CENTER for PRODUCE SAFETY

CPS 2019 RFP
FINAL PROJECT REPORT

Project Title

Analysis of the presence of *Cyclospora* in waters of the Mid-Atlantic States and evaluation of removal and inactivation by filtration

Project Period

January 1, 2020 – December 31, 2020 (extended to February 28, 2022)

Principal Investigator

Kali Kniel
University of Delaware
Department of Animal and Food Sciences
531 South College Avenue
Room 044, Townsend Hall
Newark, DE 19716-2150
T: 302-831-6513
E: kniel@udel.edu

Objectives

1. *Confirmation of the presence of Cyclospora cayetanensis in presumptive positive environmental water samples using the Sanger (dideoxy) sequencing method.*
2. *Evaluation of zero-valent iron (ZVI) and sand filtration in the removal and inactivation of C. cayetanensis surrogates in artificial agricultural water.*

NOTE: This project received an extension through January 2022 to complete additional analysis of the samples from the project and to use the funds remaining from the original budget. The Addendum results are presented on pp. 13–20.

Funding for this project provided by the Center for Produce Safety through:
CDFA 2019 Specialty Crop Block Grant Program & CPS Campaign for Research

FINAL REPORT

Abstract
Cyclospora cayetanensis is a lesser studied member of the Apicomplexan family of protozoa, which includes more well studied protozoan genera like *Cryptosporidium*, *Eimeria*, and *Toxoplasma*. However, *C. cayetanensis* is an important protozoan parasite that continues to cause outbreaks of gastroenteritis associated with contaminated fresh produce. *Cyclospora* oocysts may be transmitted through contaminated surface water, thereby making reliable detection imperative for public health and produce safety. Polymerase chain reaction (PCR) analysis is used for detection of a presumptive positive water sample. Further analysis of that PCR amplicon by sequencing provides evidence of the close familial genetic relationship of these protozoa, as numerous samples positive by PCR were not identified as *Cyclospora* following phylogenetic analysis of amplicon sequencing. This analysis is complex and does not yield simple answers. Given the difficulties with detection, and the potential for contamination based on epidemiological studies, intervention strategies for water are valuable. Filtration using a laboratory filtration unit containing sand and zero-valent iron (ZVI) can successfully remove oocysts from contaminated pond water. In this project, filtration using the laboratory-scale ZVI-sand column achieved a 4.3 log reduction of *Cryptosporidium parvum* oocysts compared to a 1.8 log reduction using a sand-only filter. Similarly, a ZVI-sand filter resulted in a 6-log removal of *Eimeria tenella* oocysts compared to an average 2.3 log removal by a sand filter. This project provides evidence of the need for continued study of *Cyclospora* detection and removal methodologies.

Background
Cyclospora cayetanensis is a protozoan parasite that causes gastrointestinal illness following consumption, which often occurs with contaminated water or fresh produce. Previous foodborne outbreaks have involved contaminated herbs like basil and cilantro, leafy greens, sugar snap peas, and raspberries and strawberries. While it is speculated that contaminated water may play a role in transmission, this has not been confirmed. *Cyclospora cayetanensis* is not a zoonotic pathogen. *Cyclospora* is still considered an emerging pathogen by many, due to the limited understanding surrounding the transmission and survival of oocysts in the environment along with difficulties and limitations in detection of oocysts. Over the past two decades there has been an increasing number of illnesses; some associated with fresh produce and some not connected to a specific food. The increase in clinical illness detection could be in part due to an increase in use of rapid culture-independent diagnostic testing. *Cyclospora* is an apicomplexan coccidian protozoan that belongs to the subphylum subphylum Apicomplexa, subclass Coccidiasina, and family Eimeriidae. At this time there are nineteen *Cyclospora* species that are known to be associated with illness in various animals, and only one species that infects humans, namely *C. cayetanensis* (Ortega and Sanchez, 2010; Lainson, 2005). The elucidation of *C. cayetanensis* has included continuing phylogenetic analysis since 1986 (Hadjilouka and Tsaltas, 2020). Phylogenetic analysis of *C. cayetanensis* small subunit rRNA gene sequences, the mitochondrial genome, and apicolast genomes show close relationships to organisms of the *Eimeria* genus as well as a strong relation to *Toxoplasma* (Li et al., 2017; Cinar et al., 2015; Zhao et al., 2013).

This project has two main objectives that address critical data gaps that affect the produce industry across multiple commodities. As the number of *C. cayetanensis* outbreaks associated with produce continue to increase, knowledge of the organism and use of intervention strategies will be crucial to ensure the safety of produce. The first objective is to provide an understanding of the impact of *C. cayetanensis* on waters in the Mid-Atlantic region of the United States, a previously unstudied area. The water sources tested included those that could be used for irrigation of raw agricultural commodities such as surface water (pond, river), reclaimed

wastewater, and reclaimed produce wash water. Previous data suggested a number of samples were presumptive positive for *C. cayetanensis* as determined by PCR. This project included further analysis of these samples by amplicon (PCR product) sequencing and subsequent analysis. The second objective includes testing the effectiveness of ZVI filtration in the removal and inactivation of parasitic pathogens, to improve pre-harvest food safety. This work will also facilitate the development of novel on-farm filtration technology and guidelines for commercial applications to control parasitic pathogens in agricultural water, thus improving the safety of produce and reducing parasitic foodborne illness. The proposed ZVI technology may offer advantages to small and midsized growers, including adaptation to existing filtration systems, feasibility, effective filtration and broad neutralization of numerous biological and chemical hazards, with minimal environmental impacts.

Research Methods and Results

Methods – Objective 1

Sample collection and detection

(While this part of the project was not funded by CPS, it is critical to the project.)
Water samples were initially collected and studied for the presence of several protozoan parasites, including *Cryptosporidium*, *Giardia intestinalis*, *Toxoplasma gondii*, and *Cyclospora cayetanensis*. Seventy-two water samples were collected from six sites in June 2017 through September 2018. Samples were collected biweekly during fruit and vegetable growing seasons (June to October 2017 and May to September 2018) and then samples were collected once during the non-growing season time periods (November 2017 to April 2018, and October 2018). Water types included surface (pond and tidal river water) and reclaimed (vegetable processing water and treated wastewater). Physicochemical data were measured in triplicate with EXO2 or ProDSS multiparameter water quality sonde/meter (data not included in this report, but available upon request). Samples were transported to the laboratory for immediate filtration or stored at 4°C for up to 24 hours. Up to 20 L of each water sample were agitated and then filtered, concentrated, and pelleted in accordance with the EPA 1623.1 method. DNA was extracted from pellets using DNeasy® PowerWater® kits (Qiagen, Hilden, Germany), and performed in accordance with recommendations and instructions from the manufacturer. Pellets equal to or less than 0.5 mL in volume were stored at −80°C until DNA was extracted. Pellets greater than 0.5 mL were subsampled into quantities of 0.5 mL or less and then stored at −80°C. Real-time PCR was performed using the QuantiNova™ Probe PCR kit (Qiagen) according to manufacturer's instructions. For all water samples, if one replicate of the unknown sample was positive, the sample was considered positive. PCR detection of *C. cayetanensis* included use of forward primer (5'- TAG TAA CCG AAC GGA TCG CAT T-3'), reverse primer (5'-AAT GCC ACG GTA GGC CAA TA-3'), and probe (5'-/56-FAM/CCG GCG ATA/ZEN/GAT CAT TCA AGT TTC TAG C/3IABkFQ/-3') (FDA BAM Chapter 19b; Murphy and DeSilva, 2016).

Presumptive positive sample analysis, sequencing, and sequence analysis

As detailed above, environmental water samples collected from the Mid-Atlantic region were collected in accordance with a modified EPA 1623.1, "*Cryptosporidium* and *Giardia* in Water by Filtration/IMS/FA," method. The method was modified to exclude immunomagnetic separation (IMS) and fluorescent antibody (FA) testing, as there are not widely accepted IMS and FA methods for *C. cayetanensis*. The confirmation of the presence of *C. cayetanensis* in presumptive positive environmental water samples was performed using the 19a:9, "PCR Analysis" method of nested PCR from the U.S. Food and Drug Administration Bacteriological Analytical Manual (Orlandi et al., 2003) (**Table 1**).

Samples were then purified using the Thermo Scientific GeneJet PCR Purification kit (Thermo Scientific, Waltham, MA). The purified samples were submitted to the Delaware Biotechnology Institute (Newark, DE) for dideoxy Sanger sequencing with ABI Prism 3130XL Genetic Analyzer (Applied Biosystems). Data were analyzed by assistance of a computational biologist, Dr. Shawn Polson, Director of the University of Delaware Bioinformatics Center. Chromatograms were imported into Geneious (v 10.2.6). Sequences were trimmed for quality (Q>0.95) and to remove primer sequences, and read pairs were assembled into one contiguous sequence per sample. Contigs were aligned using the SINA aligner (v 1.2.11) with the SILVA global SSU rRNA alignment template using default settings (Pruesse et al. 2012). Aligned sequences were to a 341-position region shared by 30 samples and a tree was constructed using FastTree (v2.1.11) using a GTR model (Price et al. 2010). Additional homology searches were performed using BLASTN against the GenBank nucleotide database (Altschul et al. 1990).

Results – Objective 1

Nested PCR amplification was performed on all presumptive positive water samples. Primers were tested using DNA from *C. cayetanensis*, *Eimeria acervulina*, *E. maxima*, and *E. tenella* as positive controls. Nuclease-free water, and DNA from *Cryptosporidium parvum* served as negative controls. Negative water samples were spiked to test for inhibitors and other factors that could interfere with PCR. Results indicated that the nested PCR method increased sensitivity of the detection. The limit of detection for *C. cayetanensis* decreased from 250 genomic units to 2.5 genomic units.

Forty-two amplicons were sent for sequence analysis. Based on initial analysis, one sample showed 99.9999% similarity to *C. cayetanensis* based on the sequence that is in GenBank; however, after further review this sample appears to include multiple sequences, including an 80 out of 81 base pair match to *C. cayetanensis*. Due to the remaining ambiguous base pairs in this sample, it is still under review. The repeated primary amplicon was used in sequence analysis and sequences compared to those in the GenBank database. Sequence analysis was not straightforward. There were strong base calls at most positions, but a significant number of positions showed a mixed signal. In these cases, there may be multiple copies of the gene, and this occurs with protozoa due to the presence of multiple life stages within an oocyst, for example. In several samples, there appear to be more than one signal in the sample, so in these cases it is consistent with having mixed amplicons in the sample. In others (like 5.3) the reverse primer is encountered much earlier than expected (i.e., the amplicon is not the correct length). These issues continue to be assessed.

Several phylogenetic trees were created to better understand the similarities across these samples and including the top 10 to 25 similar microbes from the GenBank database in a blast sequence alignment. A phylogenetic tree is a hypothesis of the evolutionary inheritance of genes across individual taxa. Trees have been used to summarize an organism's pedigree. The organisms are then grouped by similarities in genetic sequence. In terms of sample analysis, several of the samples fall into specific clades, which means that these samples contain organisms that are all the descendants of a common ancestor. The phylogenetic tree included in **Figure 1** shows the evolutional distance of these organisms. There is greatest similarity in the *Cyclospora* gene fragment (indicated as tCL) and several *Eimeria* and *Isospora* species. *Toxoplasma gondii* is included in this tree as an out-group, so one can see the distance from this organism, which is not as closely related as the others and has low to no similarity with the samples.

The finding in this dataset that this apicomplexan group of protozoa is difficult to speciate is supported by the literature. It may be possible that a combination of 18S and COX1 mitochondrial genes could be useful, as used by other scientists or if there is a means of constructing a longer amplicon for this analysis. Phylogenetic analysis will continue on these

samples with the assistance and guidance of Dr. Shawn Polson, and final results will be included in a future publication.

Methods – Objective 2

Protozoa preparation

Cryptosporidium parvum oocysts were purchased from Waterborne Inc. (New Orleans, LA) and *Eimeria tenella* oocysts were a gift from Dr. Mark Jenkins at the USDA-ARS Animal Parasitic Diseases Laboratory (Beltsville, MD). All protozoa were stored at 4°C and used within 3 months of receiving.

Sand and sand:ZVI column construction and utilization

Sand columns, which served as a control, and sand:ZVI columns, referred to as test columns, were constructed in pairs as previously described (Shearer and Kniel, 2018). A steel mesh disk (60 mesh; 60 openings per in^2 mesh, wire diameter 0.0075 in.) was placed at the bottom of each acrylic column (3.8 inside diameter by 10 cm). Columns were wet packed to the top with sterile deionized water and sand alone (for the control columns) or a 1:1 (w/w) mixture of sand and ZVI (particle size 0.005 to 0.125 in.). Columns were sealed with tubing affixed at the center of the top and bottom.

A 3-L volume of sterile water was passed over the columns before the start of the filtration trials. Each organism was analyzed in triplicate, meaning that 3 trials were performed for each of the two protozoa studied. The column pore volume was calculated to be 45 ml. Each experiment included the following parts: 1) a 45-mL initial flush with autoclaved pond water; 2) inoculation with a 45-mL parasitic suspension containing one million oocysts (1×10^6) in autoclaved pond water; 3) three individual pulses, of 45 mL each, of autoclaved pond water (3 x 45 mL). Liquid was pumped through the columns at a rate of 1 mL/min by using a Tris Peristaltic pump. Throughout the filtration study 5-mL fractions of eluate were collected in sterile glass tubes with a Foxy Jr. fraction collector. Fractions were pooled (3 tubes in each subsample, 15 ml) and stored at 4°C. Studies were performed in triplicate, control groups include a negative control, and a positive control. Removal and inactivation of oocysts were assessed using qPCR.

Identification of *Cryptosporidium parvum* oocysts and infectious *Cryptosporidium*

Oocysts were detected in pooled 5-mL samples of collected filtrate using quantitative PCR analysis. Each sample was tested in duplicate and all samples were compared to a standard curve using DNA extracted from serially diluted oocysts. The reduction in infectivity of *C. parvum* oocysts was quantified by an integrated cell culture and qPCR assay. *Cryptosporidium* oocysts were bleached, and triple rinsed with Hanks balanced salt solution (HBSS) prior to filtration. Untreated oocysts and nuclease-free water were used as positive and negative controls. Human cell culture was infected with oocysts as described previously (Craighead et al., 2020). In brief, filtrate samples were pooled as described above, and then were applied to HCT-8 cell confluent monolayers in a 6-well cell culture dish, incubated for 2 h, and then cells were triple rinsed with HBSS. Infected cells were incubated at 37°C with 5% CO_2 for 48 h. DNA extraction was completed using the Zymo Research Quick-DNA Extraction kit. Then qPCR was performed with the Qiagen SYBR Green PCR kit and Rotor-Gene Q cycler (Qiagen). In addition, each sample was run in duplicate in qPCR. A standard curve was generated using control oocysts, which was then imported and adjusted for each qPCR run. The primer sequences for *Cryptosporidium parvum* detection are as follows: forward primer (CPHSPT2F, 5' TCC TCT GCC GTA CAG GAT CTC TTA 3') and reverse primer (CPHSPT2R, 5' TGC TCT TAC CAG TAC TCT TAT CA 3') (LeChevallier et al., 2003). The PCR reaction volume of 25 uL, included 12.5 uL of Rotor-Gene SYBR Green PCR master mix, 1 uL (400 nM) of each primer, forward and reverse, 8 uL of DNA template, and 2.5 uL of nuclease-free water. PCR was performed using the Rotor-Gene Q

thermocycler (Qiagen) with the following settings: hold at 95°C for 5 min, followed by 40 cycles of denaturation at 95°C or 5 s and combined annealing and extension at 60°C for 10 s.

Identification of *Eimeria tenella* oocysts

Oocysts were detected in pooled 5-mL samples of collected filtrate using quantitative PCR analysis. Each sample was tested in duplicate and all samples were compared to a standard curve using DNA extracted from serially diluted oocysts. A bead beating step was included to break oocyst and sporocyst walls prior to DNA extraction. Primers ET_F (TGGAGGGGATTATGAGAGGGA) and ET-R (CAAGCAGCATGTAACGGAGA) were used, according to Thabet et al. (2015). The preparation of the PCR reaction solution consisted of 10 µl of SYBR green master mix (Thermo Scientific, Germany), 0.4 µl of a 25 µM stock of ET-F (final concentration, 0.5 µM), 0.4 µM stock of ET-R (final concentration, 0.5 µM), and 7.2 µl of nuclease-free water. The template volume was 2 µL, yielding a final 20-µL volume in the reaction tube. Cycling reaction was performed by 40 cycles of 30 s at 95ºC, 20 s at 62ºC and 20 s at 72ºC. A melting curve program involved heating from 60 to 95ºC at a rate of 0.1 ºC/s and was performed to create the dissociation curve.

Statistical analysis

Data were analyzed using a Matched Pairs test in JMP software. Differences were considered significant when the *P* value is < 0.05.

Results – Objective 2

Filtration of *Cryptosporidium parvum* oocysts

Cryptosporidium parvum is a zoonotic pathogen transmitted by oocysts that are spherical in shape and 4-5 microns in size. After filtration through either sand or a 1:1 sand and ZVI mixture, detection levels of *C. parvum* were obtained using quantitative PCR. Data are the average of three trials. Filtration using sand alone resulted in an 85.8% reduction of *C. parvum* oocysts across the filtration study, as determined by detection of genomic copies. Filtration using ZVI and sand resulted in a 99.9% reduction of *C. parvum* oocysts as determined by detection of genomic copies over the filtration study. Filtration using the sand:ZVI column achieved a 4.30 log reduction of *C. parvum* compared to 1.82 log reduction using sand alone for filtration. There was a statistically significant difference in filtration treatments ($P = 0.0238$) over the course of this study. This is shown in **Figure 2**.

Due to the fact that oocysts reproduce in cell culture in an undefined number of asexual reproduction events, it is difficult to state how many oocysts correlate to the detection signal associated with life stages following qPCR and integrated cell culture. These analyses show that *C. parvum* oocysts remained infectious post-filtration for both treatment columns. Oocysts collected within the first 45 mL were the most infectious, as indicated by a greater signal detected by qPCR. These oocysts had the lowest time in contact with the column, sand or sand:ZVI mix.

Filtration of *Eimeria tenella* oocysts

Eimeria tenella is an avian pathogen transmitted by oocysts that are ovoid to ellipsoid in shape and approximately 19-22 microns in size. Detection of *E. tenella* was determined using quantitative PCR from filtrate from either sand or 1:1 sand and ZVI columns. Data are the average of three trials. Filtration with a sand and ZVI mix column resulted in no detectable *E. tenella*, as determined by zero genomic copies per reaction in all samples tested. Filtration using sand-only columns resulted in a 99.5% reduction of *E. tenella*, as determined by analysis of genomic copies across all samples. From total analysis, sand:ZVI filtration achieved a 6 log reduction, whereas sand-only filtration reduced *E. tenella* by 2.3 log. This is shown in **Figure 3**.

Outcomes and Accomplishments

Environmental samples provide great challenges to microbiologists in terms of their microbial content along with potential for inhibitors. In this case, it is not the inhibitors but the complexity of detection of protozoa with numerous life stages and potentially numerous gene copies with these complex environmental samples that provide a challenge. One outcome of this project is informing the need to perform sequence analysis of these environmental samples given that a positive PCR is likely just a presumptive positive. However, sequence analysis is accompanied by a set of challenges that must be further investigated. Perhaps this project is nearly the tip of the iceberg in terms of elucidating the need to better understand detection of *Cyclospora cayetanensis*. An accomplishment of this project is providing evidence that sand filtration provides efficient removal of small and large protozoa and inclusion of ZVI with sand can increase removal efficiencies of both *Cryptosporidium* and *Eimeria*.

Summary of Findings and Recommendations

Research with *C. cayetanensis* is severely limited by the lack of resources and inability to obtain *C. cayetanensis* oocysts. Throughout history, under such conditions scientists have garnered vast amounts of data through the use of surrogate microorganisms. In this study, two protozoan surrogates were used. *Cryptosporidium* is a robust pathogen resistant to chemical disinfectants but smaller in size compared to *Cyclospora*. *Eimeria tenella* is a bit larger than *Cyclospora* but more similar in morphology to *Cyclospora*. Together these two pathogens provide useful information.

Phylogenetic analysis of apicomplexan protozoa is complex. The need for additional analysis following PCR detection of presumptive *Cyclospora* positive environmental samples is clear; however, the pathway to further identification of these samples remains murky. Computational biologies may be able to clarify this issue given more time for analysis. Microbiologists perform multiple steps of analysis on bacterial detection and confirmation from environmental samples, including biochemical tests, culture isolation, and sequencing. It seems that a similar set of multiple steps to analysis is warranted for these complex environmental samples that may be positive for *Cyclospora*. Alternate means of detection, using multiple primer sets and potential for inclusion of mitochondrial genomes should also be considered.

Sand biofilters are used for recreational water purification and on some farms. The incorporation of zero-valent iron has previously been shown to increase the removal and inactivation of bacteria and viruses by filtration. The project findings suggest that zero-valent iron can enhance the efficacy of filtration for removing protozoa of varying sizes (5 to 22 microns). Future research should address the inactivation of protozoan oocysts, including feasibility and mechanistic studies.

NOTE: This project received an extension through January 2022 to complete additional analysis of the samples from the project and to use the funds remaining from the original budget. The Addendum results are presented on pp. 13–20.

APPENDICES

Publications and Presentations

Presentations:
A. Kelly, Confirmation of the presence of *Cyclospora cayetanensis* in presumptive positive environmental water samples, Seminar presentation to the Department of Animal and Food Sciences, March 29, 2021.

A. Kelly, S. Craighead and K.E. Kniel, Confirmation of the Presence of *C. cayetanensis* and *C. parvum* from Environmental Water Samples, International Association for Food Protection Annual Meeting, Phoenix, Arizona, July 2021, P2-173.

Budget Summary

The project was severely impacted by the COVID-19 pandemic, which resulted in the closure of the University of Delaware and suspension of the majority of laboratory work from March to September 2020. This study continued to be affected by the COVID-19 pandemic after that period due to the protective safety rules in place at the University allowing low numbers of laboratory personnel in at one time. The project was additionally impacted by a move to a new laboratory, that had to be rescheduled for January 2021 from its previously scheduled time due to the pandemic and resulted in an additional 6 weeks of laboratory closure. The delays impacted spending on personnel and some supplies for some studies. The need to include funds spent on bioinformatics support became obvious following initial analysis of the sequencing data. The scope of the project did not change, and we will continue data analysis through to the publication of this work. A total of $49,170 in research funds was awarded to this project. The majority of funds were spent within the past couple of months due to constraints in the laboratory associated with the physical move and COVID-19 restrictions.

We would like to thank CPS for the no-cost extensions granted in this project associated with loss of productive time due to the pandemic. We regret that the delays impacted this project, as all were beyond the control of the research team.

Table 1 and Figures 1–3 (see below)

Table 1: DNA Primer Sequences for *Cyclospora*-specific PCR Amplification (18S rRNA gene). These are used in the nested PCR analysis, adapted from Orlandi et al., 2003.

Primer Code	Primer Specificity	Primer Sequence (5'-3')	Amplicon Size (bp)	Application
F1E (Forward)	*Cyclospora* and *Eimeria* spp.	TACCCAATGAAAACAGTTT	636	Primary Amplification
R2B (Reverse)		CAGGAGAAGCCAAGGTAGG		
F3E (Forward)	*Cyclospora* and *Eimeria* spp.	CCTTCCGCGCTTCGCTGCGT	294	Nested Amplification
R4B (Reverse)		CGTCTTCAAACCCCCTACTG		

Figure 1. Phylogenetic tree representing NCBI blast results showing genetic similarities to presumptive positive environmental water samples identified as *C. cayetanensis* through nested PCR (numerical entries) with *Toxoplasma gondii* included as an out-group. The sample tCL is the control gene fragment of *C. cayetanensis* purchased from IDT (HMgBlock135m).

Figure 2. Removal of *Cryptosporidium parvum* oocysts by sand and sand:ZVI column filtration. Values shown are indicative of genomic signal from oocysts detected. The lack of signal indicates few to no oocysts detected in those analyzed filtrate samples. Samples of 5 mL each were collected throughout filtration, as follows: the parasitic suspension in samples collected at 0-45 mL, flow through 1 in samples collected at 45-90 mL, flow through 2 in samples collected at 90-135 mL, and flow through 3 in samples collected at 135-180 mL.

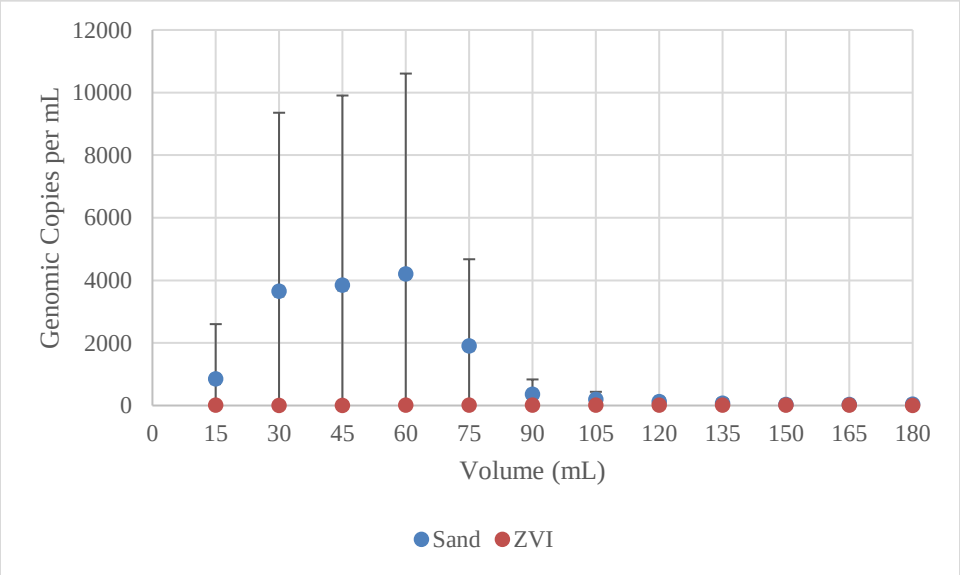
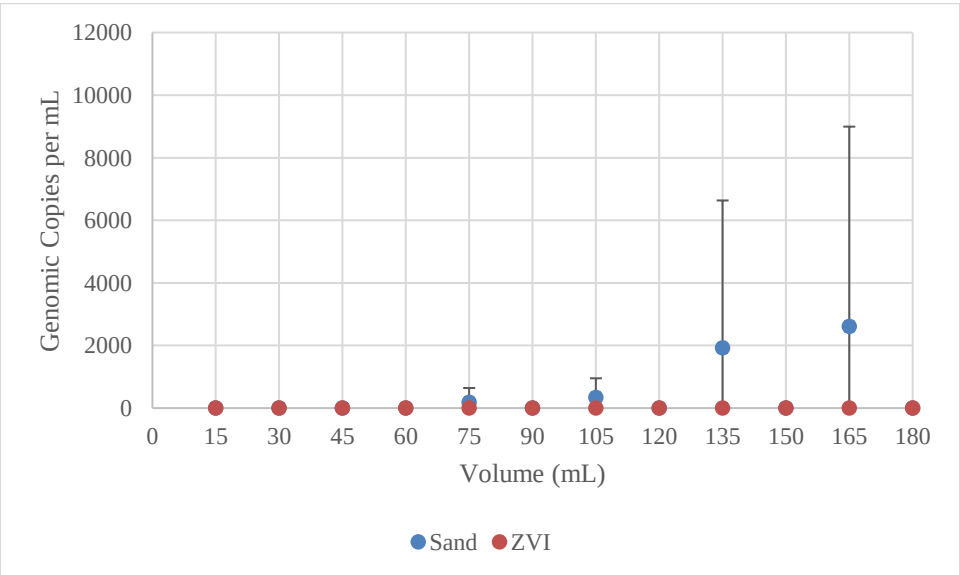


Figure 3. Removal of *Eimeria tenella* oocysts by sand and sand:ZVI column filtration. Values shown are indicative of genomic signal from oocysts detected. The lack of signal indicates few to no oocysts detected in those analyzed filtrate samples. Samples of 5 mL each were collected throughout filtration, as follows: the parasitic suspension in samples collected at 0-45 mL, flow through 1 in samples collected at 45-90 mL, flow through 2 in samples collected at 90-135 mL, and flow through 3 in samples collected at 135-180 mL.



References cited

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol*. 215:403-410.

Cinar, H.N., Gopinath, G.K., Jarvis, K., Murphy, H.R. 2015. The complete mitochondrial genome of the foodborne parasitic pathogen *Cyclospora cayetanensis*. *PLoS ONE* 10:e0128645.

Craighead, S., Hertrich, S., Boyd, G., Sites, J., Niemira, B.A., Kniel, K.E. 2020. Cold atmospheric plasma jet inactivates *Cryptosporidium parvum* oocysts on cilantro. *J. Food Prot*. 83:794-800. doi: 10.4315/0362-028X.JFP-19-442.

Hadjilouka, A., Tsaltas, D. 2020. Cyclospora cayetanensis—Major outbreaks from ready to eat fresh fruits and vegetables. *Foods* 9:1703. https://doi.org/10.3390/foods9111703.

Lainson, R. 2005. The genus *Cyclospora*: (Apicomplexa: Eimeriidae), with a description of *Cyclospora schneideri* n.sp. in the snake *Anilius scytale* (Anilüdae) from Amazonian Brazil—A review. *Mem. Inst. Oswaldo Cruz.* 100:103–115.

LeChevallier, M. W., Di Giovanni, G. D., Clancy, J. L., Bukhari, Z., Bukhari, S., Rosen, J. S., Sobrinho, J., and Frey, M. M. 2003. Comparison of method 1623 and cell culture-PCR for detection of Cryptosporidium spp. in source waters. *Appl. Environ. Microbiol*. 69(2):971-979. doi:10.1128/AEM.69.2.971-979.2003

Li, J., Chang, Y., Shi, K.E., Wang, R., Fu, K., Li, S., Xu, J., Jia, L., Guo, Z., Zhang, L. 2017. Multilocus sequence typing and clonal population genetic structure of *Cyclospora cayetanensis* in humans. *Parasitology* 144:1890–1897.

Murphy, H.R., Almeria, S., and da Silva, A.J. BAM Chapter 19b: Molecular Detection of Cyclospora cayetanensis in Fresh Produce Using Real-Time PCR in the *Bacteriological Analytical Manual (BAM)*. Accessed May 2021 at https://www.fda.gov/food/laboratory-methods-food/bam-chapter-19b-molecular-detection-cyclospora-cayetanensis-fresh-produce-using-real-time-pcr.

Murphy, H., and De Silva, A. 2016. A Real-Time PCR Assay for Detection of *Cyclospora cayetanensis* on Fresh Produce: Laboratory SOP. *U.S. Food and Drug Administration*. Available at: https://www.fda.gov

Orlandi, P. A., L. Carter, A. M. Brinker, A. J. da Silva, D. M. Chu, K. A. Lampel, and S. R. Monday. 2003. Targeting single-nucleotide polymorphisms in the 18S rRNA gene to differentiate *Cyclospora* species from *Eimeria* species by multiplex PCR. *Appl. Environ. Microbiol.* 69:4806–4813.

Ortega, Y. R. and Sanchez, R. 2010. Update on *Cyclospora cayetanensis*, a food-borne and waterborne parasite. *Clinical Microbiology Reviews* 23:218–234.

Pruesse, E., Peplies, J., Glöckner, F.O. 2012. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes, *Bioinformatics*, 28:1823-1829

Shearer, A.E.H. and Kniel, K.E. 2018. Enhanced Removal of Norovirus Surrogates, Murine Norovirus and Tulane Virus, from Aqueous Systems by Zero-Valent Iron. *J Food Prot*. 81:1432-1438. doi: 10.4315/0362-028X.JFP-18-054.

Thabet, A., Alnassan, A. A., Daugschies, A., and Bangoura, B. 2015. Combination of cell culture and qPCR to assess the efficacy of different anticoccidials on *Eimeria tenella* sporozoites. *Parasitol. Res*. 114:2155–2163. doi: 10.1007/s00436-015-4404-4

Zhao, G.H., Cong, M.M., Bian, Q.Q., Cheng, W.Y., Wang, R.J., Qi, M., Zhang, L.X., Lin, Q., and Zhu, X.Q. 2013. Molecular characterization of Cyclospora-like organisms from golden snub-nosed monkeys in Qinling Mountain in Shaanxi province, northwestern China. *PLoS ONE* 8:e58216.

ADDENDUM (03/01/2022)

In working on the project as described, it became apparent that additional questions regarding the detection of *Cyclospora cayetanensis* could be addressed by further analysis of the samples used in this study. The preliminary findings of *C. cayetanensis* oocysts in surface water in the Mid-Atlantic region of the United States is unique and warrants further exploration. Included in this addendum are the studies that were performed to address these questions, during the period of time from November 2021 to February 2022. In part, these experiments sought to 1) further understand and assess the sequence analysis of these samples, and 2) attempt to assess the issue of reagent specificity in the detection of *C. cayetanensis*.

Preparation of the CPS report and presentation discussions with FDA and industry scientists reinforced the pertinent issues associated with detection of *C. cayetanensis* from environmental water samples. The original methodology used in the CPS project included water samples processed originally in 2017 with the protocol in FDA BAM 19A. Follow-up included analysis with the primers in FDA BAM 19B and sequence analysis of the amplicons.

Three different approaches were taken to address the two broader issues (1 and 2 above) identified as the need to explore sequence analysis and reagent specificity. The first approach involves a different type of sequencing on the entire water samples. The second approach includes analysis of the samples using quantitative polymerase chain reaction (qPCR) with two different reagent kits. The third approach is analysis of the samples by digital PCR (dPCR), which includes a unique opportunity for absolute detection of the sample. These three approaches are described in this report. The methodology used in the collection of water and subsequent steps is presented in Figure A1. DNA subsamples of previously presumptive positive samples were pooled to be sure there was enough DNA for all analyses.

Approach 1. Sequencing of DNA in water samples

Previously, amplicons, which are short pieces of DNA amplified through PCR, were sequenced and the sequences were attempted to be matched to known sequences of Apicomplexan protozoa, including *C. cayetanensis*. In this study, 6M read shotgun sequencing (1Gb) was used in junction with analysis by taxonomic profiling, with specific interest in the protozoa that could be detected. Total DNA was purified from the water samples that were previously presumptive positive for the presence of *C. cayetanensis*. These samples are noted in **Table A1**. The following steps were performed by the CosmosID company, where samples were prepared, sequenced on an Illumina HiSeq platform, and data taxonomic profiling performed. DNA samples were quantified using Qubit 4 fluorometer and Qubit™ dsDNA HS Assay Kit (Thermofisher Scientific). DNA libraries were prepared using the Nextera XT DNA Library Preparation Kit (Illumina) and IDT Unique Dual Indexes with total DNA input of 1ng. Genomic DNA was fragmented using a proportional amount of Illumina Nextera XT fragmentation enzyme. Unique dual indexes were added to each sample followed by 12 cycles of PCR to construct libraries. DNA libraries were purified using AMpure magnetic beads (Beckman Coulter) and eluted in QIAGEN EB buffer. DNA libraries were quantified using Qubit 4 fluorometer and Qubit™ dsDNA HS Assay Kit. Libraries were then sequenced on an Illumina HiSeq X platform 2x150bp. Bioinformatics analysis was performed according to standard CosmosID-HUB analysis. The system utilizes a high-performance data-mining k-mer algorithm that rapidly disambiguates millions of short sequence reads into the discrete genomes engendering the particular sequences. The pipeline has two separable comparators: the first consists of a pre-computation phase for reference databases and the second is a per-sample computation. The input to the pre-computation phase are databases of reference genomes, virulence markers and antimicrobial resistance markers that are continuously curated by CosmosID scientists. The output of the pre-computational phase is a phylogeny tree of microbes, together with sets of variable length k-mer fingerprints (biomarkers) uniquely

associated with distinct branches and leaves of the tree. The second per-sample computational phase searches the hundreds of millions of short sequence reads, or alternatively contigs from draft de novo assemblies, against the fingerprint sets. This query enables the sensitive yet highly precise detection and taxonomic classification of microbial NGS reads. The resulting statistics are analyzed to return the fine-grain taxonomic and relative abundance estimates for the microbial NGS datasets. To exclude false positive identifications the results are filtered using a filtering threshold derived based on internal statistical scores that are determined by analyzing a large number of diverse metagenomes.

Table A1. Sample identification and water type for sequencing analysis and PCR

Sample ID # (P indicates pooled samples)	Sample Location/Water type	Concentration* (ng/ul) [Minimum > 20ng/ul]	"Amount (ng) [Minimum >250ng, Recommended > 1000ng]"
P1	River	87.2	4360
P2	Pond	27.6	1380
P3	Pond	11.5	575
P4	Pond	116.7	5835
P5	Pond	42	2100
P6	Reclaimed	183.5	9175
P7	River	40.2	2010
P8	Reclaimed	32.1	1605
P9	Pond	54	2700
P10	Pond	24.5	1225

Deep sequencing of these samples provides more evidence that these are complicated mixed samples. The most abundant protozoan present in the samples is *Hammondia hammondi*. This is a protozoan closely related to the zoonotic protozoan *Toxoplasma gondii* but does not cause disease in intermediate or definitive hosts (Schares et al., 2021). It is not surprising that cats are near these water sources. *Hammondia* is not genetically related to *Cyclospora*; however, *Eimeria* is in the same family as *Cyclospora*, and *Eimeria* (not species specific) was detected in some abundance in sample 6, as shown in a sunburst chart (**Figure A2**). While this is interesting information, several limitations exist in extrapolating these findings. Sequencing like this can be useful in identification of abundant microbes in a sample, which is why enrichment is often used for bacterial sequencing. Organisms can only be detected in these sequences if they exist in the database, and in this case adequate unique biomarkers must exist for each organism.

A more in-depth analysis of these samples was requested. Most interestingly, the bioinformaticians stated they identified a very weak signal to "*Cyclospora_cayetanensis*_strain_CHN_HEN01" in the sequenced data; however, the results were filtered out by their filters due to lack of enough unique biomarkers for "*Cyclospora_cayetanensis*_strain_CHN_HEN01" in their database. The bioinformaticians stated that this organism could be present but appears to be at too low an abundance to confidently detect. More information that might be available regarding these samples and what information is needed should be explored. These findings inform our use of sequencing, which cannot be relied upon solely for the detection of *Cyclospora* in complex water samples. Studies should continue to gather information on pertinent biomarkers and ways in which samples might be concentrated prior to sequence analysis.

Approach 2. Quantitative polymerase chain reaction (qPCR)

The FDA BAM chapter 19B was updated to include real-time quantitative PCR methods for the molecular detection of *C. cayetanensis*. As noted above, during various discussions in 2020-2021 regarding molecular detection, a concern for reagent specificity became apparent. To address this, samples 1–10, along with appropriate positive and negative controls, were subjected to qPCR using two different reagent systems. These are the Qiagen QuantiFast Multiplex PCR Kit (Cat No. 204654, noted in FDA BAM 19B) and Qiagen QuantiNova Probe PCR Kit (Cat No. 208254). There are a few things to consider regarding these reagents. Qiagen has discontinued sales of QuantiFast as newer reagents are now being sold, e.g., QuantiNova for qPCR and QIAcuity for dPCR. While Qiagen formulations are proprietary and a comprehensive list of reagents and their ratios is not available to the public, some differences are available on the Qiagen website. Quantifast PCR buffer contains KCl and (NH4)2SO4 while QuantiNova PCR buffer contains KCl and NH4Cl. According to the Qiagen website, the reagent that has been discontinued Quantifast "buffer contains a balanced combination of KCl and (NH 4)2SO4, which promotes a high ratio of specific to nonspecific primer binding during the annealing step of each PCR cycle. This creates stringent primer annealing conditions, leading to increased PCR specificity." Also, according to the Qiagen website, the newer reagent QuantiNova "promotes a high ratio of specific to nonspecific primer binding during the annealing step of each PCR cycle. This creates stringent primer annealing conditions, leading to increased PCR specificity." The qPCR assays using QuantiFast and QuantiNova reagents were performed on a Qiagen Rotor-Gene Q cycler, routinely used for qPCR analysis of water samples (Anderson-Coughlin et al., 2021a, 2021b).

PCR detection of *C. cayetanensis* included use of forward primer (5'- TAG TAA CCG AAC GGA TCG CAT T-3'), reverse primer (5'-AAT GCC ACG GTA GGC CAA TA-3'), and probe (5'-/56-FAM/CCG GCG ATA/ZEN/GAT CAT TCA AGT TTC TAG C/3IABkFQ/-3') (FDA BAM Chapter 19b; Murphy et al., 2017). qPCR analysis was performed on pooled samples P1–P10. A standard curve was generated in duplicate for each qPCR run, based on analysis of the synthetic positive control *Cyclospora* DNA purchased from IDT. Nuclease-free water was used as negative control. Detection for samples P1–P10 was performed in triplicate reactions across duplicate assays. Of the 13 samples for which detection was performed, including 3 negatives (data not shown), 37 (62%) and 40 (67%) replicates (n=60) were positive for QuantiFast and QuantiNova, respectively. At least one replicate was positive for all samples for each reagent, QuantiFast and QuantiNova, as shown in **Table A2**.

dCt values, which are the differences between raw Ct values (terminal Ct value-Ct value; e.g., 38-Ct value, 38-34=4) were used to perform analyses, as there were a significant number of replicates with no detection of *Cyclospora* (**Figure A3**). This is not altogether surprising given the likely very low levels of *Cyclospora* DNA in these pooled samples. To be sure there was enough volume of DNA to perform the assays, concentration techniques were not used; however, this could be further considered in the future.

Sample P2 was the only sample to have a significant difference in detection between the two reagents, QuantiNova and QuantiFast, using a T-test and significance level of $p<0.05$. Overall, QuantiNova detection was significantly greater than QuantiFast, with an average of 26.4 and 15.8 copies per reaction, average of 6.41 and 2.94 dCT (cycles prior to 38 cycles), respectively. Regarding sample P2, 6/6 replicates were positive using QuantiNova and 4/6 replicates were positive using QuantiFast. All other samples showed no significant differences between the two reagents.

Table A2. Percent of *Cyclospora* positive replicates by reagent

Sample ID	Percent of positive replicates (n=6)	
	QuantiFast	QuantiNova
P1	50% (3)	83% (5)
P2	67% (4)	100% (6)
P3	100% (6)	50% (3)
P4	50% (3)	67% (4)
P5	83% (5)	50% (3)
P6	67% (4)	67% (4)
P7	50% (3)	67% (4)
P8	50% (3)	67% (4)
P9	33% (2)	67% (4)
P10	67% (4)	50% (3)

Approach 3. Digital polymerase chain reaction (dPCR)

Digital polymerase chain reaction (dPCR) allows for improved sensitivity and precision. In dPCR, faint genetic signals can be identified even against a strong background, instead of a single positive being lost in a dense pool of negatives. The primers (FDA BAM 19B) used in the qPCR were used here as well. The sample is partitioned into a large number of individual reactions, making it easier to detect the single positive. dPCR has a lower limit of detection compared to qPCR. Detection was performed in triplicate using 1.25 µl template in the 12.5 µl reaction in a 96-well 8.5K micro-partition plate using the Qiagen QIAcuity dPCR instrument. A total of thirteen pooled samples were used: ten (P1–P10) pooled from the previous positives and three pooled from negatives (P11–P13). DNA from *Eimeria tenella*, *E. acervulina*, and *E. maxima* was included as negative controls and they were not detected. *C. cayetanensis* was detected in samples 3, 4, 5, and 10, with sample 5 being much higher than the others (**Table A3** and **Figures A4** and **A5**). Interestingly, these four samples were all collected from the same pond from between June and October 2017 (samples 3, 4, 5, and 10). This pond is in a rich agricultural area. Further analysis will be performed into any sewage incidents that may have occurred.

Table A3. Genomic copies of *C. cayetanensis* DNA detected

Sample ID	Average Genomic Copies/2 ul DNA template
P1	---
P2	---
P3	0.59
P4	1.38
P5	4.53
P6	---
P7	---
P8	---
P9	---
P10	0.46
P11	---
P12	---
P13	---

Summary

In summary, pooled samples were used in these analyses to ensure enough volume was available to conduct the analysis. This also allowed continuity in samples assessed. This may have also impacted the fact that *Cyclospora* was already at a low level in these samples, since *Cyclospora* may not have been equally represented in all fractions of one sample. The fractions that were pooled were created initiation in sample filtration. Reagent specificity was identified as a critical piece of *Cyclospora* detection during the 2021 CPS meeting, and so was assessed here. QuantiFast has been discontinued, and our results indicate that QuantiNova provided nearly identical results in the analysis of 10 presumptive positive samples. An important factor here to consider is the low sample number, which was out of our control for this study. Our findings indicate that deep sequencing may not yield direct results. To our knowledge, this is the first use of dPCR in the detection of *Cyclospora* and dPCR is advantageous over qPCR and deserves further consideration for detection of *C. cayetanensis* in complex samples.

References cited

Anderson-Coughlin, B. L., Shearer, A., Omar, A. N., Wommack, K. E., & Kniel, K. E. (2021b). Recovery of SARS-CoV-2 from Wastewater Using Centrifugal Ultrafiltration. *Methods and Protocols*, *4*(2), 32. https://doi.org/10.3390/mps4020032

Anderson-Coughlin, B. L., Craighead, S., Kelly, A., Gartley, S., Vanore, A., Johnson, G., Jiang, C., Haymaker, J., White, C., Foust, D., Duncan, R., East, C., Handy, E. T., Bradshaw, R., Murray, R., Kulkarni, P., Callahan, M. T., Solaiman, S., Betancourt, W., Gerba, C., … Kniel, K. E. (2021a). Enteric Viruses and Pepper Mild Mottle Virus Show Significant Correlation in Select Mid-Atlantic Agricultural Waters. *Applied and Environmental Microbiology*, *87*(13), e0021121. https://doi.org/10.1128/AEM.00211-21

Murphy HR, Almeria S, da Silva AJ. 2017. Molecular detection of Cyclospora cayetanensis in fresh produce using real-time PCR In *Bacteriological analytical manual*. U.S. Food and Drug Administration, Silver Spring, MD.

Schares, G., Globokar Vrhovec, M., Tuschy, M. *et al.* (2021). A real-time quantitative polymerase chain reaction for the specific detection of *Hammondia hammondi* and its differentiation from *Toxoplasma gondii*. *Parasites & Vectors 14*, 78. https://doi.org/10.1186/s13071-020-04571-8

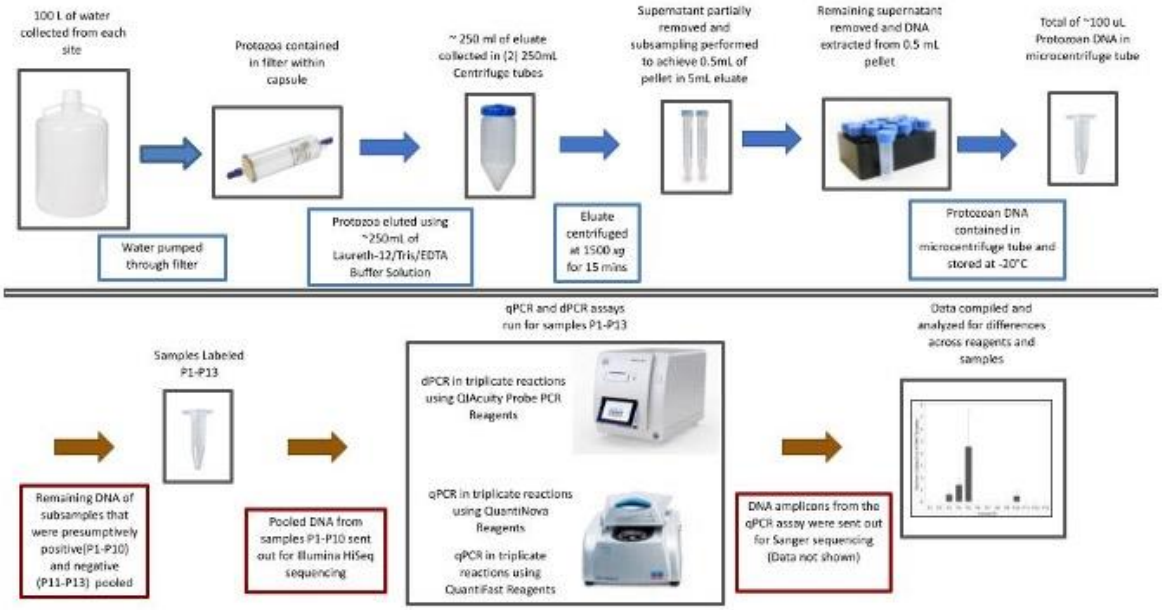
Budget Summary

CPS generously provided an extension on using the funds remaining from the original project. Due to extenuating circumstances within the University, we were unable to use the budgeted funds to support the graduate student performing the work. Funds were used to purchase the molecular biology and PCR reagents and consumables, for sequencing, and for computational analysis. Funds will also be used to support domestic travel to attend the CPS Research Symposium in June 2022.

Figures A1–A5 (see below)

(Note: Figures in the Addendum are noted as A followed by the number.)

Figure A1. Methodology from original sample collection to DNA analysis.



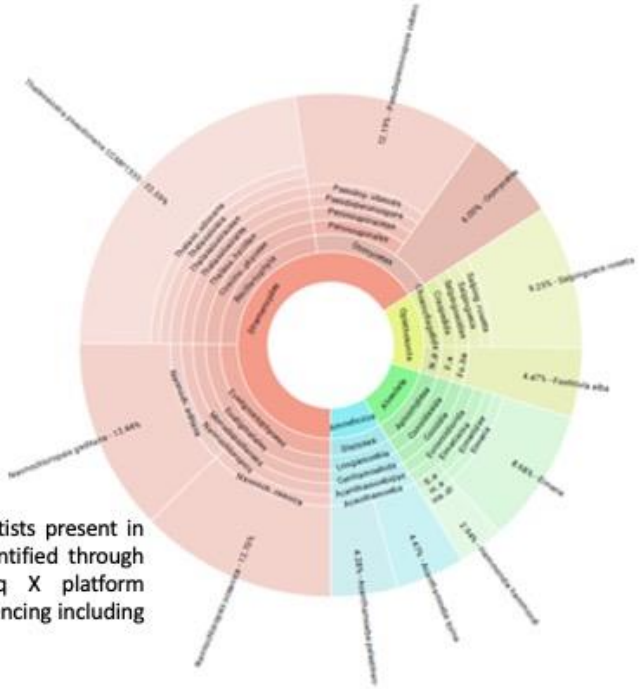


Figure A2. Protists present in sample P6 identified through Illumina HiSeq X platform 2x150bp sequencing including *Eimeria*.

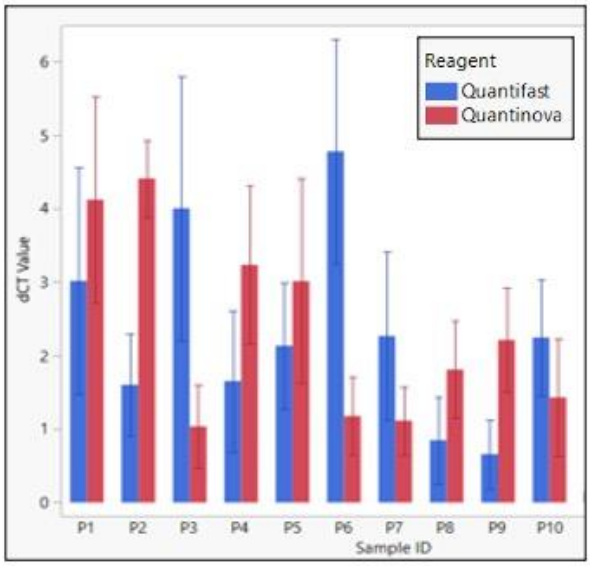


Figure A3. Detection of *C. cayetanensis* using QuantiFast (blue) and QuantiNova (red) reagents. Data are presented as dCT values (number of cycles at which detection occurs prior to the terminal cycle (38), *e.g.,*Ct value of 35 is a dCt value of 3.

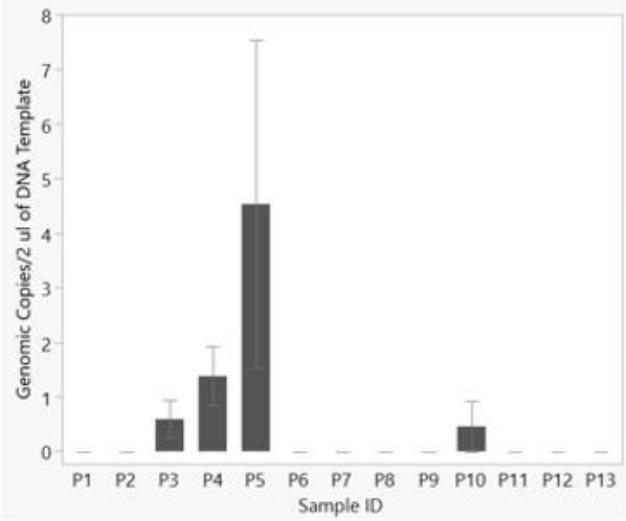


Figure A4. *Cyclospora cayetanensis* detection using digital PCR (dPCR) assay and QIAcuity Probe PCR reagents. Detection for pooled samples (n=13) was performed in triplicate. The mean genomic copies/2ul of template are shown with the associated standard error bars.

Figure A5. *Cyclospora cayetanensis* detection visual output in QIAcuity software. Sample P5 in well F2 of the plate is magnified to show the positive partitions in that well.