polymorphism (RFLP) analysis and DNA sequence of the amplified PCR products (Appendix 2 and 3). The likely source for *Cryptosporidium* oocysts in water is determined by the identity of *Cryptosporidium* species or genotypes found. Because of the existence of two types of *C. parvum* (anthroponotic subtype family IIc seen only in humans and zoonotic subtype family IIa seen mostly in calves in the US), all *C. parvum* positive samples will be further subtyped by PCR-sequencing of the 60 kDa glycoprotein (GP60) gene (Zhou et al. 2003) (Appendix 4).

At the end of the project, results of analyses with the conventional EPA Method 1622 and genotyping will be compared. The likely source of *Cryptosporidium* oocyst contamination in the

Potomac River watershed will be assessed based on results of genotyping and subtyping. Results of the study will be summarized and the final project report will be shared among Partnership members, EPA personnel, and CDC investigators. Differences between PCR and microscopy in detection rates will be compared by Fisher's exact t test. Mean oocysts level among sites will be compared by Student's t test.

Site Selection and Sample Collection

The five sites selected for monitoring include two water treatment plant intakes and three watershed sites (see Figure 1). The two water treatment plant intake sites are Fairfax Water's (FW) Corbalis Water Treatment Plant (WTP) in Herndon, VA and Washington Suburban Sanitary Commission's (WSSC) Potomac Water Filtration Plant (WFP) in Potomac, MD. These two plants were selected to reflect the characteristics of the source waters on the southern (Virginia) and northern (Maryland) sides of the River. These two intakes are located about 4 miles apart from one another on opposite sides of the mainstem Potomac River. These two plants represent the primary supply sources for two of the three major water utilities in the Washington D.C. metro region. The source water at the Potomac WFP is considered comparable to the source water for the third major water utility (the Washington Aqueduct), whose intake is also on the northern side of the Potomac River. Storm and baseflow samples at the two plant intakes will be collected by FW and WSSC staff.

The three watershed sampling sites were selected to represent contributions of agricultural (cattle) and urban (wastewater treatment) sources to the overall Cryptosporidium load. Two of the watershed sites are in Maryland upstream of the Potomac WFP while the third site is in Virginia upstream of the Corbalis WTP. The Virginia site and one of the Maryland sites represent areas influenced by agricultural sources, and the other Maryland site represents an area influenced by urban sources. The three watershed sites are as close as practical to the major impacting landuses. For the cattle influenced areas, samples immediately downstream of individual operations are not practical due to access issues, but the selected locations are draining the lands of a significant number of cattle operations, so the impact of these cattle operations is expected to be easily observed in the data. The WWTP influenced site is about 200 feet downstream of the discharge point (the effluent is considered well-mixed with the stream at this point), and it is preferable to sample from the stream rather than directly from the WWTP effluent so that upstream influences can also be captured.

The Virginia site is located in the Shenandoah subwatershed on the North Fork Shenandoah River just northeast of Edinburg, VA in Shenandoah County. This location (also known as Chapman's Landing bridge) is downstream of significant dairy and beef feedlot operations, and there are no major WWTP (>1 MGD) discharges upstream. Storm samples at this site will be collected by an EPA contractor, and baseflow samples will be collected by Virginia Department of Environmental Quality staff.

The Maryland site is located in the Monocacy subwatershed on the Monocacy River in northeastern Frederick, MD in Frederick County. This location is also downstream of significant dairy operations, but also has several major WWTP discharges upstream of it (it is estimated that 25-30% of streamflow under low flow conditions is WWTP discharges) and therefore is

somewhat of a "mixed" land use site. Storm and baseflow samples at this site will be collected by Frederick County, MD and Frederick City, MD staff.

The Maryland urban source site will be on Great Seneca Creek just downstream of the discharge from WSSC's Seneca Waste Water Treatment Plant (WWTP). This stream joins the Seneca Creek tributary whose confluence with the mainstem Potomac River is located on the opposite side of the River from FW's Corbalis WTP intake structures. Seneca Creek is not believed to influence Corbalis' source water quality because the River is too wide at this point, but it does influence the source water quality of WSSC's Potomac WFP because it is on the same side of the River and upstream. The Great Seneca Creek sample location is a more urbanized watershed location that is targeted to assess the human contributions of *Cryptosporidium* due to WWTP discharges (specifically from the Seneca WWTP, which is a tertiary WWTP with UV disinfection). The WWTP discharge comprises a significant portion of the streamflow at this point during dry weather flows (perhaps as much as 50%). Storm and baseflow samples at this site will be collected by WSSC staff.

All sample collection shall be performed in accordance with EPA Method 1622 and will follow the procedures of Appendix G (Envirochek Field Filtration for *Cryptosporidium*) or Appendix I (Collecting Bulk Water Samples for Laboratory Filtration and *Cryptosporidium* Analysis) of the "Source Water Monitoring Guidance Manual for Public Water Systems for the Long Term 2 Enhanced Surface Water Treatment Rule", Feb 2006, EPA. It is anticipated that the samples collected at the Virginia watershed location will be bulk water samples that will be shipped to the *Cryptosporidium* laboratories for filtration and analysis; samples from all other locations are expected to be field filtered followed by shipping of filters to the laboratories for analysis. For QC purposes, one matrix spike sample from each sample location will be collected for use by CDC and the commercial labs near the beginning of the monitoring period. Samples from the watershed locations will be taken from points representative of the bulk flow of the stream to the degree possible. Storm events to be sampled will be selected on the basis of the following rules:

1) at least one storm sample should be collected from each location in each of the four seasons over the course of the study period;
2) a storm event must be expected to produce at least 0.5" of rain in the local area and be preceded by at least 4 days of no significant precipitation in order to qualify for sampling;
3) only storm event samples meeting the following criteria will be analyzed by the *Cryptosporidium* laboratories:
    a) a single grab sample from the rising limb of the hydrograph (as verified by a nearby flow gaging station) after at least 0.25" of rain has fallen; or
    b) a time-weighted composite sample from grab samples collected during the rising, peak, and falling limbs of the hydrograph for a storm of at least 0.5" of rain in the local vicinity.

The DWSPP will utilize local meteorological forecasts (preferably with the assistance of a NOAA) to identify the potential storm events for sampling in each of the locations. Because local storm conditions may vary significantly from location to location, it is anticipated that the dates and times of storm samples from the various sites will not necessarily coincide. The co-PI will direct the collection of storm samples for each sample location; advance notice for sampling of 8 hours (and preferably 24 hours) will be given whenever possible.

Quantitative analysis of the above samples by EPA Method 1622 will be performed by several certified contract labs. The samples collected at the WSSC's Potomac WTP and the Great Seneca Creek sites will be analyzed by WSSC's contract lab (Environmental Associates, Ltd.), the samples collected at FW's Corbalis WTP and the North Fork Shenandoah sites will be analyzed by FW's contract lab (Analytical Services, Inc.), and the samples collected at the Monocacy site will be analyzed by either WSSC's contract lab or WAD's contract lab (Environmental Associates, Ltd.).

Oocyst isolation

Cryptosporidium oocysts present in water samples will be isolated by immunomagnetic separation, using anti-*Cryptosporidium* Dynabeads from the Dynabeads Anti-*Cryptosporidium* kit (Invitrogen, Carlsbad, CA) and procedures implemented in EPA Method 1622.

DNA extraction

DNA will be extracted from water samples using Qiagen tissue DNA extraction kits as previously described (Xiao et al. 2004). This technique was shown earlier by us to effectively remove PCR inhibitors (Xiao et al. 2000), and has been used effectively in PCR detection of *Cryptosporidium* oocysts in various water samples (Xiao et al. 2000; Xiao et al. 2001; Jiang et al. 2005a; Jiang et al. 2005b). The standardized operational procedure (SOP) is attached.

PCR

A nested PCR based on the SSU rRNA gene will be used in molecular detection of *Cryptosporidium* in the extracted DNA (Xiao et al. 2004). Positive secondary PCR products will be digested with restriction enzymes SspI and VspI to differentiate *Cryptosporidium* species and genotype. Products indicative of the presence of C. muris and C. andersoni will be digested with DdeI to further differentiate the two species. Samples with *C. parvum* will be further subtyped by a nested PCR based on the 60 kDa glycoprotein (GP60) gene, which allows the differentiation of anthroponotic *C. parvum* (IIa subtype family) from zoonotic *C. parvum* (IIa subtype family) based on sequence differences (Zhou et al. 2003). To counter residual PCR inhibitors, 400 ng/µl of non-acetylated bovine serum albumin (Sigma, St. Louis, MO) will be used in all PCR. The SOP for SSU rRNA-based PCR-RFLP and GP60-based PCR are attached.

DNA sequencing

To verify the results of RFLP genotyping and to subtype *C. parvum*, all positive SSU rRNA and GP60 products of the expected size and all PCR products of the unexpected sizes will be sequenced using an ABI 3100 genetic analyzer and the standardized operational procedures (attached).

Quality Control

Personnel and training

The project combines the expertise of environmental engineers, molecular biologists, parasitologists, and environmental microbiologists. The principal investigator, and EPA project officer and consultants have been working on the detection of waterborne pathogens for many years. The procedures needed to accomplish the objectives are routinely used by investigators involved in this project. All laboratory personnel involved in this project are trained in analysis of microorganisms in water, PCR and DNA sequencing and are active researchers in environmental microbiology and molecular parasitology. Laboratory personnel have also received yearly safety training. The postdoctoral associate work on the project will receive additional project-specific hands-on training and practices during the initial phase of the interagency agreement when the QAPP is being developed.

Equipment calibration and maintenance

A maintenance contract ($15,000/year) is in place for ABI thermocyclers and the DNA autosequencer. They will be inspected and serviced by certified technicians from Applied Biosystem once every half year and whenever problems arise. The heating blocks in thermocyclers and capillaries in the sequencers will be checked during each visit, and will be changed if needed. The performance of thermocyclers is checked periodically with PCR of positive controls and negative controls. The accuracy of the DNA autosequencer is checked every thirty runs with DNA sequencing of the PCR product of small subunit rRNA gene of a control *Cryptosporidium hominis* specimen. Pipettes used in the project will be calibrated once every six months by technicians from Precision Pipette, Atlanta, GA. The laboratory has a general contract with Four Seasons, Atlanta, GA for the maintenance and repair of all other laboratory equipment.

Reagents QC

Only commercial, molecular grade reagents and kits from reputable suppliers will be used in the project. They will be labeled with name, receiving and expiration date, and stored in designated, locked refrigerators and freezers at manufacturer-specific conditions. Care will be taken to avoid contact with contaminants during storage. The lot number of reagents used in each experiment will be recorded. Each lot of reagents will be initially tested with control or known samples in the laboratory through the entire process prior to the use in the project.

Sample handling

Water samples will be collected and filtered by DWSPP personnel and volunteers as specified in EPA method 1622. Filters with concentrated water will be sent in coolers packed with freezer blocks to certified commercial laboratories and CDC for *Cryptosporidium* detection by overnight express mail. A chain of custody form will be used to track the handling of samples (attached). Upon arrival in the laboratory, the filters will be handled only by the postdoctoral researcher working on the project, assigned a laboratory number, logged into laboratory notebook and sample database, stored in a locked refrigerator, and used in filtrate elution, oocyst isolation and DNA extraction within two days.

Quality Control Samples

Nested PCR-RFLP

(1) Positive controls. The following positive controls will be used with each sample processing runs:

Matrix spike. Matrix spikes with a known number of *Cryptosporidium* oocysts will be used to evaluate the performance of the molecular detection at the beginning of the project, and whenever false negative results are generated.

Method positive control. Reagent water seeded with a known number of *Cryptosporidium* oocysts will be included in each sample processing run prior to the start of DNA extraction. This allows us to verify whether the entire method is performing properly.

PCR positive control. To control for poor PCR amplification, one positive control will be used in each PCR run. For SSU rRNA-based nested PCR, the control we will use is DNA of *Cryptosporidium serpentis*, which is rarely seen in water samples. For GP60 nested PCR analysis of *C. parvum*, the positive control is *C. hominis* DNA.

Electrophoresis control. 100-bp DNA size ladders will be used in each electrophoresis run to verify that the electrophoresis apparatus works properly.

RFLP control. *Cryptosporidium* serpentis PCR products generated from the PCR positive control will be used in each run of restriction digestion of secondary SSU rRNA PCR products.

DNA sequencing control. The accuracy of DNA autosequencer is checked every thirty runs with DNA sequencing of the PCR product of small subunit rRNA gene of a control *C. hominis* specimen. Although it is unlikely that the misincorporation of one or two nucleotides during PCR will lead to misdiagnosis, to reduce error rates, PCR products will be sequenced directly without cloning, because our experience indicates this practice has lower error rates than sequencing of cloned PCR products.

(2) Negative control. The following negative controls will be used with each sample processing runs:

Method blank. A method blank containing only reagent water and no *Cryptosporidium* oocysts will be run with each sample processing and DNA extraction

PCR negative control. To control contamination, 2 negative controls (no DNA template) (one for primary PCR and one for secondary PCR) will be used in each PCR run.

(3) Analytic replicates. To ensure accurate diagnosis, each sample will be analyzed by PCR at least five times, and the results will be confirmed by the sequencing of all positive PCR products from each sample. All sequencing will be done in both 5' and 3' directions, using sequencing primers for every 500 bp. If there is any discrepancy (sequences from forward or reverse sequencing, sequences from two PCR products do

not match each other exactly) in the nucleotide sequence obtained from the 2 PCR products, a third PCR product from the sample will be sequenced to resolve the discrepancy. If only one positive PCR product from the five PCR replicates is available for sequencing and a unique sequence is obtained, DNA sequencing of more independent PCR products from the same sample will be conducted.

QC samples
The QC samples specified in the QA/QC guidelines of the Method 1622, such as matrix spikes, method positive control and method blanks specified above, will be used to minimize analytic errors. They will be analyzed by both microscopy and PCR in the same fashion as test samples.

False positive/negative prevention

False positive prevention. The QA/QC procedures outlined in the recent EPA publication (815-B-04-001) "Quality Assurance/Quality Control Guidance for Laboratories Performing PCR Analyses on Environmental Samples" will be followed strictly to reduce the occurrence of false positive results. In addition to the use of sample, method or PCR negative controls to ensure there will be no introduced contamination, all molecular analysis related to this project will be conducted in two designated rooms. Disposable gloves and white laboratory coat will be worn before sample handling. Sample preparation and DNA extraction will be done in areas separated from PCR preparation and DNA sequencing. Filter-plugged tips will be used in all pipetting of samples and reagents. Different pipettes will be used in sample preparation, PCR, gel electrophoresis, RFLP and sequencing. They will be decontaminated by treatment in Stratalinker before each use. Only disposable plastics free of nucleic acids and DNase and RNase will be used in DNA extraction and PCR. PCR preparation will be conducted in a laminar biological safety cabinet or PCR station, with sample loading occurring in a separated area.

False negative prevention. In addition to the use of positive controls, only analytical and preparational procedures published in the scientific community or standardized by us are used in this project. PCR methods used in the project have been optimized, and the PCR condition can be modified further if needed. The use of 400 ng/µl of non-acetylated bovine serum albumin in PCR will be employed to neutralize residual PCR inhibitors. If needed, DNA polymerase with know resistance to PCR inhibition, such as *Tth*, will replace *Taq* in PCR.

QC failures and corrective actions

If any positive control failures occur, all sample analyses associated with the control will be considered invalid. If amplification of a positive control fails to produce the specific amplification product, the integrity of the control and the PCR design will be examined to determine the reason for the failure. When determined, the reason for the failure will be documented and the controls and samples re-run. If the matrix spike produces a negative result (false negatives), the results of the entire run are also considered invalid. In this case, the matrix will be evaluated to determine whether PCR inhibition has occurred. This is done with spiking of the extracted DNA into PCR of positive control DNA extracted from low number of *C. parvum*

oocysts in reagent water. Corrections cannot be made if the PCR inhibition positive control shows evidence of inhibition. Corrections also cannot be extrapolated to samples for which matrix spike data are not available.

If PCR negative controls or method blanks produce specific amplification products, all sample analyses associated with the failed controls will be considered invalid. The source of contamination will be identified and eliminated. Once determined, the source of the contamination will be documented, and the samples in the batch will be recollected and reanalyzed, if possible. If the source of the contamination cannot be identified, additional types of negative controls will be added at various steps in the method to determine where the contamination is being introduced.

If false positives and false negatives persist after several attempts and causes of their occurrence cannot be decided, the project officer will be notified and peer scientists will be consulted.

Peer Review/Audits

Results including QA/QC problems will be summarized quarterly and reported to the EPA project officer for review. Results of the project will be presented at international scientific conferences attended by researchers in environmental microbiology. Manuscripts will be submitted for publication in peer-reviewed international scientific journals. Authorship of submitted peer reviewed report will be agreed upon by all investigators involved. These manuscripts and reports will also be reviewed internally by personnel in CDC and EPA. The overall project has been reviewed internally by both EPA and CDC scientists. Audits of the project can be conducted by the project officer or other agency representatives.

Data Management

Detailed record keeping will be enforced in this project. All experiments will strictly follow standard operational procedures. All numeric and graphic data generated from the studies will be kept in designated computers and/or notebooks, and analyzed by staff statisticians if necessary. Weekly meetings will be conducted at CDC to review QA/QC and work progress, plan daily work, and solve possible problems, and conference calls involving all investigators, project manager and consultants will be conducted quarterly and as needed. Raw data (pictorial, numerical, electropherogram, sequence) will be provided to the project officer and QAPP manager upon request. PCR, RFLP and sequence results will be initially interpreted by the postdoctoral associate involved in the analysis of samples, and confirmed by the principal investigators. Results of the study will be summarized once every three months, with written report submitted to the project officer. The PI will be responsible for the collection, validation, reduction and reporting of raw data. Significant findings will be summarized and incorporated into manuscripts in collaboration with the project officer, and they will be published in peer-reviewed scientific journals after internal agency reviews and clearance.

The laboratory notebook will be kept by the principal investigators recording the purpose of data, the procedure, the data file names, observations, conditions, reagent tracking and a brief analysis of the results. The notebook will be kept in accordance with the notebook policies of the laboratory. Hardcopies of relevant data will be stored in the laboratory in binders filed under the project name. Hardcopies of the data will also be taped into the principal investigator's laboratory notebook as a figure at the end of each day's experiments. Since the data are mostly generated by computers and instruments and mostly qualitative, transcription errors will not be an issue.

The data will be acquired on a number of instruments, including a UVP EPI Chemi II Darkroom gel documentation system and an Applied Biosystems 3100 Genetic Analyzer. All these instruments acquire data through linked computers and specialized software. For all the instruments, the data are stored on the instrument's computer hard drive, with copies transferred to network link drive by CD-ROM. The link drive files are backed-up automatically once everyday. The data file name will include project name, sample/DNA number, PCR target and number, and date of analysis. The raw data will be stored in network link drive for at least 10 years. The reduced data will be checked against raw when ambiguity arises.

Work Plan, Deliverable, and Timeline

Work Plan:

We will sample water at five sites in the Potomac River watershed once a month at base flow and six times during storm events during a one year period. The number of *Cryptosporidium* oocysts present will be determined by EPA method 1622, and *Cryptosporidium* genotypes and *C. parvum* subtypes will be determined by PCR, RFLP and DNA sequencing. The sources of *Cryptosporidium* contamination will be determined by determination of *Cryptosporidium* genotypes and subtypes found.

Deliverables

The following will be delivered by PIs to EPA project officer during the pilot project:

1. A quality assurance project plan (QAPP) to be submitted in consultation with the EPA project officer after the award of the agreement, but at least two weeks prior to any sample analysis. The QAPP is expected to be amended and approved prior to the commencement of sampling and analysis.
2. Quarterly reports describing technical progress and any delays that have occurred or are expected to occur in the coming quarter. This report should also include an accounting of funds expended in the preceding quarter.
3. A report detailing the *Cryptosporidium* genotyping and subtyping results of water samples collected at the study sites and the interpretation of the test results. It is expected that results of the project will be presented at a scientific conference and the report will be submitted to a peer reviewed journal for publication. Authorship of submitted peer reviewed report will be agreed upon by all principal investigators involved. The report is due prior to the end of the IAG project period.
4. A proposal for future research in more extensive monitoring and tracking of *Cryptosporidium* contamination in Potomac River watershed is due at the end

of the project.

Timeline

The initial pilot project is for 1.5 years, from October 1, 2006 to March 31, 2008. The following is the timeline of the project:

1. Preparation of QAPP: September 1, 2006- September 14, 2006
2. Collection of water samples: October 1, 2006-September 30, 2007
3. *Cryptosporidium* genotyping and subtyping: November 1, 2006-November 30, 2007
4. Final report: March 1, 2008

References

Jiang J, Alderisio KA, Singh A, Xiao L (2005a) Development of procedures for direct extraction of *Cryptosporidium* DNA from water concentrates and for relief of PCR inhibitors. Appl Environ Microbiol 71:1135-1141

Jiang J, Alderisio KA, Xiao L (2005b) Distribution of *Cryptosporidium* genotypes in storm event water samples from three watersheds in new york. Appl Environ Microbiol 71:4446-4454

Ruecker NJ, Bounsombath N, Wallis P, Ong CS, Isaac-Renton JL, Neumann NF (2005) Molecular forensic profiling of *Cryptosporidium* species and genotypes in raw water. Appl Environ Microbiol 71:8991-8994

Ruecker NJ et al. (2006) Tracking host sources of *Cryptosporidium* spp. in raw waterfor improved health risk assessment. Appl Environ Microbiol 72

USEPA (2001) Method 1622: *Cryptosporidium* in water by filtration/IMS/FA. In. Office of Water, U.S. Environmental Protection Agency, Washington D.C., Washington, D.C.

USEPA (2004) Quality Assurance/Quality Control Guidance for Laboratories Performing PCR Analyses on Environmental Samples Office of Water, USEPA, Washington, DC, 815-B-04-001, 56pp.

USEPA (2006) National primary drinking water regulations: Long Term 2 Enhanced Surface Water Treatment Rule. Federal Register 71:654-786

Xiao L, Alderisio K, Limor J, Royer M, Lal AA (2000) Identification of species and sources of *Cryptosporidium* oocysts in storm waters with a small-subunit rRNA-based diagnostic and genotyping tool. Appl Environ Microbiol 66:5492-5498.

Xiao L, Lal AA, Jiang J (2004) Detection and differentiation of *Cryptosporidium* oocysts in water by PCR-RFLP. Methods Mol Biol 268:163-176

Xiao L, Singh A, Limor J, Graczyk TK, Gradus S, Lal A (2001) Molecular characterization of *Cryptosporidium* oocysts in samples of raw surface water and wastewater. Appl Environ Microbiol 67:1097-1101.

Zhou L, Singh A, Jiang J, Xiao L (2003) Molecular surveillance of *Cryptosporidium* spp. in raw wastewater in Milwaukee: Implications for understanding outbreak occurrence and transmission dynamics. J Clin Microbiol 41:5254-5257

Appendix 1: Chain of Custody Form for Water Samples

Washington Suburban Sanitary Commission
14501 Sweitzer Lane, Laurel, MD 20707
301-206-8064 (phone); 301-206-8057 (fax)

Project: *Cryptosporidium* Source Tracking

DWSPP No.	Collection site	Volume collected	Date of collection	Sample description	Filter No.	Date of filtration	Comments

Relinquished by (Signature) Date/Time	Received by (Signature) Date/Time
Relinquished by (Signature) Date/Time	Received by (Signature) Date/Time
Relinquished by (Signature) Date/Time	Received by (Signature) Date/Time

Appendix 2: **SOP: Detection and differentiation of *Cryptosporidium* oocysts in water by a PCR-RFLP technique**

1. Materials

1.1. Supplies for immunomagnetic separation (IMS)
1. Dynabeads® anti-*Cryptosporidium* kit (for the isolation of *Cryptosporidium* oocysts only; product No. 730.01 for10 tests or No. 730.11 for 50 tests), or Dynabeads® GC-Combo kit (for the isolation of both *Cryptosporidium* oocysts and *Giardia* cysts; product No. 730.02 for 10 tests or No. 730.12 for 50 tests), Invitrogen, Carlsbad, CA.
2. Dynal Magnetic Particle Concentrators (Dynal MPC): Dynal MPC-S (product No. 120.20) and Dynal MPC-1 (product No. 120.01), Invitrogen.

1. 2. Supplies for DNA extraction
QIAamp DNA mini kit: Product No. 51304 (50 tests) or 51306 (250 tests), Qiagen Inc., Valencia, CA.

1.3. Supplies for PCR-RFLP
1. Primary PCR primers:
    a. Forward (F1): 5'-TTCTAGAGCTAATACATGCG-3'
    b. Reverse (R1): 5'-CCCATTTCCTTCGAAACAGGA-3'
2. Secondary PCR primers:
    a. Forward (F2): 5'-GGAAGGGTTGTATTTATTAGATAAAG-3'
    b. Reverse (R2): 5'- CTCATAAGGTGCTGAAGGAGTA -3'
3. 10X PCR Buffer with 15 mM Mg^{++}, Product No. N808-0129, PE Applied Biosystems, Foster City, CA.
4. 100 mM dNTP, Product No. U1240, Promega, Madison, WI. To make a 1.25 mM working solution, add 12.5 µL of each dNTP to 950 µL of distilled water. Store the working solution at –20ºC before use.
5. Taq polymerase, Product No. M2665, Promega, Madison, WI.
6. 25 mM $MgCl_2$, Product No. A351F, Promega, Madison, WI.
7. SspI, Product No. R0132L, New England BioLabs, Beverly, MA.
8. VspI, Product No. R6851, Promega, Madison, WI.
9. DdeI, Product No R0175L, New England BioLabs, Beverly, MA.

2. Methods

2.1. Immunomagnetic separation of *Cryptosporidium* oocysts from water pellets

1. Process 10-L of water samples through filtration, elution and concentration steps, following method 1622 or 1623 by the U.S. Environmental Protection Agency (see **Note 1**).
2. Wash the concentrated water pellets in 15 ml polypropylene tube twice with distilled water by centrifugation at 1,500 g for 10 min.
3. Equilibrate the washed samples, 10 X Buffer A and 10 X Buffer B from the Dynabeads kit to room temperature.

4. Add 1 ml of 10 X Buffer A and 1 ml of 10 X Buffer B into the 15 ml tube containing washed sample. Use only 0.5 ml of the water concentrate if the pellet is bigger than 0.5 ml in volume.
5. Resuspend the beads fully by vortexing the vial for 10 sec, and add 100 µl of Dynabeads to the 15 ml tube.
6. Add distilled water to give a final volume of 10 ml.
7. Rotate at 15~20 rpm for 1 h at room temperature.
8. Prepare the 1 X dilution of Buffer A. One ml of 1 X Buffer A will be required for each sample.
9. At the end of incubation, capture the Dynabeads in the 15 ml tube using MPC-1, and decant the solution in the tube
10. Resuspend the Dynabeads with 1 ml of 1 X Buffer A, and transfer the suspension into 1.5 ml microfuge tube.
11. Capture the Dynabeads in microfuge using MPC-S, and decant the solution in the tube. The Dynabeads with bound *Cryptosporidium* oocysts will be used in DNA extraction (see **Note 2**).

2.2. DNA extraction using QIAamp® DNA Mini kit

1. Add 180 µL of Buffer ATL to a 1.5 ml microfuge tube containing IMS-isolated *Cryptosporidium* oocysts, and vortex for 30 sec.
2. Freeze-thaw five times at –70℃ (or dry ice) and 56℃.
3. Add 20 µL of proteinase K to the tube, vortex for 10 sec, and incubate at 56℃ overnight.
4. Add 200 µL of Buffer AL to the sample, vortex, and incubate the tube at 70℃ for 10 min.
5. Centrifuge at full speed to precipitate the undigested pellet.
6. Transfer the supernatant into a new 1.5 ml tube.
7. Add 200 µL of ethanol to the sample and vortex for 15 sec.
8. Carefully transfer the mixture to a QIAamp spin column without wetting the rim, and centrifuge the column at 6,000 g for 1 min.
9. Place the spin column in a clean 2 ml collection tube, and discard the tube containing the filtrate.
10. Add 500 µL of Buffer AW1 without wetting the rim, and centrifuge at 6,000 g for 1 min.
11. Place the spin column in a clean 2 ml collection tube, and discard the tube containing the filtrate.
12. Add 500 µl of Buffer AW2 without wetting the rim, and centrifuge at full speed for 3 min.
13. Place the spin column into a clean 1.5 ml microfuge tube, and discard the tube containing the filtrate.
14. Add 100 µL of Buffer AE, and incubate the tube at room temperature for 1 min.
15. Centrifuge the tube at 6,000 g for 1 min.
16. Save the filtrate containing DNA and the store the extraction at –20ºC.

2.3. PCR-RFLP analysis of the SSU rRNA gene

1. Primary PCR:

A. Preparation of master mixture. For each PCR reaction, prepare the following (see **Note 3**):

10X Perkin-Elmer PCR buffer	10 µL
dNTP (1.25 mM)	16 µL
F1 primer (40 ng/μl)	2.5 µL
R1 primer (40 ng/μl)	2.5 µL
$MgCl_2$ (25 mM)	6 µL
Bovine serum albumin (10 mg/ml)	4 µL
Distilled water	56.5 µL
Taq polymerase	0.5 µL
Total	98 µL

B. Add 98 µL of the master mixture to each PCR tube.
C. Add 2 µL of DNA sample to each tube.
D. Run the following PCR program:

94°C: 3 min
35 cycles of: 94°C for 45", 55°C for 45" and 72°C for 1 min
72 °C for 7 min
4 °C soaking

2. Secondary PCR
A. Preparation of master mixture. For each PCR reaction, prepare the following:

10X Perkin-Elmer PCR buffer	10 µL
dNTP (1.25 mM)	16 µL
F2 primer (40 ng/μl)	5 µL
R2 primer (40 ng/μl)	5 µL
$MgCl_2$ (25 mM)	6 µL
Distilled water	55.5 µL
Taq polymerase	0.5 µL
Total	98 µL

B. Add 98 µL of the master mixture to each PCR tube.
C. Add 2 µL of the primary PCR reaction to each tube.
D. Run the following PCR program:

94°C: 3 min
35 cycles of: 94°C for 45", 58°C for 45" and 72°C for 1 min
72 °C for 7 min
4 °C soaking

E. Run electrophoresis on 1.5% agarose gel with 20 µL of the PCR product.

2.4. RFLP:

1. Preparation of master mixture using the following formula, which is for one restriction digestion reaction (**see Note 4**).

	Buffer	Water	Enzyme
Ssp I	4 µL of New England BioLabs Buffer SspI	22 µL	4 µL

Vsp I	4 µL of Promega Buffer D	24 µL	2 µL
Dde I	4 µL of New England BioLabs Buffer 3	24 µL	2 µL

2. Transfer 30 µL of master mixture to each tube, add 10 µL of secondary PCR reaction to the tube, and mix well.
3. Incubate in 37°C waterbath for 2 h or overnight.
4. Run electrophoresis on 1.2% argrose gel with the entire 40 µL of restriction digestion reaction.
5. Identify *Cryptosporidium* species and genotypes based RFLP banding patterns (see **Note 5**)

3. Notes

1. Protocols for EPA methods 1622 and 1623 can be downloaded from the EPA website (http://www.epa.gov/waterscience/methods/1622.pdf and http://www.epa.gov/waterscience/methods/1623.pdf). Even though filtration of 10 L water samples is recommended, larger volumes of finished water can and should be filtered and used in the analysis. For raw wastewater, filtration with the standard Envirocheck capsule filters can be problematic. We normally process the pellets from 50 ml of grab samples of raw wastewater directly for IMS without filtration, after they were washed twice by centrifugation at 1,500 g for 10 min.
2. IMS-oocyst pellets can be stored at -20ºC before they are used in DNA extraction. No detachment of Dynabeads from oocysts is needed prior to the DNA extraction.
3. The magnesium concentration used in both primary and secondary PCR is 3 mM, which is higher than normal PCR. Even though concentrations lower than 3 mM generally do not work well for the SSU rRNA-based PCR, it is recommended that magnesium concentration should be optimized in each laboratory prior to sample analysis.
4. We generally do *Ssp* I and *Vsp* I restriction digestions to differentiate common *Cryptosporidium* species and genotypes. *Dde* I digestion is to differentiate *C. andersoni* from *C. muris*, and is done only when results of *Ssp* I digestion have shown the RFLP pattern of *C. andersoni*/*C. muris*. In most areas, *C. andersoni* is found much more frequently in water than *C. muris*.
5. *Ssp I* and *Vsp I* RFLP patterns for some common *Cryptosporidium* species and genotypes are shown in Table 1 and Figure 1, and the *Dde* I RFLP patterns for *C. andersoni* and *C. muris* are shown in Figure 2. Occasionally, due to the genetic heterogeneity between parasites of the same genotype and the multicopy nature of the SSU rRNA gene, RFLP banding patterns deviations from the characteristic patterns can occur. New and unusual *Cryptosporidium* parasites are also sometimes seen in water samples. A few genotypes can not be differentiated from each other by RFLP analysis. These require DNA sequencing of the secondary PCR product using the secondary forward and reverse primers for confirmation. The most polymorphic region of the SSU rRNA sequences of known *Cryptosporidium* parasites are shown in Figure 3. Likewise, isolates within each genotype can differ somewhat in SSU rRNA sequences due to the presence of heterogeneous copies of the gene and intragenotypic variations. Thus minor differences (<5 bp) in the SSU rRNA sequences generally do not warrant new genotype designation.

4. References

USEPA. 2004. Quality Assurance/Quality Control Guidance for Laboratories Performing PCR Analyses on Environmental Samples. EPA Office of Water, Washington, DC. 815-B-04-001; 56pp.

Xiao L, Lal AA, Jiang J (2004) Detection and differentiation of *Cryptosporidium* oocysts in water by PCR-RFLP. Methods Mol Biol 268:163-176

Table1. Restriction fragment length (in base pairs; only sizes of visible bands are shown) polymorphism in the SSU rRNA gene of common *Cryptosporidium* spp. and genotypes.

Species	PCR fragment	*Ssp* I digestion*	*Vsp* I digestion*
C. muris/C. andersoni	833	385, 448	102, 731
C. serpentis	831	370, 414	102, 729
C. baileyi	826	254, 572	102/104, 620
C. felis	864	390, 426	102/104, 182, 476
C. meleagridis	833	108, 254, 449	102/104, 171, 456
C. wrairi	834	109, 254, 449	102/104, 628
C. saurophilum	834	109, 255, 418	102/104, 628
C. canis	829	105, 254, 417	94/102, 633
Cryptosporidium ferret genotype	837	111, 254, 449	102/104, 174, 457
C. suis	838	365, 453	102/104, 632
Cryptosporidium marsupial genotype	837	109, 254, 441*	102/104, 631
C. hominis	837	111, 254, 449	70, 102/104, 561
C. parvum A gene	834	108, 254, 449	102/104, 628
C. parvum B gene	831	119, 254, 449	102/104, 625

Cryptosporidium mouse genotype	838	112, 254, 449	102/104, 175, 457

*An additional upper band (about 583 bp) from the heterogeneous copy of the gene is usually present.

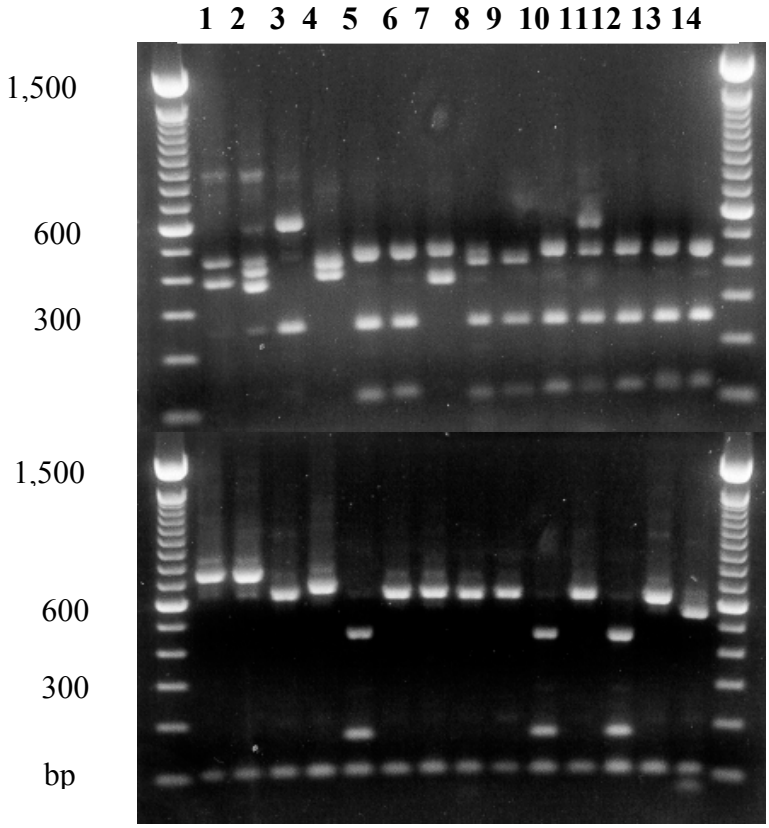


Figure 1. Differentiation of common *Cryptosporidium* species and genotypes by a nested PCR-RFLP procedure based on the SSU rRNA gene. Lane 1: *C. muris* or *C. andersoni*; lane 2: *C. serpentis*; lane 3: *C. baileyi*; lane 4: *C. felis*; lane 5: *C. meleagridis*; lane 6: *C. wrairi*; lane 7: *C. suis*; lane 8: *C. canis*; lane 9: *C. saurophilum*; lane 10: *Cryptosporidium* ferret genotype; lane 11: *Cryptosporidium* marsupial genotype; lane 12: *Cryptosporidium* mouse genotype; lane 13: *C. parvum*; and lane 14: *C. hominis*. The upper panel are *Ssp* I digestion products, and the lower panel are *Vsp* I digestion products. Molecular markers are 100-bp ladders.

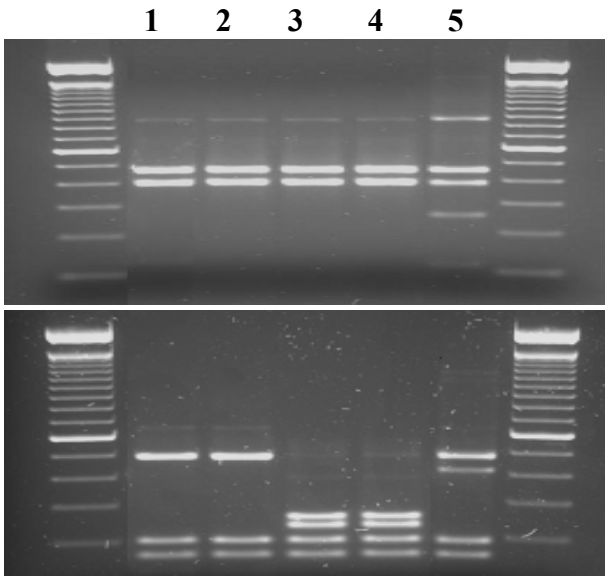


Figure 2. Differentiation of *C. andersoni* and *C. muris* by RFLP analysis of SSU rRNA gene PCR products using *Dde* I. Lanes 1 and 2: *C. andersoni*; lanes 3 and 4: *C. muris*; and lane 5: *C. andersoni* and *C. hominis*. The upper panel are *Ssp* I digestion products, and the lower panel are *Dde* I digestion products. The top band in lane 5 of the *Ssp* I products was due to partial digestion. Molecular markers are 100-bp ladders.

```
Human              TTGGATTTCTGTTA-ATAATTTA----TATAAAATATTTTG------ATGAATATT----TATATAATATTAACATAATT
Rabbit             .............-........----.....T........------..AG.....----....................
Bovine             .............-........----.............------.........----....................
Mouse              .............-........----.............A------..T......----....................
Ferret             .............-........----.............------..T......----....................
C. wrairi          .............-........----....T........------.A-......----....................
C. meleagridis     .............-........----....T.......-.------..T......----....................
Pig                .............-........----....T......---------T.T......----....................
Marsupial          .............-........----....TT...C...T------TAAGG.G..----....................
Opossum I          .............-........----....G....C...T------T.AGG.G..----....................
Coyote             .............-........----....T.......-A--------AC.....----....................
C. canis           .............-........----....T.......-A--------AC.....----....................
Bear               .............-........----....T........A-------AT.....----....................
Deer mouse         ............C.-...G...T---G...GT....C..CA--------ACG....----C....G..G..G........
Opossum II         .............-.......TTATG...GC.....A.A.TTAAAA..ATG....GTTG...G................
Fox                .............-........----.....T..T..---------.ACG.....-----...................
Skunk              .............-........----....T........A------..T......----....................
C. felis           .............-...CC...----....T........TTTT---T.A......A--A...G...G............
C. saurophilum     .............-........----....T......AC.----------G....----....................
Deer               ..AAT.........-..-T....----....C....C.AC.----------G....----....................
Cattle             ..AATC........-..-T....----....T.....CAC.----------.....----....................
Goose              ...........C..-..-T...G----C...C....CCAC.----------G....----...G...A....G......C
C.baileyi          .............-..-.C...----....C....CCAC.----------G....----.......C............
Snake              .............-..-TC.CG----....TT...C.AC.----------G...G----.....G.C............
C. andersoni       .............GT........----..-.T.T..CCAA.----------G..A.---TAT....T...C.....CC..
C. muris           .............GT.....C..----..-.T.T..C.AA.----------G...A---TAT....T...C.....CC..
C. galli           .............GC..C..-..----..-TT.TC.CCAA.----------G..A----TA.........C.....CC.C
C. serpentis       .............GT..T-....----..-.T.T....AA.----------G..A.---A.T........C.....CC..
Tortoise           .............GT..G...AT----A.-TTG.....AA.----------....----TAT....TC..C.....CC..

Human              CAT-ATTACTATTTTT-------------TTTTTTA----------------------------GTATATGAAATTTTAC
Rabbit             ...-............-------------....--.----------------------------................
Bovine             ...-........A-----------------.A....----------------------------................
Mouse              ...-........AA..A------------.......----------------------------................
Ferret             ...-.......AA...-------------..G...G----------------------------................
C. wrairi          ...-........A.A.-------------..--...----------------------------................
C. meleagridis     ...-.......AA...A------------..----.----------------------------................
Pig                ...-........AA..-------------...A...----------------------------................
Marsupial          ...-........A..---------------......----------------------------................
Opossum I          ...-........--.---------------......----------------------------................
Coyote             ...-........------------------.-.A..----------------------------................
C. canis           ...-........------------------.-.A..----------------------------..........C.....
Bear               ...-........------------------AAC...----------------------------................
Deer mouse         .G.-......--------------------......----------------------------.....C..........
Opossum II         ...-........A.GGTGGATTGGTGAAG.......CTTTTTTCCAGTCACACCGGGAATTATG................
Fox                ...-........A..-------------A..-...----------------------------.......G.......
Skunk              ...-........A..A-------------......----------------------------................
C. felis           ...-...TT..AGAC.GA-----------A......GTTTTG---------------------ATA.............
C. saurophilum     ...-......T.A-----------------......G--------------------------A................
Deer               ...-......---------------------.....----------------------------..........C.....
Cattle             ...-......---------------------.....----------------------------..........C.....
Goose              .GC-.....CTCGC-------------------------------------------------....GC.G........
C.baileyi          ..C-......--------------------.A....A--------------------------A....G.....C.....
Snake              ...-......---------------------...A..----------------------------................
C. andersoni       .C.-....TAT.--------------------C.A.AT--------------------------A....G..........
C. muris           .C.-....TAT..-------------------C.A.AT--------------------------A....G....C.....
C. galli           .C.T....TATC----------------------..AT--------------------------A..G.G.G........
C. serpentis       .C.-....TAT.--------------------....AT--------------------------A....G..G.......
Tortoise           .C.-.....A..---------------------...TT--------------------------.....G..........
```

Figure 3. Sequence diversity among *Cryptosporidium* species and genotypes in the polymorphic region of the SSU rRNA gene. Dots denote sequence identity to the *C. hominis* (human) (top sequence) and dashes denote deletions. Human: *C. hominis*; rabbit: *Cryptosporidium* rabbit genotype; bovine: *C. parvum*; mouse: *Cryptosporidium* mouse genotype; ferret: *Cryptosporidium* ferret genotype; pig: *C. suis*; marsupial: *Cryptosporidium* marsupial genotype; opossum I: *Cryptosporidium* opossum genotype I; coyote*: C. canis* coyote genotype; bear: *Cryptosporidium* bear genotype; deer mouse: *Cryptosporidium* deer mouse genotype; opossum II: *Cryptosporidium* opossum genotype II; fox: an unnamed *Cryptosporidium* sp. in foxes; deer: an unnamed *Cryptosporidium* sp. in deer; cattle: *C. bovis*; goose: an unnamed *Cryptosporidium* sp. in geese; snake: an unnamed intestinal *Cryptosporidium* sp. in snakes; and tortoise: an unnamed gastric *Cryptosporidium* sp. in tortoises.

Appendix 3: **SOP: Direct sequencing PCR product**

1. Materials

1.1 Supplies for Cleaning PCR product

Montage-PCR (Catalog No UFC7PC250, Millipore, Bedford, MA)

1.2 Supplies for DNA sequencing

A. 3100 AB Prism® Genetic Analyzer (Applied Biosystems, Foster City, CA)
B. BigDye® Terminator v3.1 Cycle Sequencing Kit (Product No 4337455, Applied Biosystems)
C. Centrisep spin Columns (Catalog No CS-901, Princeton Separations, Adelphia, NJ)
D. MicroAmp Optical 96 well reaction plate (Part No N801-0560, Applied Biosystems)
E. 3100 Genetic Analyzer Plate Septa 96-well (Part No 4315933, Applied Biosystems)

2. Methods

2.1 Cleaning PCR product

A. Use secondary PCR product of selected gene to sequence.
B. Insert reservoir from Montage-PCR kit into Spin vial (purple on top, white side bottom).
C. Pipette 400 µl of water into Reservoir.
D. Add 100 µl PCR product. Seal and spin at 1000 X g for 15 min.
E. Remove sample reservoir and place it upright (purple on top) onto new spin vial.
F. Add 20 µl water (pre-heated to 60 ℃).
G. Invert the reservoir (white on top) and spin at 1000 X g for 2 minutes.
H. Save cleaned PCR product.

2.2 Sequencing reaction

A. Preparation of master mixtures. Forward and reverse primer sequenced each PCR product. For each sequencing reaction, prepare the following:

   BigDye terminator buffer: 4 µl
   BigDye Terminator: 1 µl
   Forward or reverse primer (40 ng/µl): 2 µl
   Water: 9.5 µl

B. Add 19.5 µl of the master mixture to each tube.
C. Add 0.5 µl cleaned PCR products to each tube.
D. Run the following program:

   25 cycles of: 90℃ for 10 sec, 50℃ for 5 sec, and 60℃ for 4 min

1 cycle of 4℃ soaking

2.3 Cleaning sequencing reaction.

A. Rehydrate the columns to be used: tap dehydrated contents down, add 0.8 ml reagent water, cap the top, mix well by vortexing, and allow hydration for at least 30 min.
B. Remove bubbles by inverting and sharply tapping the column.
C. Remove top cap, and then remove the column end stopper from the bottom. Allow excess column fluid to drain (gravity) into a Wash Tube (2 ml) for at least 20 min, and remove excess fluid (decant a bit) if the fluid does not flow down through the end of the column.
D. Remove bottom cap and centrifuge at 750 X g for 3 min. Keep track of position/orientation of column with respect to G forces, so next time it is placed in same orientation.
E. Transfer 20 µl of amplified sequencing reaction products on top of gel of each Centri-Sep, do not disturb top of gel. Do not touch sides of tube.
F. Place loaded column in Sample Collection Tube, with proper column orientation. Centrifuge 3 min at 750 X g.
G. Dry sample in vacuum centrifuge. Store in freezer.

2.4 Running DNA sequencing.

A. Prior to running, add 15 µl of BigDye Formamide to dehydrated DNA.
B. Transfer to MicroAmp Optical 96 well reaction plate.
C. Place plate lid (3100 Genetic Analyzer Plate Septa 96-well).
D. Heat at 94℃ for 5 min, then -20℃ for at least 3 min.
E. Load the reaction products to 3100 AB Prism Genetic Analyzer
F. Run the sequencing for using the software installed on the sequencing computer, following the prompts.

Appendix 4: **SOP: Subtyping *CCryptosporidium parvum* and *Cryptosporidium hominis* by Sequence Analysis of the 60 kDa Glycoprotein (GP60) Gene**

1. MATERIALS

1.1. Supplies for GP60 PCR

10. Primary PCR primers:
    a. Forward (F1): 5'-ATAGTCTCCGCTGTATTC -3'
    b. Reverse (R1): 5'-GGAAGGAACGATGTATCT -3'
11. Secondary PCR primers:
    a. Forward (F2): 5'-TCCGCTGTATTCTCAGCC-3'
    b. Reverse (R2): 5'-GCAGAGGAACCAGCATC-3'
12. 10X PCR Buffer with 15 mM Mg++, Product No. N808-0129, PE Applied Biosystems, Foster City, CA.
13. 100 mM dNTP, Product No. U1240, Promega, Madison, WI. To make a 1.25 mM working solution, add 12.5 µL of each dNTP to 950 µL of distilled water. Store the working solution at –20ºC before use.
14. Taq polymerase, Product No. M2665, Promega, Madison, WI.
15. 25 mM $MgCl_2$, Product No. A351F, Promega, Madison, WI.
16. BigDye® Terminator V3.1 Cycle Sequencing Kit, Product Np. 4336917, Applied Biosystems, Foster, CA.
17. Bovine serum albumin (10 mg/ml), Product No. B-6917, Sigma, St. Louis, MO.

1.2. Supplies for Cleaning PCR Product

Montage-PCR (Catalog No UFC7PC250, Millipore, Bedford, MA)

1.3. Supplies for DNA Sequencing

F. Intermediary sequencing primer (R3): 5'-GAGATATATCTTGTTGCG-3'.
G. 3100 AB Prism® Genetic Analyzer (Applied Biosystems, Foster City, CA)
H. BigDye® Terminator v3.1 Cycle Sequencing Kit (Product No 4337455, Applied Biosystems)
I. Centrisep spin Columns (Catalog No CS-901, Princeton Separations, Adelphia, NJ; CDC glassware Cat# 95103)
J. MicroAmp Optical 96 well reaction plate (Part No N801-0560, Applied Biosystems)
K. 3100 Genetic Analyzer Plate Septa 96-well (Part No 4315933, Applied Biosystems)

2. PROCEDURES

2.1. PCR of GP60 Gene

2.1.1. Primary PCR:

A. Preparation of master mixture. For each PCR reaction, prepare the following:

10X Perkin-Elmer PCR buffer	10 µL
dNTP (1.5 mM)	16 µL
F1 primer (40 ng/μl)	5 µL
R1 primer (40 ng/μl)	5 µL
$MgCl_2$ (25 mM)	6 µL
Bovine serum albumin (10 mg/ml)	4 µL
Distilled water	52.5 µL
Taq polymerase	0.5 µL
Total	99 µL

B. Add 99 µL of the master mixture to each PCR tube.
C. Add 1 µL of DNA sample to each tube.
D. Run the following PCR program:

94°C: 3 min
35 cycles of: 94°C for 45", 50°C for 45" and 72°C for 1 min
72 °C for 7 min
4 °C soaking

2.1.2. Secondary PCR

A. Preparation of master mixture. For each PCR reaction, prepare the following:

10X Perkin-Elmer PCR buffer	10 µL
dNTP (1.5 mM)	16 µL
F2 primer (40 ng/μl)	5 µL
R2 primer (40 ng/μl)	5 µL
$MgCl_2$ (25 mM)	6 µL
Distilled water	54.5 µL
Taq polymerase	0.5 µL
Total	97.5 µL

B. Add 97.5 µL of the master mixture to each PCR tube.
C. Add 2.5 µL of the primary PCR reaction to each tube.
D. Run the following PCR program:

94°C: 3 min
35 cycles of: 94°C for 45", 50°C for 45" and 72°C for 1 min
72 °C for 7 min
4 °C soaking

2.1.3. Detection of secondary PCR products

Run electrophoresis on 1.5% agarose gel with 15 µL of the secondary PCR product.

2.2. Cleaning of Secondary PCR Products

I. Use positive secondary PCR products of the expected size in DNA sequencing.

J. Insert reservoir from Montage-PCR kit into Spin vial (purple on top, white side bottom).
K. Pipette 400 µl of water into the reservoir.
L. Add up to 100 µl PCR product into the reservoir. Seal and spin at 1000 X g for 15 min.
M. Remove sample reservoir and place it upright (purple on top) onto new spin vial.
N. Add 20 µl water (pre-heated to 60 ℃).
O. Invert the reservoir (white on top) and spin at 1000 X g for 2 minutes.
P. Save cleaned PCR product at -20C.

2.3. DNA Sequencing Reaction

E. Preparation of master mixtures. The cleaned PCR products will be sequenced using three sequencing primers (the forward and reverse PCR primers and the intermittent primer).

   For each sequencing reaction, prepare the following:

   BigDye terminator buffer: 4 µl
   BigDye Terminator: 1 µl
   Primer (40 ng/µl): 2 µl
   Water: 9.5 µl

F. Add 19.5 µl of the master mixture to each tube.
G. Add 0.5 µl cleaned PCR products to each tube.
H. Run the following program:

   25 cycles of: 96℃ for 10 sec, 50℃ for 5 sec, and 60℃ for 4 min

   1 cycle of 4℃ soaking

2.4. Cleaning of Sequencing Reaction

H. Rehydrate the columns to be used: tap dehydrated contents down, add 0.8 ml reagent water, cap the top, mix well by vortexing, allow hydration for at least 30 min.
I. Remove bubbles by inverting and sharply tapping the column.
J. Remove top cap, and then remove the column end stopper from the bottom. Allow excess column fluid to drain (gravity) into a Wash Tube (2 ml) for at least 20 min, and remove excess fluid (decant a bit) if the fluid does not flow down through the end of the column.
K. Remove bottom cap and centrifuge at 800 X g for 2 min. Keep track of position/orientation of column with respect to G forces, so next time it is placed in same orientation.
L. Transfer 20 µl of amplified sequencing reaction products on top of gel of each Centri-Sep, do not disturb top of gel. Do not touch sides of tube.
M. Place loaded column in Sample Collection Tube, with proper column orientation. Centrifuge 3 min at 800 X g.
N. Dry sample in vacuum centrifuge. Store in freezer.

2.5. Running DNA Sequencing

G. Prior to running, add 15 µl of BigDye Formamide to dehydrated DNA.
H. Transfer to MicroAmp Optical 96 well reaction plate.
I. Place plate lid (3100 Genetic Analyzer Plate Septa 96-well).
J. Heat at 94℃ for 5 min, then -20℃ for at least 3 min.
K. Load to 3100 AB Prism Genetic Analyzer for sequencing using proper program.

2.6. Sequence Analysis

A. Read out the electropherograms generated by the sequencer using the ChromasPro software (www.technelysium.com.au/ChromasPro.html).
B. Aligned the GP60 sequences generated with each other and reference sequences using the software ClustalX (ftp://ftp-igbmc.u-strasbg.fr/pub/ClustalX/).
C. Check the sequence alignment for sequencing accuracy using the software BioEdit (www.mbio.ncsu.edu/BioEdit/**bioedit**.html).
D. Recheck the electropherograms for any sequence uncertainty.
E. Determine subtype designation based on sequence identity to reference sequences and the number of trinucleotide repeats (see below).

2.7. Subtype Nomenclatures

The GP60 subtype nomenclature was previously described (Sulaiman et al., 2005). This gene has a trinucleotide repeat region in the 5' end. It categorizes *C. hominis* and *C. parvum* to several subtype families based on nucleotide sequences in the non-repeat regions, such as Ia, Ib, Id, Ie and If for *C. hominis* and IIa, IIb, IIc and IId for *C. parvum*. Within each subtype family, subtypes differ from each mostly in the number of the trinucleotide repeats TCA, TCG and TCT (only seen in Ie). For each subtype, the name starts with the subtype family designation (Ia, Ib, Id, Ie, and If for *C. hominis* and IIa, IIb, IIc and IId for *C. parvum*) followed by the number of TCA (represented by the letter A), TCG (represented by the letter G) and TCT(represented by the letter T) repeats found. Thus, the name IbA10G2 indicates that parasite belongs to *C. hominis* subtype family Ib and has 10 copies of the TCA repeat and 2 copies of the TCG repeat in the trinucleotide repeat region of the GP60 gene. In IIa subtype family, a few subtypes also has two copies of the AACATCA sequence right after the trinucleotide repeats, which are represented by "R2" (R1 for most subtypes). The subtype family Ia also has different copies of a 15-bp repetitive sequence 5'-AAA/G ACG GTG GTA AGG- 3' (the last copy is 13-bp) shortly after the trinucleotide repeats, which is represented by the letter "R".Thus, R4 indicates the presence of four copies of the 13-15-bp repeat in the GP60 gene. The subtype family IIc was previously known as Ic. It differs from other subtype families by having no variation in the number of trinucleotide repeats (all A5G3) and by having extensive sequence polymorphism in the 3' end of the gene. Thus, subtypes in the family are differentiated from each other by the additional letters a to g, such as IIcA5G3a and IIcA5G3b. Occasionally, a few subtypes in other families have the same trinucleotide repeat sequence, but differ from each other by one or two nucleotides

in the sequence after the repeat region. They are also differentiated from each other by the addition of small letters, such as IIdA18G1a, IIdA18G1b, IIdA18G1c and IIdA18G1d.

References

1. **Alves, M., L. Xiao, I. Sulaiman, A. A. Lal, O. Matos, and F. Antunes.** 2003. Subgenotype analysis of *Cryptosporidium* isolates from humans, cattle, and zoo ruminants in Portugal. J Clin Microbiol **41:**2744-2747. (Reference for the methodology)

2. **Sulaiman, I. M., P. R. Hira, L. Zhou, F. M. Al-Ali, F. A. Al-Shelahi, H. M. Shweiki, J. Iqbal, N. Khalid, and L. Xiao.** 2005. Unique endemicity of cryptosporidiosis in children in Kuwait. J Clin Microbiol **43:**2805-2809. (Reference for the subtype nomenclature)