

Testing Methods used for DNA Barcoding of Environmental Samples from the Eightmile River for Diatom (Bacillariophyta) Identification.

Basic Information

Title:	Testing Methods used for DNA Barcoding of Environmental Samples from the Eightmile River for Diatom (Bacillariophyta) Identification.
Project Number:	2011CT226B
Start Date:	3/1/2011
End Date:	2/28/2012
Funding Source:	104B
Congressional District:	2
Research Category:	Ecological Processes
Focus Category:	Water Quality, Ecology, Methods
Descriptors:	None
Principal Investigators:	Diba A. Khan-Bureau, Louise Lewis, Gary A. Robbins

Publications

There are no publications.

Testing Methods used for DNA Barcoding of Environmental Samples from the Eightmile River for Diatom (Bacillariophyta) Identification

In this study we will be contrasting molecular techniques (DNA barcoding) with SEM and LM. Morphology alone to distinguish diatom species can be challenging because morphology of a species can be influenced by the life cycle stage and phenotypic plasticity in response to environmental conditions. The use of DNA barcoding is now more common for biodiversity studies of algae, although the feasibility of this method for water quality assessment has not yet been tested. This study will contrast the performance of barcoding and morphological methods to distinguish diatom taxa from fresh water samples.

Newer DNA-based methods combined with SEM have led to some revisions of the LM morphology-based taxonomy, including descriptions of new genera and species that are morphologically cryptic. Many researchers believe that DNA barcoding would be a valuable tool to provide a consistent identification of diatoms and make the data from different studies directly comparable, even if taxonomy changes. This protocol could provide information about the diversity of diatoms found in a river sample. Establishing a reliable diatom DNA barcoding protocol may provide a more uniform identification process. The resulting nucleotide data can be compared with existing, publicly available data that are archived through the Canadian Barcode of Life and National Center for Biotechnology Information (NCBI) database (a United States Government funded resource for molecular biology information) in order to determine accurate taxonomy. This information can be shared and made available for anyone interested in genomic regions of similarity between biological sequences and can also help to identify organisms (NCBI).

We will also investigate *Didymosphenia geminata*, an invasive sp of diatom. The recent discovery in 2011 of *Didymosphenia geminata* in the West Branch of the Farmington River (a tributary of the Connecticut River) has caused concern of the possibility of the impact to the ecology specifically the food web and biodiversity of the Connecticut River, and to the economy of the state of Connecticut if *D. geminata* were to persist, bloom and continue to expand its range. Invasive species such as *D. geminata* has cost over \$125-138 billion dollars in economic loss and ecological damage in the United States (Pimentel et al, 2000; Pimentel et al, 2005). *D. geminata* and its stalks are believed to have a negative effect on benthic macroinvertebrates and fish as well as other aquatic organisms and plants affecting our ecosystems (Spaulding and Elwood, 2007).

OBJECTIVES

The primary objective of this study is to apply DNA-based methods for identification of diatoms taken from environmental samples to assess water quality and contrast molecular with morphological analysis. The long-term goal is for diatoms to be used as biological indicators for monitoring and assessing watercourses including pollution, acidification, turbidity, salinization, nutrient loading and thermal alterations.

To achieve the primary goal of this study, we will specifically;

1. Extract DNA, amplify, and using the V 4 secondary structure of the ribosomal RNA gene 18S, clone, sequence and BLAST to identify diatom taxa.
2. Compare the performance of different methodologies LM, SEM and barcoding for diatom identification in an environmental sample.
3. Contribute new diatom DNA sequences to GenBank to help expand the diatom DNA database.
4. Long-term goal will be to establish a water quality protocol using both microscopy and molecular methods.

The primary objective of this section of the study is to apply DNA-based methods for identification of this potentially invasive diatom, *Didymosphenia geminata*, taken from direct samples and environmental samples to contrast molecular with morphological analysis. Molecular and morphological data of *D. geminata* found in Connecticut will be compared to the molecular and morphological data from other states.

Major goal of this study;

1. Work with the CT DEEP to perform surveillance of the West Branch of the Farmington River for evidence, absence or presence, of *Didymosphenia geminata*.

If *Didymosphenia geminata* is present:

2. Take direct visual samples *D. geminata* if available and environmental samples. Extract DNA, amplify, and using the V 4 secondary structure of the ribosomal RNA gene 18S, clone, sequence and BLAST to identify *Didymosphenia geminata*.
3. Make molecular comparisons of *D. geminata* found in other states with samples found in Connecticut.
4. Compare the performance of different methodologies LM, SEM and barcoding for *D. geminata* identification in an environmental sample.
5. Contribute *D. geminata* DNA sequences to GenBank to help expand the diatom DNA database.
6. Long-term goals include determining if *Didymosphenia geminata* has the characteristics to become a nuisance species (i.e. physical, geographical, and chemical conditions) in the Connecticut River and if so the cause(s) of *D. geminata* invasiveness and formulate ways of controlling the spread of this species.

METHODS

Sampling Diatoms in a River

A location is chosen where the substrate is no more than 33 cm/13 inches deep. The sampling location must have reasonable flow with a cobble substrate since **epilithic** diatoms are found in this habitat and are the desired specimen for the study organism. Random collection of 5-6 cobbles that can fit without difficulty in a Whirl-Pak will be collected. The cobbles will be scrubbed with a clean toothbrush and rinsed with distilled water into 1 Liter container to take back to the lab for analysis. Each sample will be divided, stored at 4°C until processing for morph examination. The other sample will be processed immediately for extraction or will be stored in -20°C until processing can occur. This sample will be used for sequence collection and DNA analysis.

PCR, Cloning, and Sequencing

The slurry sample will be homogenized then split into several 50 µL eppendorf microtubes for further analysis while several microtubes will be stored in -80 freezer to preserve for future use. From the sample, extraction of the DNA will be accomplished by using the MoBio PowerLyser Soil Extraction kit. Amplification of the target gene, 18S V 4, will then be achieved. PCR products will be visualized on Syber Safe stained gel, cleaned using Clontech NucleoSpin Kit, and DNA quantitated by the Nano Drop spectrophotometer. Cloning will be performed using Invitrogene TOPO TA; following the procedures as directed. Plasmid Prep will follow Qiagen mini prep procedures using the QIAprep® Spin Miniprep Kit. Sequences will be analyzed using the ABI3100 analyzer. Resulting sequences will be compared to GenBank, the public sequence database, using BLAST to obtain an accession number.

The methods of Zimmermann et al. (2011) and Jahn et al. (2008) will be used to examine the efficacy of a collective approach for molecular analysis and contrasting LM, EM, and DNA examination of diatoms.

In Jahn et al. (2008) environmental samples of diatoms were collected in freshwater streams. The DNA was extracted and amplified followed by cloning and sequencing. Jahn et al. was able to identify diatoms using the 18S region although finding a gene region that could differentiate between diatom species needed to be established. In Zimmermann et al. (2011), Zimmermann developed diatom-specific primers for the V4 region of the 18S rDNA gene. This region is variable among eukaryotes and provides an appropriate level of signal in diatoms to differentiate among taxa. I will be using Zimmermann's primers, D512 forward (ATT CCA GCT CCA ATA GCG) and D978 reverse (GAC TAC GAT GGTATC TAATC). Zimmerman was able to amplify and sequence with success 100% of the time with the fragment length approximately 390 and 410 bp (Zimmermann et al., 2011).

I will continue to clone colonies and sequence until I have reached a plateau of the number of different species of diatoms obtained is evident on the species accumulation curve. A phylogenetic tree was created for alignment and analysis (see figure 1). The phylogenetic analysis indicates that we were able to successfully extract and amplify DNA and obtain signal from sequencing.

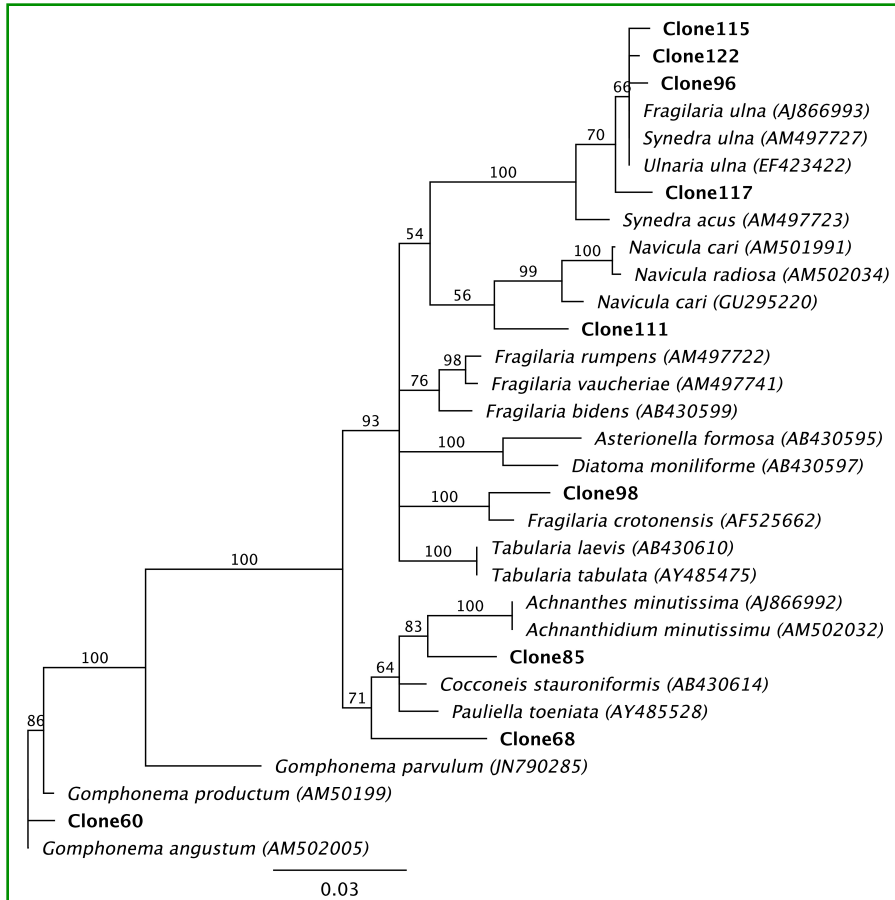


Figure 1: Phylogenetic Tree of sequences of the 18S V4 data set. The clones, 60, 68, 85, 96, 98, 111, 115, 117, and 122 are samples from an environmental sample from the Eightmile River in Lyme. Their DNA from the 18S gene was extracted, amplified, cloned and sequenced. Once this was accomplished the clones were BLASTED. Close BLAST matches from the NCBI database were included in this tree.

Morphological Analysis

Concentrated river samples will be simmered in a 1:1 ratio of water and nitric acid to oxidize organic matter. A hot plate will be used to simmer the sample after which the sample will be cooled. Acid will be rinsed with deionized water to dilute the acid, and centrifugation will be used for concentration following the protocol of Lowe (personal correspondence, 2009). The siliceous remains; the frustules, will then be prepared for slides.

LM preparation

After neutralizing and drying diatom samples, frustules will be mounted in NAPHRAX® (mounting medium), and placed on a hotplate until the solvent evaporates. The slide will be cooled immediately, by removing it from the hotplate. The diatom slide obtained is then permanent. Cleaned and prepared frustules will be examined by using an Olympus CX41 Phase Microscope and photographed using an Olympus MicroFire Digital Camera and a MicroFire CCD Color Camera at 1600x 1200 Pixels.

SEM Preparation

Aliquots of each slurry sample will be dried onto several pieces of aluminum foil which will be trimmed to fit and mounted on stub with double sided tape (Morales et al. 2001; Morales, personal correspondence 2010). Following Sputter Coat Instructions and using the Morales methodology (Morales et al. 2001) the stubs will be prepared by being coated for 30 seconds at 1.8 kV with gold/palladium using a Polaron sputter coater. The stubs can then be viewed with the Leo/Zeiss DSM 982 scanning electron microscope (SEM). In this study, the primary references for diatom taxonomy and morphology will be Krammer and Lange-Bertalot (1988), Round et al. (1990), and two online databases, ANSP Algae Image Database (<http://diatom.acnatsci.org/AlgaeImage/>) and the Great Lakes Image Database: (<http://www.umich.edu/~phytolab/GreatLakesDiatomHomePage/top.html>). These databases are in continual development and are always being updated.

PHYSICAL AND CHEMICAL ANALYSIS

Physical and chemical properties of river water, the pH, temperature, nutrients, and conductivity, will be measured using YSI meters. The physical characteristics of the river will be described such as water depth, geomorphology, riparian zone description, canopy and other physical attributes that could influence diatom species and population. Sampling of stream micro algae will take place when water table conditions are acceptable unless otherwise specified due to inclement environmental conditions. The USGS has sampling stations at 2 locations on the Eightmile River and that information will be utilized as needed. Samples will be collected and placed on ice in a cooler and transported to the laboratory for processing. The long-term goal will incorporate data collected to be used to establish the relationship of the species of diatoms found in the river with diatom autecology and correlate diatom species absence or presence with pH and chemical influences and other variables by using statistical analysis tools, ANOVA and multiple regression correlation models (Potapova, 2007).

PRELIMINARY RESULTS

In our initial investigation we were able to extract DNA and amplify the 18S V 4 marker. We cloned the amplicon and sequenced with success. The NCBI BLAST tool identified our sequences as diatoms.

The preliminary studies in the Lewis lab at the University of Connecticut, show promise but also indicate that further methodological refinement is needed. Fresh water samples subjected to DNA extraction yielded an abundant amount of genomic DNA, which was subsequently used to PCR amplify. We employed diatom-specific nucleotide primers for the V 4 secondary structure of the 18S region. The amplified PCR fragments were sequenced and yielded diatom-specific signal that was verified through BLAST comparisons against the public NCBI sequence database to matched published diatom sequences; *Synedra sp.*, *Frustrulia sp.*, *Gomphonema sp.*, *Cocconeis sp.* and *Navicula sp.*.

Microscopic examination of the samples verified the presence of several species. Ten diatoms were verified to sp. and 3 taxa* identified to sp. with some ambiguity due to LM. These include: *Tabellaria flocculosa*, *Discotella stelligera*, *Gomphonema parvulum*, *Nitzschia hantzschiana**, *Achnanthes*

pusilla, *Synedra goulardii*, *Achnanthes chlidanos**, *Gomphonema angustatum*, *Navicula rhyncocephala*, *Gomphonema truncatum*, *Tabellaria flocculosa*, *Synedra ulna*, *Frustrulia krammeri**.

Despite this success, however, further methodological development is needed. Preliminary results indicate that the 18S V 4 region amplifies well. We will continue to examine the utility of 18S V 4 markers to identify diatoms.

OUTCOMES AND SIGNIFICANCE

In this study, we assess barcoding methods in a broad taxonomic spectrum of diatoms from environmental samples from the Eightmile River and the Connecticut River in Connecticut. The results contrasted with light microscopy (LM) and scanning electron microscopy (SEM). Such comparisons have been made previously but with limited taxon sampling and with the use of cultures rather than environmental samples.

The use of molecular analysis in conjunction with LM and SEM could provide useful and new information about the diversity and the identity of diatoms found in a river sample. DNA-based methods combined with SEM have led to some revisions of the LM morphology-based taxonomy, including descriptions of new genera and species that are morphologically cryptic. DNA based methods also have been used to examine and probe the accuracy of diatom identification using morphology alone. Many researchers believe that DNA barcoding would be a valuable tool to provide a consistent identification of diatoms and make the data from different studies directly comparable, even if taxonomy changes.

Establishing a reliable diatom DNA barcoding protocol may provide a more uniform identification process. The resulting nucleotide data can be compared with existing, publicly available data that are archived through the Canadian Barcode of Life and National Center for Biotechnology Information (NCBI) database (a United States Government funded resource for molecular biology information) in order to determine accurate taxonomy. This information can be shared and made available for anyone interested in genomic regions of similarity between biological sequences and can also help to identify organisms (NCBI).

Since diatoms are important bioindicators and are often used for routine water quality assessments it is crucial to identify diatoms accurately (Pniewski et al., 2010; Morales et al., 2001, Mann et al., 2010) but it is also important that a cost effective and efficient method be developed to accurately identify diatoms (Jahn et al., 2007; Kaczmarska et al., 2007). The few researchers currently working with diatom DNA amplification find the work challenging for several reasons including the necessity to expand the taxon reference libraries and sequences in Genbank and the refinement of laboratory protocols for optimal extraction and amplification methods (Jahn et al., 2007; Evans et al., 2007). If this method can be consistently applied and optimized, it may have a significant impact on the accurate identification of these important organisms and their use as water quality bioindicators.

CHALLENGES

Throughout the extraction and PCR reaction process, inconsistent bands of varying brightness and amplification were noted. After extraction the DNA bands were often sheared and were light in color. PCR amplification could only be accomplished directly after successful extraction was realized. This led to inefficient lab work. Inhibitor problems were subsequently confirmed and corrected after working with a scientist and designer of the MOBIO extraction kits. We determined that using an “up-front” inhibitor removal method would help with extraction and amplification issues. Consequently, PCR amplification has been repeated several times consistently with acceptable results for both rbCL and 18S regions.

Cloning has been a challenge since it is a “hands on” applied technique. With this technique there is a learning curve which can require an extensive period to learn the process and become proficient in cloning methods. Practicing this technique is essential for optimum cloning and results. I continue to utilize the cloning kits to enhance my technique and gain exposure and knowledge of the cloning process. I have had several cloning kits that were ineffective in producing the desired results and required new kits consequently creating hours of trouble shooting and starting the PCR and cloning procedure over again. Results are improving appreciably as further refinement of appropriate materials and amounts used in the process and methodology of cloning are employed.

TIMELINE

Cloning will continue to ensure a consistent cloning methodology. The sampling, LM, SEM, molecular work, cloning and sequencing will be accomplished within the next year and a half 2014 and data analyzed over a six month period thereafter. This research project should take approximately one and a half years to two years to complete if there are no major obstacles and under good conditions.

Poster Presented at the 34th Annual Meeting & Environmental Conference of the Connecticut Association of Conservation & Inland Wetlands Commissions, November 12, 2011, Wallingford, CT

APPLICATION OF THE V4 REGION OF THE 18S RRNA GENE TO DNA BARCODE DIATOMS (BACILLARIOPHYTA) FROM ENVIRONMENTAL SAMPLES OF THE EIGHTMILE RIVER IN CONNECTICUT.

Diba Khan-Bureau¹, Louise Lewis¹, Gary Robbins¹
¹ University of Connecticut, Storrs, CT 06269, U.S.A.

Diatoms are ubiquitous and ecologically important. Diatoms are accepted as biological indicators for monitoring and assessing watercourses but can be used additionally to evaluate other important ecological questions. Various approaches have been used to identify diatoms and more recently the use of molecular analysis has been employed. Using morphology alone to distinguish diatom species can be difficult, thus many researchers have proposed the use of DNA barcoding to provide consistent identification of diatoms and make the data from different studies directly comparable, even if taxonomy changes. In this investigation, we assess barcoding methods in a broad taxonomic array of diatoms from environmental samples from a river in Connecticut. We continue to contrast light microscopy (LM), scanning electron microscopy (SEM) and molecular approaches to estimate diatom diversity. Comparisons have been made previously but with limited taxonomic capacity and with the use of cultures. The use of molecular analysis in conjunction with LM and SEM could provide useful and new information about the diversity of diatoms found in a river sample. In this study we have extracted DNA and PCR amplified using the diatom specific 18S V4 marker region. We cloned the amplicon and sequenced with success. We continue to collect data to support and illustrate phylogenetic diversity. When we used the NCBI website we were able to BLAST our sequences to diatoms. We will continue to examine the practicability and efficacy of the 18S V4 marker for accurate identification of diatoms. If barcoding can consistently be applied effectively and economically, the use of this procedure may have a significant impact on the accurate classification of these important organisms as water quality biological indicators. Environmental sampling and DNA barcoding shows promise as one of the tools in the taxonomic and phylogenetic toolbox.

Application of the V4 Subregion of the 18S rRNA to DNA Barcode Environmental Samples for Diatom Identification in the Eightmile River



Diba A. Khan-Bureau¹, Louise A. Lewis², Gary A. Robbins³
 University of Connecticut, NRE^{1,2} and EEB² Department, Storrs, CT 06269 USA
 diba.khan-bureau@uconn.edu⁴



INTRODUCTION AND BACKGROUND

Diatoms are important biological indicators of water quality. They are significant since they are considered the base of the food web in aquatic ecosystems and are a carbon sink (1). Diatom identification can be difficult since there are hundreds of thousands of species including many that are undescribed. Additionally, phenotypic plasticity and cryptic species can confuse morphological identifications and estimates of diatom biodiversity (2).

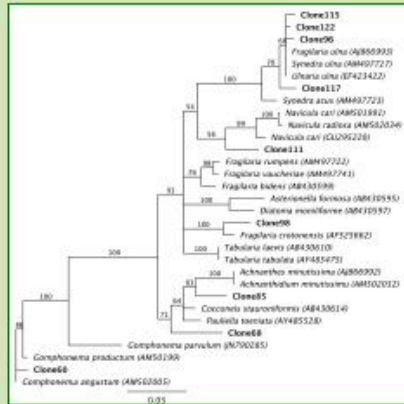
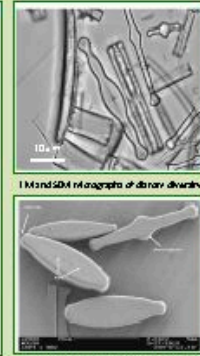
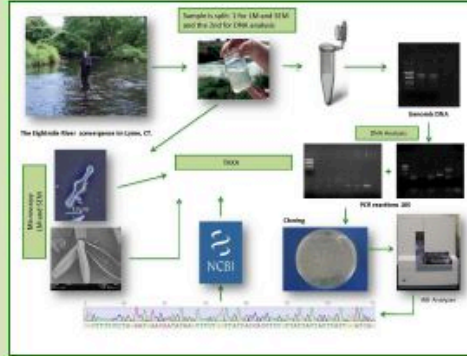
Amato et al. (2007) contrasted the resolving power of Light Microscopy (LM), Scanning Electron Microscopy (SEM), and DNA barcoding to identify diatoms (3). They used diatom specific markers with limited taxonomic resolving abilities. In this study we will employ the approach of Amato et al. using environmental samples without culturing. We will evaluate the feasibility of obtaining representative diatom sequences from environmental samples. This study will contrast LM, EM and DNA sequencing for identifying field collected diatoms using the 18S gene region. Zimmerman et al. (2011) evaluated the V4 subregion of the 18S rRNA gene and developed new primers and protocols which we have incorporated into our study (4). If diatoms are used as biological indicators of water quality then proper identification is essential.

MAJOR QUESTIONS AND GOALS

- > **Questions:**
 - Can environmental sampling provide sufficient diatom DNA for barcoding?
 - Is the 18S V4 region variable enough to differentiate species of interest?
- > **Goals:**
 - Refine known primers for environmental sampling of diatoms of interest.
 - Construct PCR, amplify, clone, and sequence specimens to differentiate and identify numerous diatom species in an environmental sample.

METHODOLOGY

> Sampling was conducted at the convergence of the East and Main Branches of the Eightmile River in Lyme Connecticut. Samples were split into two; one sample was used for LM and SEM. The second sample was used for DNA extraction followed by amplification, cloning, sequencing, and NCBI BLAST.



Neighbor-joining tree based on analysis of the diatom 18S V4 sequences observed from the Eightmile River. Labels are "clone" IDs assigned with the diatom name for observed from the NCBI database (with associated GenBank accession numbers indicated in parentheses). Below LM micrograph of diatom diversity from the riverbank.



PRELIMINARY RESULTS

- > Extractions were accomplished using MD BIO Powerfood and Powerizer Soil after difficulties with inhibitors requiring upfront anti-inhibitor extraction kits.
- > Successful and repeatable amplification was obtained using primer pairs D812F 18S and D878R 18S.
- > Successful amplification products were sequenced using MJ3 derivatives and matched published diatom sequences.
- > Sequences were aligned using MUSCLE (Edgar 2004) to construct neighbor-joining tree (5).

DISCUSSION

- > Further development of primers and cloning techniques of an environmental sample will continue to ensure efficiency and accuracy.
- > Barcoding from environmental samples has potential to become more reliable and less time consuming with more data and experience.
- > Molecular analysis of environmental sampling is becoming a successful method for diatom identification and could become the preferred method for aquatic biodiversity assessment.

REFERENCES

(1) Stevenson, R. J. and T. Pan. 1999. Assessing ecological conditions in rivers and streams with diatoms. Pages 11-40 in Swamer, C.F. and Smal, J.R. editors. The Diatoms: Applications to the Oceanic and Terrestrial Environments. Cambridge University Press, Cambridge, UK.
 (2) Round, F.E., Crawford, R. M. and Howe, D. G. 1990. The Diatoms: Biology and Morphology of the Genera. Cambridge University Press, Cambridge, 747 pp.
 (3) Amato, A., Watabe, J.C.T.R., Leibold, C.H., Chiriac, David G., Mann, P., Prashad, T., and Mironov, M., 2007. Reproducible diatom among sympatric Cryptosporidium Inhabitant Diatoms. Protist, Vol. 158, 190-207.
 (4) Zimmerman, J., John, R., Gopal, M., et al., 2011. Barcoding diatom evolution of the V4 subregion on the 18S rRNA gene, in developing new primers and protocols. Open Access Journal of Diatoms DOI: 10.1080/1522-0224.2011.602946
 (5) Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32: 1792-1797.



ACKNOWLEDGMENTS

The Lewis lab; Shelley Elm, Dr. Indy Leach, Dr. Carolina Polakow, Larry Turley, SIMON-G-CI Three Rivers Community College A.S. scholar; Dr. Jerry Sgro, John Carroll University Ohio; Dr. Nic Louie, Bowling Green University for those who have financially supported this research; The Eightmile River Wild and Scenic Watershed Committee and the Connecticut Institute of Water Resources (CIWR).
 *FCP-31 BP Green support

