EPA/600/R-03/047
September 2002

Animal Source Identification Using A *Cryptosporidium* DNA Characterization Technique

by

Michael Royer
U.S. Environmental Protection Agency
Edison, New Jersey 08837

Lihua Xiao and Altaf Lal
Centers for Disease Control and Prevention
Atlanta, Georgia 30341

NATIONAL RISK MANAGEMENT RESEARCH LABORATORY
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
CINCINNATI, OHIO 45268

NATIONAL CENTER FOR INFECTIOUS DISEASES
DIVISION OF PARASITIC DISEASES
CENTERS FOR DISEASE CONTROL and PREVENTION
U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
ATLANTA, GEORGIA 30341

NOTICE

The U.S. Environmental Protection Agency through its Office of Research and Development partially funded and collaborated in the research described here under EPA/NRMRL-HHS/CDC interagency agreement DW 75937984. It has been subjected to the Agencies' peer and administrative review and has been approved for publication as an EPA document.

FOREWORD

The U.S. Environmental Protection Agency is charged by Congress with protecting the Nation's land, air, and water resources. Under a mandate of national environmental laws, the Agency strives to formulate and implement actions leading to a compatible balance between human activities and the ability of natural systems to support and nurture life. To meet this mandate, EPA's research program is providing data and technical support for solving environmental problems today and building a science knowledge base necessary to manage our ecological resources wisely, understand how pollutants affect our health, and prevent or reduce environmental risks in the future.

The National Risk Management Research Laboratory is the Agency's center for investigation of technological and management approaches for preventing and reducing risks from pollution that threatens human health and the environment. The focus of the Laboratory's research program is on methods and their cost-effectiveness for prevention and control of pollution to air, land, water, and subsurface resources; protection of water quality in public water systems; remediation of contaminated sites, sediments and ground water; prevention and control of indoor air pollution; and restoration of ecosystems. NRMRL collaborates with both public and private sector partners to foster technologies that reduce the cost of compliance and to anticipate emerging problems. NRMRL's research provides solutions to environmental problems by: developing and promoting technologies that protect and improve the environment; advancing scientific and engineering information to support regulatory and policy decisions; and providing the technical support and information transfer to ensure implementation of environmental regulations and strategies at the national, state, and community levels.

This publication has been produced as part of the Laboratory's strategic long-term research plan. It is published and made available by EPA's Office of Research and Development to assist the user community and to link researchers with their clients.

Hugh W. McKinnon, Director
National Risk Management Research Laboratory

ABSTRACT

This document summarizes the application of a particular molecular method to improve detection and differentiation of species and genotypes of *Cryptosporidium* oocysts found in environmental samples. Of particular interest is the method's potential for determining the source animal types of oocysts in water samples. The molecular method is a nested polymerase chain reaction (PCR)-restriction fragment length polymorphism (RFLP) procedure that characterizes the small sub-unit (SSU) ribosomal RNA gene. The method was previously developed for characterizing oocyst DNA from clinical samples. The current project explores the method's applicability to environmental water samples, which have greater diversity of oocyst species and strains, lower concentrations of oocysts, and different interferents than clinical samples. Results include demonstrating that the method is capable of detection and differentiation of at least 10 species and 22 genotypes of *Cryptosporidium*; method sensitivity demonstrated to a single oocyst with laboratory samples; and detection and differentiation of oocysts from oyster gill washings and hemolymph, storm water, surface water, and raw waste water. The method's capability to determine an oocyst's source animal type was demonstrated by identification in environmental water samples of host-adapted *Cryptosporidium* species and genotypes that were consistent with the source animal types (i.e., humans, farm animals, wildlife, and/or pets) inhabiting the sampled watersheds.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGMENTS

Much of the work described in this report on the SSU rRNA nested PCR-RFLP method was conducted under EPA/NRMRL-HHS/CDC interagency agreement DW 75937984. The work conducted under this interagency agreement built upon method development work previously conducted under EPA/Office of Water-HHS/CDC interagency agreement 75937730, and a substantial amount of collaboration also occurred. Key collaborating researchers include Ronald Fayer, Agriculture Research Service of the U.S. Department of Agriculture; Kerri Alderisio, New York City Department of Environmental Protection; Una Ryan and R.C. Andrew Thompson, Murdoch University, Western Australia; Steve Gradus and Ajaib Singh, City of Milwaukee Public Health Laboratories; Thaddeus K. Graczyk, Johns Hopkins School of Hygiene and Public Health, and Joseph Limor, and Irshad Sulaiman, Centers for Disease Control and Prevention.

INTRODUCTION

The purpose of this document is to summarize progress to date on the application to water samples of a molecular method for improved detection and differentiation of species and genotypes of *Cryptosporidium* oocysts. Of particular interest is the method's potential for determining the source animals of oocysts found in water samples. The molecular method is a small sub-unit ribosomal RNA (SSU rRNA) gene-based nested polymerase chain reaction (PCR)-restriction fragment length polymorphism (RFLP) procedure previously developed for characterizing oocyst DNA from clinical samples. Exploring its applicability beyond clinical samples to water samples – where the diversity of oocysts is greater, the concentration of oocysts is much lower, and the interferents are different -- was undertaken as part of EPA-CDC interagency agreement 75937984. The development and testing of the method are reported in greater detail in the referenced peer-reviewed journal articles. A number of the references are available in full text on the Internet.

Cryptosporidiosis and *Cryptosporidium*

Cryptosporidiosis is a protozoan infection of humans, domestic animals and other vertebrates. In young farm animals, especially pre-weaned diary calves, it causes severe enteritis resulting in significant morbidity, mortality and economic loss. In humans *C. parvum* results in acute infection of the digestive system in immunocompetent individuals, and chronic, life-threatening disease in immunocompromised patients. Several transmission routes, including person-to-person, contamination of water or food, and zoonotic infection, are possible (Fayer et al., 1997).

Waterborne cryptosporidiosis outbreaks can occur when large numbers of pathogenic *Cryptosporidium* oocysts from the intestinal tracts of infected animals or humans are discharged into the environment, transported into the water supply and through water treatment processes and the distribution system, then out the tap in concentrations exceeding the infectious dose. The infectious dose may be as low as 10 oocysts for some strains for healthy individuals, and presumably less for immunocompromised persons. Oocysts may remain viable for months in the environment. They may be removed by filtration and are susceptible to ozone and ultraviolet light treatment, but their oocyst wall helps them survive routine chlorine disinfection. Therefore, water systems that chlorinate, but do not filter, must be particularly cautious about monitoring for oocysts and preventing their entry into source waters. A waterborne cryptosporidiosis outbreak in Milwaukee, WI in 1993 resulted in approximately 403,000 illnesses (Working Group on Waterborne Cryptosporidiosis, 1997). Numerous other cryptosporidiosis outbreaks in the U.S. and other countries have occurred.

Benefits of Identifying Host Range of *Cryptosporidium* Oocysts in Water

When *Cryptosporidium* oocysts are found at levels of concern in source or treated drinking water, officials responsible for watershed, utility, or public health management desire to quickly determine or confirm the source of contamination. An important piece of evidence is the oocyst's host range, which is the range of animal types that can be infected by the oocyst's *Cryptosporidium* species. While the host range does not directly identify the specific source or its location, it does enable a more focused and efficient investigation of the most likely sources. Not only does the host range indicate the potential upstream sources, but also the susceptible downstream hosts.

Cryptosporidium species that are adapted to only one or a small number of host animal types are valuable indicators of the specific animal source of these oocysts in water. Recently it has been discovered that *Cryptosporidium parvum*, which was thought to be extremely non-host-specific (i.e., it was thought to infect 79 different species of mammals (Fayer et al., 1997)) may in fact be a multi-species complex that contains many genotypes that are very host-specific. If so, detection of these C. *parvum* genotypes in water would help eliminate all other source animal types from further investigation, except those known to be suitable hosts. The C. *parvum* human genotype (now known as C. *hominis*) is one of the recently discovered host-specific genotypes. Host specificity data indicate that this genotype, which is difficult to distinguish from other C. *parvum* genotypes based on morphology, is almost exclusively infective for humans and non-human primates. Therefore, when it is found in U.S. waters, then

with very few exceptions where non-human primates are potential contributors to the oocyst load, it can be concluded that the oocysts came from human sewage.

For those *Cryptosporidium* species that have very broad host specificity (i.e., they can infect many members of a taxonomic class, e.g., mammals or birds or reptiles), the benefits of host range determination are limited, but still may be useful for decisions that only require knowledge of the host at the class level (e.g., mammals or birds) and for the determination whether it belongs to a known human-pathogenic *Cryptosporidium* species.

Determining the Host Range of Oocysts in Water Samples

Determining the host range of the recovered oocysts is accomplished by (1) characterizing the oocyst, (2) assigning the oocyst to its appropriate species or genotype, and (3) consulting available *Cryptosporidium*-host infection data to determine the potential host animals for the *Cryptosporidium* species or genotype that was found. Before the host range of oocysts recovered in a water sample can be determined, three requirements must be met: (1) development of a collection of infection data for the pairs of hosts and *Cryptosporidium* species or genotypes of interest in the watershed; (2) development of methods capable of detailed characterization of oocysts, if necessary to the molecular level, to enable reliable identification and differentiation of oocyst types; and, (3) discovery of correlations between measured characteristics of the oocysts' and their host ranges.

These three requirements are interdependent. The number of host-*Cryptosporidium* pairs in the infection database is influenced by the number of *Cryptosporidium* species and genotypes as well as the number of hosts and their immunocompetence categories (e.g. newborns, children, adults, elderly, AIDs, chemotherapy patients). The *Cryptosporidium* species and genotype categories are affected by the capabilities and limitations of methods for measuring oocyst characteristics. The oocyst characteristics are only useful for predicting host range if a reliable correlation exists between the chosen characteristics and the host range of the oocyst. A brief discussion of the individual requirements follows.

Developing a Collection of Infection Data for Cryptosporidium-Host Pairs

The host range of a particular *Cryptosporidium* species or genotype can be determined by experimental infections of host animals where feasible. Identification of the *Cryptosporidium* species or genotype in naturally infected animals also proves they are in the host range, but absence of *Cryptosporidium* in a host in a natural setting does not confirm non-infectivity, since exposure is not strictly controlled. A considerable amount of host specificity data has been collected. However, as *Cryptosporidium* species and genotype categories are split or joined together, it will be necessary to re-assess previous conclusions about the host range of *Cryptosporidium* species and genotypes. Undertaking human subjects testing is particularly rigorous, lengthy, and costly. The human infection studies cannot be done if there is excessive risk to the participants, which would be the case for immunosuppressed persons. The ideal, complete host specificity database would be a matrix with all of the relevant *Cryptosporidium* species and genotypes in the row headings, all the host animals with their relevant immunocompetence levels in the column headings, and results of infection studies in each corresponding cell. Generating all the data required to populate the ideal database is probably not attainable, and the recent discoveries of new *Cryptosporidium* species and the tentative identification by molecular tools of host-adapted genotypes indicates that the *Cryptosporidium* species designations will remain in flux for some time. However, only a relatively small portion of the ideal database is required for any particular watershed or situation, so the inability to fully populate the database is not a critical problem. One type of infection data that is lacking is the characterization of *Cryptosporidium* species and genotypes that infect wild animals.

Methods for Detailed Characterization of Oocysts

The second prerequisite for determining the host range of an oocyst in a water sample is the existence of methods to determine *Cryptosporidium* species and genotype directly from the oocyst. This includes characterizing the species based on the small number of oocysts likely to be present in the sample, eliminating or overcoming interferences that may be present, and detecting unique features that define the *Cryptosporidium* species, including its host specificity.

Conventional Detection and Differentiation of *Cryptosporidium* Oocysts

The methods commonly used to detect and differentiate *Cryptosporidium* oocysts often do not enable correlation of oocyst characteristics with its host range. Currently the identification of *Cryptosporidium* oocysts in water samples is largely made by the use of immunofluorescent assay (IFA) after concentration processes (ICR method, EPA method 1622/1623, flow cytometric method, solid-phase cytometric method, etc.) (Lindquist, et al., 2001) (Xiao et al., 2002a). Because IFA detects oocysts from most *Cryptosporidium* parasites, the species distribution of *Cryptosporidium* parasites in water samples cannot be assessed. In addition, diagnosis of *Cryptosporidium* parasites to the species specific level is a challenge because many of the *Cryptosporidium* species are morphologically similar. For example, it is very difficult for an experienced parasitologist to differentiate C. *parvum*, C. *wrairi*, C. *meleagridis*, C. *felis*, C. *canis* , and C. *saurophilum* under a microscope. Morphometric measurements are also needed for the differentiation of C. *muris* and C. *andersoni* from C. *parvum* and C. *parvum*-related species, which can be problematic with water samples that normally have only a few oocysts. As a result, conventional diagnostic practices rely on the presumed host specificity of *Cryptosporidium* parasites in addition to morphology.

Molecular Tools for Detection and Differentiation of *Cryptosporidium* Oocysts

The use of molecular tools enables differentiation between oocysts that was not previously possible. The term "molecular tools" is used here to refer to the procedures that are used, separately or in combination, to characterize DNA sequences of selected portions of an organism's genome. The *Cryptosporidium* genome contains approximately 10 to 20 million base pairs(bp) of DNA (Jenkins and Petersen, 1997). This project focused on detection and differentiation of *Cryptosporidium* species and genotypes based on polymorphism in the SSU rRNA gene, which is approximately 1733 to 1750 bp in length. Of particular relevance to this project are Polymerase chain reaction (PCR) procedures, restriction fragment length polymorphism (RFLP) procedures, and DNA sequencing. PCR rapidly generates thousands to billions of copies of targeted DNA sequences for use in other procedures (e.g., DNA sequencing and RFLP). In addition to copying target gene sequences for use in other methods, PCR can also be used to detect the presence/absence of particular organisms by copying/not copying their DNA. DNA sequencing determines the exact order of the nucleotides in the DNA molecule. The RFLP procedure is a faster, less costly, and less detailed approach to characterizing DNA. The RFLP method cuts a selected segment of DNA into fragments using restriction enzymes. The differing DNA fragment lengths are separated by electrophoresis and visualized by staining procedures. If a unique number and length of fragments are formed for a particular species and genotype then it can be identified by this method. Additional description of molecular tools is in Appendix 1.

Discovery of Correlations Between Characteristics of Oocysts and Their Host Ranges

The third prerequisite for determining the host range of an oocyst in a water sample is the discovery of correlations between measured characteristics of the oocysts and their host ranges. Previous attempts to establish a fully reliable correlation have been unsuccessful. Numerous instances have occurred where presumably identical *Cryptosporidium* species differed significantly in host specificity. These past failures occurred in large part due to the inability to characterize the oocysts in sufficient detail to detect all differences relevant to host specificity. As indicated above, molecular tools now enable the DNA sequences of oocysts to be characterized. The rationale for using molecular tools to characterize oocyst DNA in order to determine its host animal range is that the oocyst is pre-disposed, by the structural proteins and enzymes that are coded for in its DNA, to survive and reproduce in a limited range of host environments. These host-adapted *Cryptosporidium* may develop over long periods of time through co-evolution with the host animal. If *Cryptosporidium* species or genotypes differ in host specificity, then DNA differences should occur between these species or genotypes at one or more gene loci. The challenge is to find one or more genes where the presence of a particular set of DNA sequences is unique to all *Cryptosporidium* oocysts that have the same host range. In this and other projects a number of different gene loci have been investigated for the presence of these host-range-indicating sequences. The gene that codes for the small subunit ribosomal RNA has been found to be a particularly promising site, since it contains DNA sequences that appear to be unique to the *Cryptosporidium* genus, to particular *Cryptosporidium* species, and to host-adapted species and genotypes.

PROJECT RATIONALE, OBJECTIVES, AND TASKS

The development and use of molecular tools for genetic analysis of *Cryptosporidium* oocysts is relatively new, as is the documentation of the host-specific behavior of the various genotypes that have been (or will be) discovered. Even less developed is the application of the new molecular tools to the investigation of the distribution of *Cryptosporidium* species and genotypes in water samples, which is a much different matrix than clinical samples. Specifically, water samples have low concentrations of oocysts, different interferents, and probably a wider range of oocyst species and strains from not only humans, but also farm animals, companion animals, and wild animals.

Recognizing: (1) the investigative value of determining, directly from the oocyst, the host animals in which waterborne oocysts were produced; (2) the promise of molecular tools for characterizing *Cryptosporidium* species and genotypes of waterborne oocysts, and (3) the need for further development and evaluation of molecular tools before they can be used for oocyst source animal determination, an interagency agreement was initiated between the U.S. EPA/NRMRL and HHS/CDC. The ultimate objective of the project was to improve the techniques available to investigate and prevent waterborne cryptosporidiosis. The particular focus of the interagency agreement was to determine whether PCR-restriction enzyme digestion and sequencing assays developed by CDC (and partially funded under EPA/Office of Water - HHS/CDC interagency agreement 75937730), which had been successfully used on clinical samples, could also be applied to water samples and be incorporated into investigative approaches for determining the sources of *Cryptosporidium* in water supplies.

Key tasks were: (1) confirmation that the genus *Cryptosporidium* is a multi-species complex by completion of a phylogenetic analysis based on characterization and comparison of the SSU rRNA genes of four types of *Cryptosporidium*; (2) development of an SSU rRNA-based nested PCR-RFLP method for detecting and differentiating all known species of *Cryptosporidium* and multiple C. *parvum* genotypes; (3) testing the SSU rRNA-based nested PCR-RFLP method on multiple samples from (a) hemolymph and gill washings from oysters, (b) storm water, (c) surface water, and (d) wastewater to determine its ability to detect and differentiate oocysts and to determine whether the oocyst types matched the probable source animal types; and (4) evaluate and briefly summarize the promise and the challenges for molecular tools for detection of *Cryptosporidium* oocysts in water.

MATERIALS AND METHODS

The materials and methods used for the phylogenetic analysis of the *Cryptosporidium* genus and the development and evaluation of the SSU rRNA-based nested PCR-RFLP method for identification and differentiation of *Cryptosporidium* species and genotypes are described in detail in the referenced articles (Xiao et al., 1998, 1999 a and b, 2000, 2001, and 2002a; and, Sulaiman et al., 1999). In a similar or identical manner the method was also applied to gill washings and hemolymph from oysters, surface water from several states, and raw wastewater. Other gene loci (e.g., beta-tubulin, actin, 70kDa heat shock protein, and thrombospondin anonymous protein genes) were also investigated (Sulaiman et al., 1998 and 1999a, 2000, and 2002) under the interagency agreement and other projects for species detection and differentiation, but these efforts are not discussed further here. The review of the SSU rRNA method and other molecular methods for oocyst DNA characterization was completed based on review of the literature and summarization of relevant research by the review authors.

Materials and Methods for Phylogenetic Analysis of *Cryptosporidium* Genus Based on SSU rRNA Genes

The SSU rRNA gene was selected for the phylogenetic analysis for several reasons (Xiao et al., 1999a and b). Ample sequence data are available for this gene since it has been extensively studied because it is present in all eukaryotic organisms and it plays an important role in protein synthesis. The gene has conserved regions, which may contain genus/species specific detection sites. The gene also has polymorphic regions, which may contain species/genotype differentiation sites. There are five copies of the rRNA gene per sporozoite, which increases sensitivity of detection by PCR.

Isolates of *Cryptosporidium* parasites for SSU rRNA gene sequencing were obtained from humans, cattle, calf, snakes, lizards, a guinea pig, a camel, a hyrax, a chicken, a rhesus monkey, a ferret, a pig, a dog, a kangaroo, a turkey, and a cat. Partial sequences covering the most polymorphic regions were obtained from C. parvum human genotype isolates, bovine isolates, and one C. *baileyi* isolate. Oocyst type determination was by morphology and/or host specificity. DNA extraction was extracted by freeze-thaw and adsorption procedures.

The full-length SSU rRNA gene was amplified from each sample by conventional PCR by using forward and reverse primers with 25 nucleotides each. The SSU rRNA gene is 1733 to 1750 bp long, depending on species and genotype. Each PCR consisted of 35 cycles of denaturation at 94/C for 45 s, annealing at 60/C for 45 s, and extension at 72/C for 60 s; an initial denaturation step consisting of incubation at 94/C for 5 min and a final extension step consisting of incubation at 72/C for 10 min were also included. The copied DNA segments were sequenced with an ABI377 autosequencer (Perkin Elmer, Foster City, Calif.)

The SSU rRNA sequences of *Cryptosporidium* from this study were aligned and compared to other apicomplexan parasites by a neighbor joining (NJ) tree analysis to assess the genus status of *Cryptosporidium*. The SSU rRNA sequences of *Cryptosporidium* species from this study were also aligned and compared to each other by a neighbor joining tree analysis. The analysis checked for evolutionary distances indicating separate species, clustering within species, and clustering between species with similar host specificity, and similarities in locations of mutations.

Development Process for SSU rRNA-based Nested PCR-RFLP Method for *Cryptosporidium* Detection and Differentiation

The conceptual design of the SSU rRNA-based nested PCR-RFLP method is illustrated in Figure 1. The SSU rRNA gene was selected as the target gene for the reasons previously described. DNA sequences from four *Cryptosporidium* species (i.e., C. *parvum*, C. *serpentis*, C. *muris*, and C. *baileyi)*, which were obtained during the previously described phylogenetic analysis, were used in the initial method development process (Xiao et al., 1999a).

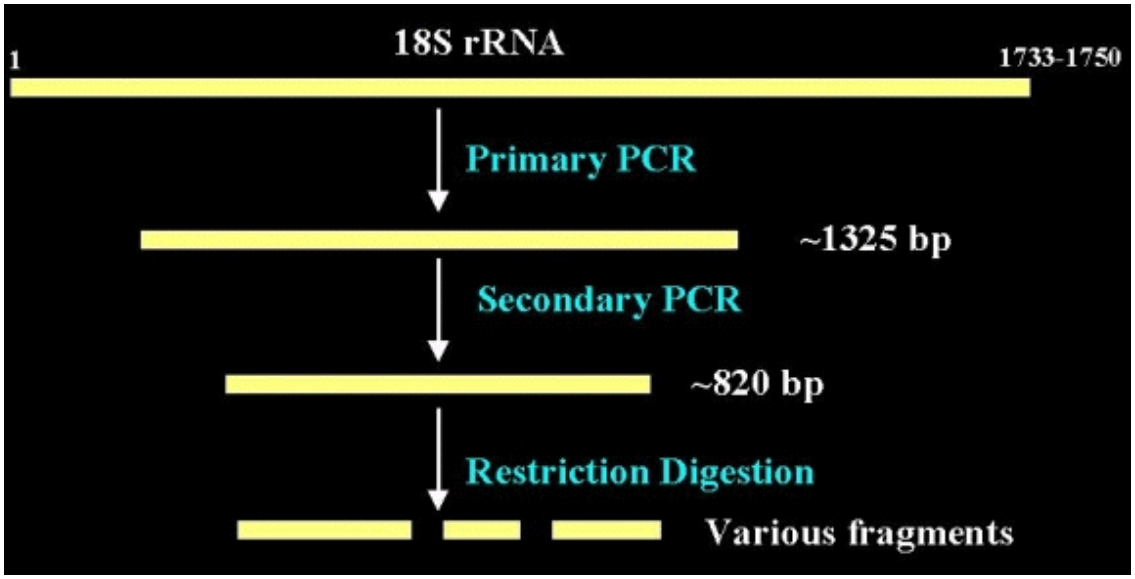

Figure 1. Detection and Diagnosis of *Cryptosporidium* Parasites by Nested PCR-RFLP

The DNA sequences were aligned and examined for restriction sites (i.e. , sites at which available restriction enzymes will cut the DNA). SspI restriction enzyme was selected for species diagnosis. VspI restriction enzyme was selected for genotyping of C. *parvum*. A DNA sequence (~ 1325 bp) in the SSU rRNA gene was identified for amplification by primary PCR and within that sequence an ~820 bp target segment was selected for secondary PCR. The forward and reverse primers (~ 20 bp in length) selected for primary and secondary PCR are: (1) common to the *Cryptosporidium* genus, but not present in the DNA of other microorganisms, and (2) bracket the target region (~ 820 bp) of the SSU rRNA gene that contains unique sequences that enable species and genotype differentiation by RFLP and/or DNA sequencing. Nested PCR was chosen to maximize sensitivity of detection. Nested PCR first amplifies the larger (~1325 bp) segment, and then, starting with some of the primary PCR product, amplifies the smaller (~ 820 bp) segment, which improves the overall efficiency of target DNA amplification compared to single-round PCR. Primer selectivity was confirmed by detection (i.e., PCR amplification) when *Cryptosporidium* DNA (i.e., C. *parvum* bovine and human genotypes; C. *muris*, C. *serpentis*, and C. *baileyi*) were present and non-detection (i.e., no PCR amplification) when non-*Cryptosporidium* parasite DNA (i.e., *Eimeria* and *Giardia)* was present. DNA from *Cryptosporidium* species and genotypes were also digested with SspI and VspI restriction enzymes to determine whether they produced the predicted and unique restriction fragment patterns upon electrophoresis and visualization by ethidium bromide. Sensitivity of the SSU PCR-RFLP method was confirmed by testing on serial dilutions of DNA to an equivalent of one oocyst.

For *Cryptosporidium* species and strains not detected or differentiated by this PCR-RFLP approach there are several options to address the problem. *Cryptosporidium* oocyst types not detected indicates sequence differences in primer regions, and this problem may be addressed by different primers to detect those *Cryptosporidium* species or genotypes. No examples of non-detectable *Cryptosporidium* species have been found to date. If *Cryptosporidium* oocyst types are detected, but not differentiated, then these oocyst types may be differentiated by (1) use of different restriction enzymes or (2) direct sequencing. *Cryptosporidium* oocyst types that are known to be different (e.g. morphology or host specificity), but do not have SSU rRNA sequence differences will not be differentiable by RFLP at the SSU rRNA gene locus, but may be detectable and differentiable by PCR-RFLP at other gene loci. No examples of this situation have been found to date.

Evaluation of SSU rRNA-based Nested PCR-RFLP Method for *Cryptosporidium* Detection and Differentiation in Storm Water Samples

The key animal sources of oocysts were identified by the environmental setting and general knowledge of the sampling sites (Xiao et al., 2000). Grab or composite samples were taken during storm flows. The Information Collection Rule (ICR) method was used for storm water sample collection/processing. (In subsequent evaluations EPA Method 1623 was used for surface water sample collection and processing, and centrifugation of grab samples was employed for wastewater processing). Oocysts in water samples that were concentrated by filtration, Percoll-sucrose floatation or centrifugation were further purified by immunomagnetic separation (IMS). Direct DNA extraction without IMS interfered with PCR because of the presence of PCR inhibitors. IMS-purified oocysts were subjected to 5 freeze-thaw cycles, incubated with 1 mg/ml of proteinase K at 56 °C for at least 1 h, and diluted with equal volume of pure ethanol. Oocyst DNA was extracted by passing the oocyst-ethanol suspension through QIAamp DNA Mini isolation columns.

For the PCR-RFLP analysis a primary PCR product of about 1,325 bp was amplified. Thirty-five (35) replication cycles were completed at about 2.5 minutes per cycle. Secondary PCR product of 826-864 bp (depending on isolates) was then amplified from 2 Fl of the primary PCR reaction, using different primers. For restriction fragment analysis, 20 Fl of the secondary PCR product was digested in a total of 50 Fl reaction mix, consisting of 20 units of *Ssp* I or *Vsp* I and 5 Fl of respective restriction buffer at 37 EC for 1 hr. The digested products were fractionated on 2.0% agarose gel and visualized by ethidium bromide staining and/or characterized by DNA sequencing. A modified procedure was employed for multiple species in a single sample. Each sample was analyzed at least 3 times by PCR-RFLP, using different volumes of DNA preparation (0.25, 0.5, and 1 µl) for PCR. Where multiple species occur in a single sample, then multiple additional bands occur after electrophoresis, which causes difficulty in interpreting the results. Since one oocyst type usually predominates in a sample, dilution of DNA prior to PCR and multiple PCR assays can increase detection of the non-dominant type. If suitable separation cannot be obtained, then the fragments can be cloned and multiple clones sequenced. Both approaches increase time and expense.

For confirmation, the secondary PCR products were sequenced using an ABI377 autosequencer. Nucleotide sequences generated were aligned with each other and with known *Cryptosporidium* species and *C. parvum* genotypes previously obtained, using computer software Wisconsin Package Version 9.0 (Genetics Computer Group, Wisconsin) and manual adjustment. Phylogenetic analysis (i.e. construction of evolutionary tree based on gene sequence similarity) was performed on the aligned sequences to assess relationship among isolates.

KEY RESULTS

Results of Phylogenetic Analysis of *Cryptosporidium* Genus Based on SSU rRNA Genes of Five Types of *Cryptosporidium*

Although biological data support the hypothesis that there are multiple species in the genus *Cryptosporidium*, a previous analysis of the available genetic data suggested that there is insufficient evidence for species differentiation, and hence that it would be infeasible to determine animal source types by genetic characterization. However, this study (Xiao et al., 1999a) revealed that the genus *Cryptosporidium* is distinct from other apicomplexan parasites. Also, *Cryptosporidium* contains phylogenetically distinct species such as *C. parvum, C. muris, C. baileyi , and C. serpentis*, which is consistent with the biological characteristics and host specificity data. The *Cryptosporidium* species formed two clades (i.e., groups), with C. *parvum* and C. *baileyi* belonging to one clade and C. *muris* and C. *serpentis* belonging to the other clade (Figure 2). Another study (Xiao et al., 1999b) extended the phylogenetic analysis to include *C. felis*, *C. meleagridis*, and some additional host-adapted *C. parvum* genotypes (dog, pig, kangaroo, ferret, mouse, and monkey). Subsequent SSU rRNA gene sequencing and phylogenetic analyses have confirmed the groupings described above and produced a more detailed characterization of the phylogenetic relationship of *Cryptosporidium* parasites as shown in Figure 3 (Xiao et al., 2002b).

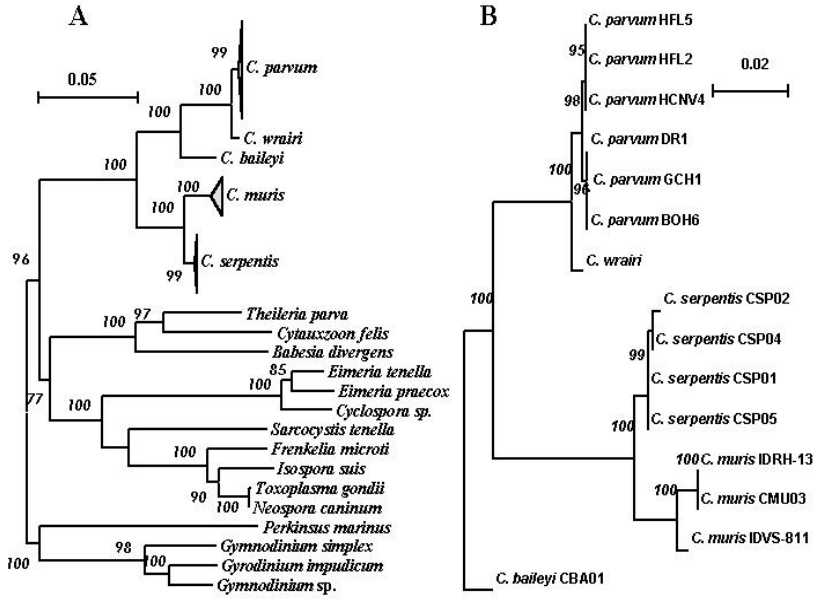

Figure 2. Phylogenetic Relationships of Cryptosporidium Parasites to Other Apicomplexans(A) and Each Other(B) (Xiao et al., 1999a)

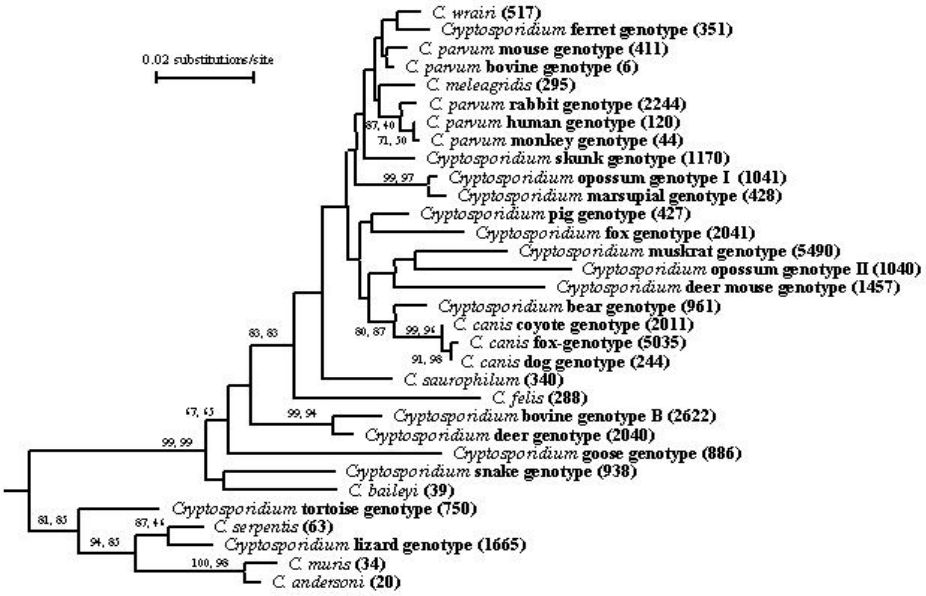

Figure 3. Updated phylogenetic relationship of Cryptosporidium parasites inferred by the neighbor-joining analysis of the SSU rRNA gene based on genetic distances calculated by the Kimura two-parameter model. The tree was rooted with an SSU rRNA sequence from *Eimeria tenella*, and the root was removed to show the details of the relationship among Cryptosporidium parasites. Bootstrap values (in percentage) above 50 from 1,000 pseudo-replicates are shown for both the neighbor-joining (the first value) and maximum parsimony analyses (the second value). Figures in parentheses are isolates' designation numbers. (Xiao et al., 2002b)

Results of Development of SSU rRNA Nested PCR-RFLP Diagnostic Tool

Based on the genetic information obtained and analyzed during the phylogenetic analysis, a species- and strain-specific PCR-RFLP diagnostic tool was developed. The diagnostic tool reveals the presence of *Cryptosporidium* oocysts by amplification of the primary and secondary PCR target DNA sequences. If the target sequences are not produced, then *Cryptosporidium* DNA are not present in the sample. Figure 4 illustrates the positive response for 5 types of *Cryptosporidium* oocysts and absence of target DNA replication when two non-*Cryptosporidium* parasites' DNA are tested (Xiao et al., 1999a). In the RFLP procedure, digestion of the secondary PCR products with SspI and VspI restriction enzymes produces unique patterns that often enable differentiation of *Cryptosporidium* species and *C. parvum* genotypes. Differentiation of some species and genotypes may require either a third digestion with a different restriction enzyme or direct sequencing (Xiao et al., 1999b; Xiao et al., 2001).

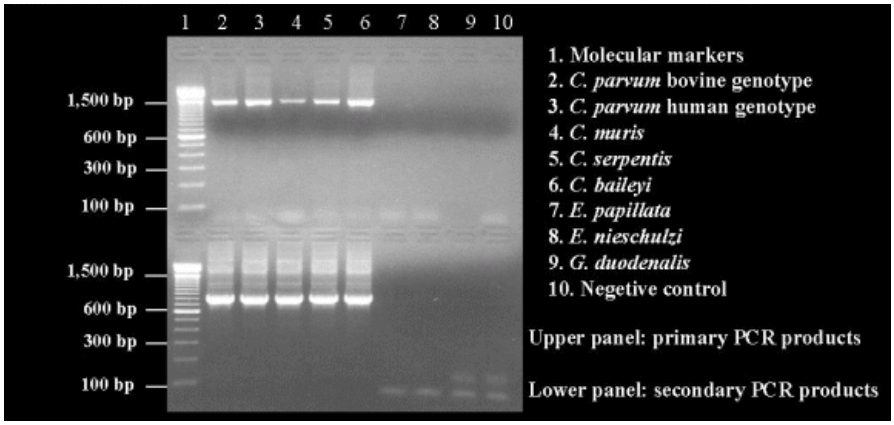

Figure 4. Detection of *Cryptosporidium* spp. by SSU rRNA-based Nested PCR

Figure 5 shows the results of the SSU rRNA-based PCR-RFLP procedure on 14 types of *Cryptosporidium* oocyst DNA. There were seven non-C. parvum species and seven C. parvum genotypes. Eight of the types, comprised of four species (i.e., *C. muris, C. serpentis, C. baileyi, and C. felis*) and four *C. parvum* genotypes (i.e., pig, marsupial, bovine, and human) produced unique restriction fragment patterns that enabled them to be differentiated. The remaining six types fell into two groups that could be differentiated from the first eight, but not from each other, by electrophoresis and ethidium bromide staining. These remaining types could be differentiated by DNA sequencing of the secondary PCR fragments. One group contained *C. meleagridis*, *C. parvum* ferret, and *C. parvum* mouse and the other group was *C. parvum* dog, C. lizard sp., and *C. wrairi*.

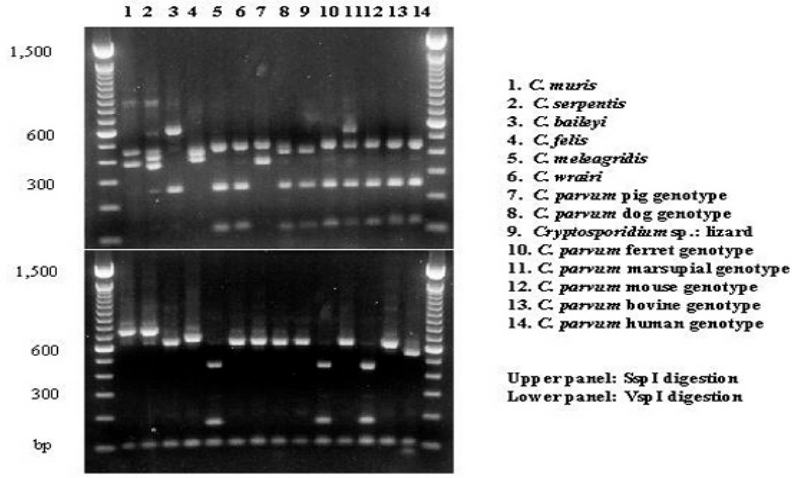

Figure 5. Differentiation of *Cryptosporidium* Species and Genotypes by SSU rRNA-based PCR-RFLP

The SSU-rRNA nested PCR-RFLP method was demonstrated to detect a single oocyst in laboratory dilution tests (Figure 6). In one field test it indicated presence of oocysts with slightly higher frequency than microscopic examination (Sulaiman et al., 1999b).

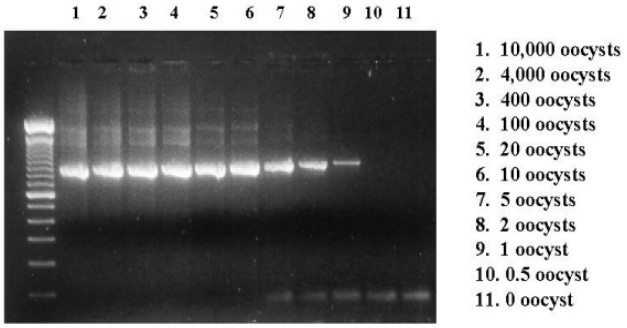

Figure 6. Sensitivity of the SSU rRNA-based *Cryptosporidium* PCR-RFLP Genotyping Technique

Initially oocysts concentrated by the ICR procedure could not be amplified by PCR, but it was then found that the immunomagnetic separation (IMS) technique utilized in the newer protozoa concentration/separation methods (EPA 1622 and 1623) successfully removed interferents and produced oocysts that can be amplified by PCR. An exhaustive examination of the effects of a wide range of interferents on detection and differentiation of oocysts was not within the scope of the project, although it was shown that IMS-PCR detects and differentiates oocysts in surface water, storm water, and wastewater (Xiao et al., 2000).

Discrimination of different genotypes within a single sample was done by dilution and multiple PCR-RFLP analyses. If discrimination is not possible by PCR-RFLP, then it may be possible by DNA sequencing (Xiao et al., 2000 and 2001).

Evaluation of the SSU rRNA-based Nested PCR-RFLP Diagnostic Tool

The SSU rRNA-PCR-RFLP method was successfully used to differentiate *Cryptosporidium* species and *C. parvum* genotypes in gill washings and hemolymph from oysters, storm water, raw surface water, and wastewater.

Gill Washings and Hemolymph from Oysters

The diagnostic tool was used to analyze oocysts recovered from the hemolymph and gill washings of oysters collected from the Chesapeake Bay. Oysters are filter feeders that concentrate and accumulate oocysts from surface water, thus enabling the investigator to avoid these tasks. Sixty-five pooled oyster samples were analyzed. *Cryptosporidium* oocysts were present in 26 samples. Twenty-four of the samples contained *C. parvum* and each of the other samples contained *C. baileyi* (typically found in birds) and *C. serpentis* (typically found in snakes). Of the *C. parvum* positive samples, 22 of 24 were the bovine genotype (also known as genotype 2), which is typically from cattle, humans, and other ruminants. Two samples were positive for *C. parvum* human genotype (also known as genotype 1), which only circulates among humans. (Xiao et al., 1998)

Storm Stream Flow Samples

When the molecular tool was applied to water samples from storm stream flows in the New York City Watershed, 12 genotypes were found in 27 of 29 samples. Four of the 12 genotypes matched sequences from known *Cryptosporidium* parasites: C. *baileyi* (from birds), *Cryptosporidium* from snakes, and two *Cryptosporidium* genotypes from opossums. No genotypes found in the storm samples matched those from humans, farm animals, or companion animals (i.e., *C. felis*, *C. meleagridis*, *C. andersoni*, and the human, bovine, and pig genotypes of *C. parvum*), indicating that genotypes in storm water were probably from wildlife (Figure 7). This conclusion is consistent with the environmental setting of the sampling sites and the presence of the four genotypes with known

(wild) animal sources (Xiao et al., 2000). The presence of unknown genotypes indicates the existence of, and need to fill, gaps in the genotyping of *Cryptosporidium* oocysts in wildlife.

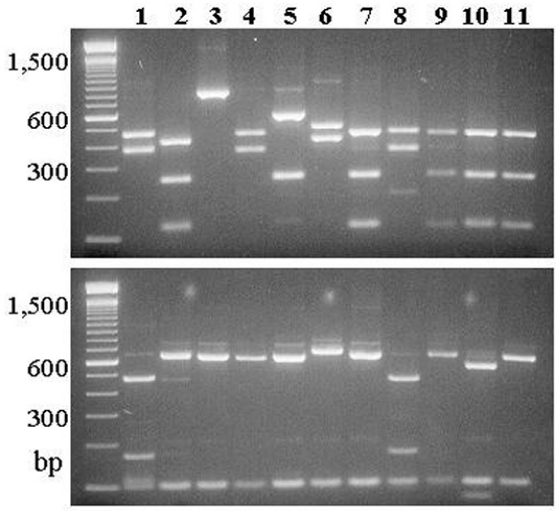

Figure 7. Differentiation of the *Cryptosporidium* Parasites in Storm Water Samples by SSU rRNA-based PCR-RFLP.
Lanes 1,2,4,8,10, and 11, unknown *Cryptosporidium* spp.;
lane 3, *Cryptosporidium* from snakes;
lane 5, *C. baileyi*;
lane 6, *Cryptosporidium* opossum genotype 2;
lanes 7 and 9, *C. parvum* bovine-like genotype.

Raw Surface Water Samples

When the molecular tool was applied to raw surface water samples that were collected from several states, 25 of the 55 surface water samples were positive for *Cryptosporidium*. The species and/or genotypes found in raw surface water were: *C. parvum* human and bovine genotypes, *C. baileyi, and C. andersoni*. *C. parvum* (both human and bovine genotypes) was the predominant (21 of 25 samples) *Cryptosporidium* species found in surface water. *C. andersoni*, which occurs in juvenile and adult cattle, was found at a moderate frequency (5 samples). Many surface water samples, particularly those from the Chesapeake Bay sampling sites, contained more than one genotype. The distribution of host-adapted *Cryptosporidium* species detected was consistent with the observed potential sources (i.e., cattle farm runoff or wastewater discharges) of water contamination (Xiao et al., 2001).

Raw Wastewater Samples

When the molecular tool was applied to raw wastewater samples from a wastewater treatment plant in Milwaukee, WI, 12 of the 49 samples were positive for *Cryptosporidium*. *C. parvum* human, bovine, and dog genotypes (now known as C. *canis* (Fayer, et al., 2001)); *C. felis*, *C. andersoni*, *C. muris*, and an unknown genotype (now known to be *Cryptosporidium* deer genotype) were found in the wastewater samples. *C. andersoni* was found at the highest frequency in the wastewater samples, which is consistent with (1) known occurrence of *C. andersoni* in juvenile and adult cattle and (2) the presence of a large slaughterhouse (1800 beef cattle per day) that discharged pre-treated effluent into the sewer system several miles upstream of the treatment plant. The appearance of C. *muris*, probably from rodents, and C. *parvum* dog genotype is also consistent with potential fecal microorganism contributors to an urban sewer system. The low observed frequency of C. *parvum* human genotype was surprising, but may be explained by the sampling period occurring in April to July, when the incidence of human cryptosporidiosis tends to be low (Xiao et al., 2001).

Comparison of PCR Protocols for Species Detection, Differentiation, and Genotyping of *Cryptosporidium*

In 1999 the specificity and sensitivity of 11 PCR protocols were evaluated for species detection, differentiation and genotyping of *Cryptosporidium* parasites in clinical samples. Although many of the protocols performed well in their particular niches, the SSU rRNA nested PCR offered a wider range of detection, differentiation, and genotyping capability and better sensitivity than the other tools. Ten protocols amplified C. *parvum* genotypes 1(human) and 2(bovine), and the expected fragment sizes were obtained. Two species-

differentiating protocols were not *Cryptosporidium* specific, as the primers used in these protocols also amplified the DNA of *Eimeria* species. Six *C. parvum* genotyping protocols that were based on six different gene loci were limited to *C. parvum*, since the primers used in these protocols amplified only the DNA of genotype 1 and/or genotype 2 isolates of C. *parvum*, but not the DNA of non-*C.parvum* oocysts. Sensitivity studies revealed that two nested PCR-RFLP protocols – the one based on the SSU rRNA gene described in this summary and the other on the dihydrofolate reductase genes – are more sensitive than single-round PCR or PCR-RFLP protocols (Sulaiman et al., 1999b). A literature review of molecular detection of *Cryptosporidium* oocysts in water was also completed in 2002 (Xiao et al., 2002a). Twelve protocols were identified. These methods include PCR, nested PCR, IMS-PCR, IMS-nested PCR, IMS-nested PCR-RFLP, Reverse Transcription (RT)-PCR, and cell culture (CC)-PCR. Other method differences include the target genes (e.g., SSU rRNA, Oocyst wall protein, HSP 70, undefined sequences, and TRAP-C2) and target gene regionis, the DNA extraction method, and the amount of testing to date on environmental samples. Again, the SSU rRNA nested PCR methods, of which there now are several, offer a wider range of detection, differentiation, and genotyping capability, as well as improved sensitivity compared to the other tools.

CONCLUSIONS AND RECOMMENDATIONS

Specific Conclusions

A molecular method (SSU rRNA-based nested PCR-RFLP) has been developed and demonstrated to be capable of detecting and differentiating, in clinical or environmental samples, at least 10 species and 22 genotypes of *Cryptosporidium*. The species detected and differentiated to date by this method are: *C. andersoni*, C. *baileyi*, C. *canis*, C. *felis*, C. *meleagridis*, C. *muris*, C. *parvum*, *C. saurophilum*, C. *serpentis*, and C. *wrairi*. The *Cryptosporidium* genotypes detected and differentiated are: bovine (2 genotypes), ferret, human (C. hominis), pig, marsupial, rabbit, mouse, deer (2 genotypes), deer mouse, bear, skunk, opossum (2 genotypes), fox (two genotypes), muskrat, goose, snake, lizard, and tortoise.

The SSU rRNA-based nested PCR-RFLP method has been used to show the presence of five (5) species or genotypes of *Cryptosporidium* that have been found in human patients. The confirmed human-infective species and genotypes are: *C. parvum* human genotype, *C. parvum* bovine genotype, *C. felis*, *C. canis*, and *C. meleagridis*. Other *Cryptosporidium* parasites such as *C. andersoni*, *C. muris*, *Cryptosporidium* cervine and pig genotypes have also been found in humans, but much less frequently than the 5 common *Cryptosporidium* parasites.

The SSU rRNA-based nested PCR-RFLP method can help prevent over-estimation of the human-pathogenic potential of oocysts found in water samples by enabling oocysts to be grouped, depending on available genetic and host-specificity data as: (1) known human pathogenic, (2) suspected human pathogenic, and (3) known non-human pathogenic. This capability is of particular value where the oocysts are difficult to distinguish by typical microscopic techniques, and the conservative assumption that all oocysts are human pathogenic is much costlier than a more accurate assessment.

For *Cryptosporidium* species and genotypes that are known with some confidence to strictly or primarily infect particular animal species, then it is possible to determine the source animal species based on DNA characterization. Based on available host-specificity information, plus the ability of the SSU rRNA-based nested PCR-RFLP method to detect and differentiate these species and genotypes, the following *Cryptosporidium* - host animal species pairs can be determined from DNA characterization of the oocyst: C. *parvum* human genotype - humans and non-human primates; *C. wrairi* - guinea pigs; *C. felis* - primarily cats; *C. andersoni* - primarily juvenile and adult cattle; and, C. canis - primarily dogs.

For *Cryptosporidium* species and genotypes that are known to strictly or primarily infect a particular animal class, then it is possible to determine the source animal class based on DNA characterization. Based on available host-specificity information, plus the ability of the SSU rRNA-based nested PCR-RFLP method to detect and differentiate these *Cryptosporidium* species and genotypes, the following crypto-host animal class pairs can be determined from DNA characterization of the oocyst: C. *baileyi* - birds; C. *meleagridis* - birds; C. *parvum* bovine genotype - mammals (including humans); C. *saurophilum* - reptiles (lizards); and C. *serpentis* - reptiles (snakes). Exceptions are possible for C. *meleagridis*, which has been found in humans in a limited number of cases.

Although their host specificity is not yet well-documented, the following *Cryptosporidium* genotypes have been detected, some for only a limited number of times, in the same hosts. If one accepts that these genotypes are tentatively host-specific, then the following additional source animal species can be tentatively identified by the SSU rRNA-based nested PCR-RFLP method or DNA sequencing of the oocysts: C. *parvum* mouse genotype - mice; *Cryptosporidium* ferret genotype -ferrets; *Cryptosporidium* fox genotype - foxes; C. *parvum* monkey genotype - monkeys; *Cryptosporidium* skunk genotype - skunks; *Cryptosporidium* opossum genotypes - opossums; *Cryptosporidium* deer mouse genotype - deer mice; *Cryptosporidium* deer genotype - deer; *Cryptosporidium* goose genotype - geese; *Cryptosporidium* bear genotype - bear; C. canis fox genotype - foxes; C. canis coyote genotype - coyotes; C. parvum pig genotype - pigs; *Cryptosporidium* muskrat genotype - muskrats.

The existence of *Cryptosporidium* species and genotypes that have limited host-specificity makes it impossible, when they are present in a sample, to identify a unique source-animal species based on DNA characterization alone. Cross-transmission studies have demonstrated that there are non-host specific *Cryptosporidium* species and genotypes. Specifically, it has been demonstrated at the genotype level by the C. *parvum* bovine genotype, which has been found in both infected humans and infected cattle.

From a public health standpoint, it would be very useful to be able to identify from the oocysts the following four sources of mammalian oocysts: humans, cattle, wildlife, and pets. Based on data collected to date, the current capability of making the distinctions between several key pairs of potential sources (human vs. cattle, human vs. wildlife, human vs. pets, and cattle vs. wildlife) is as follows:

- **Human vs. Cattle**
  - If C. *parvum* human genotype is found, then humans are the likely sources of these oocysts.
  - If C. *parvum* bovine genotype is found, then the source of these oocysts could be either humans, cattle, sheep or perhaps some other mammals as well. Subgenotype analysis can be useful because some bovine genotype isolates have only been found in humans.
  - If C. *andersoni* is found, this could indicate that the source is juvenile or adult cattle. C. *andersoni* is very rarely found in humans.

- **Humans vs. Wildlife**
  - If the *C. parvum* human genotype is found, then only humans or non-human primates are known sources of these oocysts.
  - If C. *parvum* bovine genotype is found, then the source could be either humans or cattle, or perhaps other mammals, including wildlife.
  - If the tentative wildlife host-specificity of several *Cryptosporidium* genotypes is accepted, then oocysts can be tentatively linked to specific wildlife species (e.g., deer, ferret, fox, mouse, opossum, raccoon, bear, muskrats, birds, reptiles).
  - Numerous wildlife *Cryptosporidium* types have not yet been characterized.

- **Human vs. Pets (Cats or Dogs)**
  - If the C. *parvum* human genotype is found, then only humans or non-human primates are the known source of these oocysts.
  - If the C. *parvum* bovine genotype is found, the source of these oocysts could be humans or pets.
  - If C. *felis* and C. *canis* are confirmed host-specific to cats and dogs respectively, then it will be possible to link the oocysts to these pets, although not directly determine if they are domesticated or feral, which may be apparent from the watershed.
  - Since C. *felis* and C. *canis* have been found in clinical samples from small numbers of humans, this potential source would need to be investigated as well.

- **Cattle vs. Wildlife**
  - If C. *parvum* bovine genotype is found, then the source could be either humans or cattle, or perhaps other mammals as well.
  - If C. *andersoni* is found, this could indicate that the source is juvenile or adult cattle, but this may change if C. *andersoni* is found to be human pathogenic.
  - If several *Cryptosporidium* genotypes that have been tentatively identified as wildlife are host-specific, then oocysts can be tentatively be linked to specific wildlife species (e..g, deer, raccoon, opossum, ferret, fox).
  - Numerous wildlife *Cryptosporidium* oocyst DNA types are not yet characterized. Further characterization of wildlife samples are needed to pinpoint this source.

The feasibility of identifying source animal contributors of oocysts collected in water samples using the SSU rRNA-based nested PCR-RFLP method was successfully demonstrated with multiple samples from four different water matrices and combinations of source animals.

This method was not expected to and does not provide the number of oocysts of each species and genotype that is present in a water sample. Quantitative accuracy was not a capability expected or explored in the project, although method sensitivity (i.e., the ability to detect the presence of as few as 1 oocyst/sample) is important and was examined. Also, the total and species-specific recovery efficiency of the concentration and separation methods has a bearing on conclusions that can be drawn about oocyst distributions that are found. As recovery, separation, and concentration methods improve in speed, economy, and effectiveness, the feasibility and value of molecular tools will also improve.

This method was not expected to and does not provide information about the viability of the oocysts, nor does it provide routine on-line monitoring of *Cryptosporidium* oocyst contamination.

General Conclusions

The effective application of molecular methods to determination of source animal types from oocysts in water samples is dependent on the existence and understanding of the host-specificity of the current and growing number of *Cryptosporidium* genotypes in humans, cattle, pets, and wildlife.

- If more host-specific *Cryptosporidium* genotypes are identified, then it may be possible to identify additional host animals and human pathogenicity by oocyst DNA characterization, as was the case when, for examples: C. *parvum* was found to consist of a human-specific genotype (now a named species, C. hominis) and a non-specific bovine genotype; and C. *wrairi* was found to be host-specific for guinea pigs.
- Conversely, it may be found that some of the current species and genotypes considered to be host-specific are not, and therefore molecular techniques may have to be used in combination with conventional analytical techniques to determine host animals.
- Even if a method was available to determine all *Cryptosporidium* species and genotypes, this would not help determine the source (e.g., human vs cattle vs companion animals vs. wildlife) for *Cryptosporidium* oocysts from non-host-specific species and genotypes.

Assuming that the number of host-specific genotypes increases significantly, there will probably need to be a concurrent increase in the speed, specificity, and economy of methods capable of detecting and differentiating the oocyst DNA to determine the source animal type. There are currently efforts underway to address these needs (e.g., Limor et al., 2002; McDonald et al., 2002).

The discovery of host-specific genotypes and the availability of methods for species and genotype detection and differentiation indicates a possible need to re-evaluate *Cryptosporidium* characterization practices for feeding studies. For example, comparative evaluations of the infectivity and pathogenicity of the 5 human-pathogenic *Cryptosporidium* species and genotypes are needed.

At the present time the *Cryptosporidium* species and genotype detection and differentiation methods and the supporting host specificity data described in this document and elsewhere are best suited to generating, supporting, or refuting preliminary hypotheses or conclusions about the type of source animals (i.e., humans, farm animals, pets, wild animals) in which the oocysts were produced and their human-pathogenic potential. Preliminary hypotheses about source type or pathogenicity can be very important for orienting investigation or response resources to the most likely contamination sources. Also, preliminary conclusions based on other data about the potential source animals or pathogenicity can be checked against the molecular evidence. Sufficient resources are required for representative sampling and analysis; collection and sequencing of fecal samples may be required if *Cryptosporidium* genotype data are not available for animals known to be in the watershed; and adequate time and lab capacity must be available from the limited number of laboratories that perform the procedures and data analysis.

At the present time the *Cryptosporidium* species and genotype detection and differentiation methods and the supporting host specificity data described in this document are not well-suited for the following applications.

- Conclusively proving the source animal type of reputed host-specific oocysts found in water. The difficulty here stems from a lack of comprehensive *Cryptosporidium* host-specificity data and *Cryptosporidium* genotype characterization data.
- Routine water monitoring – the time, cost, equipment, and experience required to successfully perform these methods makes them inappropriate for routine water monitoring. A rough estimate of the present cost/sample is $ 400 for materials (including filters and IMS and sequencing) plus labor and equipment.
- Quantitative characterization of the distribution of *Cryptosporidium* oocyst genotypes in water. Even assuming that an assembly of samples are analyzed that are representative of the water body, the PCR-RFLP process involves steps that are not at present readily quantifiable. For example, the PCR process makes thousands to billions of copies of the target gene and the replication efficiency is not suitably defined to enable accurate calculation of the initial number of DNA molecules. Efforts are underway to develop real-time, quantitative PCR (Limor et al., 2002; McDonald et al., 2002)

It is considered unlikely at any time that *Cryptosporidium* species and genotype detection and differentiation methods and supporting host specificity data as described in this document or elsewhere will be suitable for:

- Identifying the particular animal species that produced non-host-specific *Cryptosporidium* genotypes
- Matching the oocysts in water to a specific individual animal.

Recommendations

While molecular detection and characterization of *Cryptosporidium* oocysts has made substantial progress and shows considerable promise, there are some current and future issues that should be addressed to enable and accelerate the use of molecular tools to their full potential for generating data that are useful in the risk assessment of various waters in different environmental settings, and for watershed management and source water protection (Xiao et al., 2002a).

Current problems in molecular detection of Cryptosporidium oocysts

- Only a limited number of tools for species differentiation, most of which are based on the small subunit rRNA gene
- Nonspecificity of some species differentiation tools
- Misinterpretation of data because of outdated knowledge of the evolving research field
- Existence of erroneous data in the database and publications
- Lack of laboratory and field data on host specificity of *Cryptosporidium* species and genotypes.

Actions needed to enable routine use of molecular tools in water sample analysis

- Rigorous standardization and testing have yet to be carried out in order to develop quality assurance and quality control procedures
- Development of protocols that allow the extraction of PCR-quality DNA without using the expensive and pathogen-specific IMS
- Turnaround times have to be reduced to allow close to real-time detection for routine monitoring
- Quantitative and high resolution typing procedures (i.e., subgenotyping) need to be incorporated for analysis of samples in special situations (such as outbreaks or bioterrorism)
- Utilization of new techniques such as real-time PCR, biosensors, and microarrays.
- Continuous molecular characterization of *Cryptosporidium* parasites from various wildlife to expand current data on host-specifity
- More extensive use of SSU rRNA PCR-RFLP tools in the analysis of different environmental samples to allow more confidence in the association of common genotypes with environmental settings.

REFERENCES

Fayer, Ronald, C.A. Speer, and J.P. Dubey. 1997. The General Biology of *Cryptosporidium*. In: *Cryptosporidium* and Cryptosporidiosis. Ed., Ronald Fayer. CRC Press. New York. pp. 1-41.

Fayer, R., J.M. Trout, L. Xiao, U.M. Morgan, A.A. Lal, and J.P. Dubey. 2001. *Cryptosporidium canis* n.sp. from domestic dogs. J. Parasitol. Dec., **87**(6):1415-1422.

Fayer, R. , T.K. Graczyk, E.J. Lewis, J.M. Trout, C.A. Farley. 1998. Survival of infectious *Cryptosporidium* parvum oocysts in seawater and eastern oysters (*Crassostrea virginica*) in the Chesapeake Bay. Appl. Env. Microbiol. 64: 1070-1074.; http://aem.asm.org/cgi/content/full/64/3/1070

Jenkins, M.C. and C. Petersen. Molecular Biology of *Cryptosporidium* . 1997. In: *Cryptosporidium* and Cryptosporidiosis. Ed., Ronald Fayer. CRC Press. New York. pp. 225-232

Limor, J.R., A.A. Lal, and L. Xiao. 2002. Detection and differentiation of *Cryptosporidium* parasites that are pathogenic for humans by real-time PCR. J. Clin. Microbiol. **40**(7):2335-8

Lindquist, H.D. J.W. Bennett, M. Ware, R.E. Stetler, M. Gauci, and F.W. Schaefer. 2001. Testing methods for detection of *Cryptosporidium* spp. in water samples. Southeast Asian J Trop Med Public Health **32** (Suppl 2): 190-4.

McDonald, L.M., K. Sargent, A. Armson, R.C. Tompson, J.A. Reynoldson. 2002. The development of a real-time quantitative PCR method for characterization of *Cryptosporidium parvum* in vitro culturing system and assessment of drug efficiency. Molecular and Biochemical Parasitology. (2):279-282.

*Sulaiman, I.M., L. Xiao, C. Yang, L. Escalante, A. Moore, C.B. Beard, M.J. Arrowood, and A.A. Lal. 1998. Differentiating Human from Animal Isolates of *Cryptosporidium parvum*. Emerging Infectious Diseases, Vol 4, No. 4, Oct-Dec, 1998. pp. 681-685; http://www.cdc.gov/ncidod/eid/vol4no4/sulaiman.htm

*Sulaiman, I.M., A.A. Lal, M.J. Arrowood, and L. Xiao. 1999a. Biallelic Polymorphism in the Intron Region of β-tubulin Gene of *Cryptosporidium* Parasites. Journal of Parasitology, Volume 85, Issue 1, 1999, pp. 154-157

*Sulaiman, I.M., L. Xiao, A.A. Lal. 1999b. Evaluation of *Cryptosporidium parvum* Genotyping Techniques. Journal of Applied and Environmental Microbiology, Vol 65, Issue 10, pp. 4431-4435; http://aem.asm.org/cgi/content/full/65/10/4431

Sulaiman, I.M., U.M. Morgan, R.C. Thompson, A.A. Lal, L. Xiao. 2000. Phylogenetic Relationships of *Cryptosporidium* Parasites Based on the 70-Kilodalton Heat Shock Protein (HSP70) gene. Journal of Applied and Environmental Microbiology, Vol 66, Issue 6, pp. 2385-91; http://aem.asm.org/cgi/content/full/66/6/2385

Sulaiman, I.M., A.A. Lal, L. Xiao. 2002. Molecular Phylogeny and Evolutionary Relationships of Cryptosporidium Parasites at the Actin Locus. J. Parasitology. Vol 88, pp. 388094

Working Group on Waterborne Cryptosporidiosis. 1997. *Cryptosporidium* and Water: A Public Health Handbook. Atlanta, Georgia; http://www.cdc.gov/ncidod/diseases/crypto/crypto.pdf

*Xiao, L., I. Sulaiman, R. Fayer, A.A. Lal. 1998. Species and Strain-Specific Typing of *Cryptosporidium* Parasites in Clinical and Environmental Samples. In: Mem Inst Oswaldo Cruz, Rio De Janeiro, Vol 93(5), Sep/Oct 1998; http://www.dbbm.fiocruz.br/www-mem/935/23m.pdf

*Xiao, L., L. Escalante, C. Yang, I. Sulaiman, A.A. Escalante, R.J. Montali, R. Fayer, and A.A. Lal. 1999a. Phylogenetic Analysis of *Cryptosporidium* Parasites Based on the Small Subunit Ribosomal RNA Gene Locus. Journal of Applied and Environmental Microbiology, Volume 65, Issue 4, 1999, pp. 1578-1583; http://aem.asm.org/cgi/content/full/65/4/1578#T2

Xiao, L., U.M. Morgan, J. Limor, A. Escalante, M. Arrowood, W. Shulaw, R.C.A. Thompson, R. Fayer, and A.A. Lal. 1999b. Genetic Diversity within *Cryptosporidium parvum* and Related *Cryptosporidium* species. Applied and Environmental Microbiology. Vol 65, No. 8, pp. 3386-3391). http://aem.asm.org/cgi/content/full/65/8/3386

*Xiao, L., K. Alderisio, J. Limor, M. Royer, and A.A. Lal. 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; http://aem.asm.org/cgi/content/full/66/12/5492

*Xiao, L. A. Singh, J. Limor, T. Graczyk, S. Gradus, and A. Lal. 2001. Molecular characterization of *Cryptosporidium* Oocysts in Samples of Raw Surface Water and Wastewater. Appl. Env. Microbiol. 67:1097-1101; http://aem.asm.org/cgi/content/full/67/3/1097

*Xiao, L. and A.A. Lal. 2002a. Molecular detection of *Cryptosporidium* oocysts in water: the challenge and promise. (Draft) WQTC Conference, 11/02.

Xiao, L., I.M. Sulaiman, U.M. Ryan, L. Zhou, E.R. Atwill, M.L. Tischler, X. Zhang, R. Fayer, A.A. Lal. 2002b. Host Adaptation and Host-Parasite Co-evolution in *Cryptosporidium*: Implications for Taxonomy and Public Health. 2002. International Journal for Parasitology. In Press.

---

* Work funded partially or completely under EPA/NRMRL-HHS/CDC interagency agreement 75937984

APPENDIX 1 – Molecular Tools

DNA sequencing determines the exact order of the nucleotides in the DNA molecule. DNA sequencing is important for a variety of reasons, including the identification or confirmation of particular genotypes. DNA sequencing is commonly performed by variations on the Sanger method, which involves producing a sequence of DNA fragments that are progressively longer by one, known nucleotide. By ordering the fragment lengths by electrophoresis, which is sufficiently sensitive to separate DNA fragments differing in length by one nucleotide, and then matching the corresponding terminal nucleotide, the DNA sequence is determined. Sequencing gives the most detailed information, but it is a relatively slow and expensive method.

Restriction Fragment Length Polymorphism (RFLP) is a faster, less costly, and less detailed approach to characterizing DNA. The RFLP method cuts a selected segment of DNA into fragments using restriction enzymes, which only cut the DNA at locations straddled by specific 4- or 6-nucleotide sequences. Hence, the number and length of the fragments generated in the RFLP process depend on (1) the selection of the restriction enzyme and its particular target sequence, and (2) the existence, number, and location on the original DNA segment of the target sequence for the restriction enzyme. The differing DNA fragment lengths are separated by electrophoresis and visualized by staining procedures. An RFLP process can be designed to produce a known number and length of fragments by first sequencing the target gene, then identifying the locations of the target sequences for the available restriction enzymes, and finally, selecting the restriction enzyme that gives the best fragment combination for detecting or differentiating particular genotypes.

DNA sequencing and RFLP procedures both require very large numbers of copies of the target gene sequence. These copies are prepared by another critical molecular tool – the **polymerase chain reaction (PCR)**. The PCR procedure uses heating and cooling cycles, specific DNA primer sequences, and a supply of nucleotides to mimic the DNA replication process and rapidly generate multiple copies of the desired DNA sequence. The number of copies produced is 2^n, where n is the number of cycles. Starting with one double-strand of DNA, 20 cycles of 100% efficient PCR will produce about 1 million copies, and 35 cycles will produce about 35 billion copies. In addition to copying target gene sequences for use in other methods, PCR can also be used to detect the presence of particular organisms. When used for detection purposes, DNA primers are selected that will bind to unique DNA sequences of the target organism. If the target DNA is present, then replication of the DNA primer sequences and the intervening DNA sequence will occur. If the target DNA is not present, then the primers will not attach and the replication process will not begin. **Nested PCR** is a technique that uses two rounds of PCR on the same target sequence to increase the sensitivity of the method by increasing the number of copies of the target DNA sequence that can be produced. In the first round of PCR the DNA primers are chosen so that the DNA segment copied includes the target sequence as well as a substantial amount of DNA on both sides of the target sequence. In the second round of PCR the primers are selected to copy only the target sequences on a portion of the copies of the first-round PCR products.

Cloning is another approach for making copies of a portion of a genome. Cloning involves the cutting and insertion of the target DNA sequence into a vector, which is then inserted into a bacterium where the target DNA sequence is replicated along with the vector DNA and the bacterium. Following replication the target DNA is separated from the bacteria and vector and the desired manipulations or analyses are performed. Cloning has the advantage of not requiring prior knowledge of the cloned sequence, whereas PCR requires that an approximately 20-nucleotide sequence be known for the primer sequence. The molecular tools for determining species and genotype of *Cryptosporidium* oocysts are not quantitative methods, so they characterize the DNA sequences, but do not provide accurate data on the initial concentration of each genotype present.