1-29-2015

Two Novel Diatoms in Connecticut and Comparisons of Morphological and Molecular Approaches of Species Diversity Estimation

Diba A. Khan-Bureau
UCONN, diba.khan-bureau@uconn.edu

Two Novel Diatoms in Connecticut and Comparisons of Morphological and Molecular Approaches of Species Diversity Estimation

Diba Khan-Bureau, PhD

University of Connecticut, 2015

ABSTRACT

Diatoms (Bacillariophyta) are ubiquitous and can be found wherever there is water and are ecologically important eukaryotic microalgae. Because many diatom species have been shown to be associated with particular environmental conditions, these taxa are accepted as biological indicators for assessing water quality. In order to address water quality and other applications using diatoms, accurate taxonomic identification is essential. The dominant approach used to identify diatom species is morphological characterization with light (LM) and scanning electron microscopy (SEM). However, using morphology alone to distinguish diatom species can be challenging because the phenotype of a species is often influenced by the life cycle stage and the environment. DNA barcoding is a method that compares a short section of a genome region. There is an increasing use of DNA barcoding for biodiversity studies, although the information provided by DNA barcoding of diatoms has not yet been compared with that from morphology, except from cultured material. This research contrasted the performance of DNA barcoding and morphological methods to distinguish diatom taxa in a freshwater sample of the Eightmile River, Connecticut. The research examined the utility of DNA barcoding to identify and document the presence of nuisance diatoms *Cymbella janischii* (A. Schmidt) De Toni and other stalk forming diatoms in The West Branch Farmington River, Connecticut and reports on a putatively new species in the genus *Didymosphenia*.

Two Novel Diatoms in Connecticut and Comparisons of Morphological and Molecular Approaches of Species Diversity Estimation

Diba Khan-Bureau

A.S., Thames Valley State Technical College, 1994

B.S., University of Connecticut, 1996

M.S., Rensselaer Polytechnic Institute, 2001

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2015

APPROVAL PAGE

Doctor of Philosophy Dissertation

Two Novel Diatoms in Connecticut and Comparisons of Morphological and Molecular Approaches of Species Diversity Estimation

Presented by

Diba Khan-Bureau, A.S., B.S., M.S.

Major Advisor

Louise A. Lewis

Associate Advisor

Michael Dietz

Associate Advisor

Richard Anyah

Associate Advisor

Rex Lowe

University of Connecticut
2015

ACKNOWLEDGMENTS

This research was financially supported by the Connecticut Institute of Water Resources (CTIWR) grant no. 5606350 at UCONN College of Agriculture and Natural Resources.

This dissertation would not have been possible without the support from my children, friends, colleagues and so many others. I am grateful to all of the undergrads, graduate students and post docs in the Ecology and Evolutionary Biology (EEB) department. They have been true supporters of my endeavors through thick and thin. I am grateful for the support that I received from Three Rivers Community College (TRCC) where I am a professor. Without the support from my administration and in particular Dr. Grace Sawyer Jones, TRCC president for many years, I could not have completed this work. I am indebted to Dr. Louise Lewis for taking me into her lab as an EEB student when I was in the Natural Resources and the Environment department. Louise was patient and believed that I was doing important work even when I questioned myself. If it weren't for Dr. Frank Trainor I would not have met Dr. Lewis. It was he who told me that if I really wanted to know about micro-algae and understand DNA barcoding that I would need to meet Dr. Louise Lewis. Dr. Trainor promised me that he would be around to give advice when I needed it, and he did until his passing in 2013. I am forever grateful to my committee members, Drs. Mike Dietz, Richard Anyah, and Rex Lowe. They were always there to support and guide me. I thank my collaborators and co-authors Dr. Eduardo Morales, Luc Ector, and Mike Beauchene for their guidance.

My lab-mates, some graduated and some not, have been a tremendous influence and have inspired me and encouraged me to keep moving forward even during difficult times. Drs. Molly Letsch, Hilary McManus, Karolina Fučíková have all been instrumental in my research. I thank Shelley Olm and Chien Lo for their invaluable lab assistance. I will not forget the valuable mantra Shelley would say, “Follow the lab protocols even though you have performed these hundreds of times,” and her continuous support. My new lab mates Nic Schultz and Jordan Bishop have been wonderful additions to the lab especially for their coffee making abilities and their love of micro and macro algae discussions.

I particularly want to thank the Simon Lab’s Dr. Beth Wade, Ph.D. students Geert Geomans and Russ Meister for their frequent and thoughtful discussions. They were always available for a hug too when needed and to Lily Lewis and Ursula King as well as many other under graduate, graduate students and post docs who made me feel at home.

There are EEB faculty who have supported me because they are true teachers. Dr. Elizabeth Jockusch was an invaluable mentor. She always had time to discuss the ABI, sequencing methods and cloning methods, and almost anything else that had to do with molecular work. Dr. Eldridge Adams was always available to chat about R and statistics and thanks to Dr. Robert Colwell for walking me through EstimateS even when he was away in Madagascar. Dr. James Green and Dr. Mike Dietz were also so very helpful with the statistics package SPSS for my 2nd chapter.

My colleagues from TRCC were tremendous supporters and I must name them. My friends Kacey McCarthy-Zaremba, Marge Valentin, Jessie McCoy, faculty Dr. Terry Delaney, Jon Brammer, and Kent Harding were all supportive and always had time to make me laugh. To my union, the AFT, who supported my sabbatical in the fall of 2009, the IT department's Kem Barfield, Skye Cohen, Dean Steve Goetchius, and Diane Jewett from administrative services for their support and Luba Bugbee at the UCONN Graduate School for her continuous reinforcement of my efforts.

I thank those who have been in my life and left to soon, my dad, David and Jeffery. I miss them and love them. I thank Paul for helping with the kids scheduling. And most of all I thank my family and my children, Michaela Angel and John Michael, for their patience with their mom. I felt at times that I didn't pay enough attention or that I should have been playing with them when I had to get research work done. They have shown me what unconditional love is all about. I dedicate this to them.

TABLE OF CONTENTS

Testing DNA Barcoding Methods for Diatom (Bacillariophyta) Identification from Environmental Samples of the Eightmile River and the Farmington River in Connecticut

INTRODUCTION

Diatoms are ubiquitous and ecologically important eukaryotic microalgae. They are microscopic, autotrophic eukaryotes in the phylum Bacillariophyta (Gibbs et al., 1981; Saunders et al., 1995; Medlin et al., 1996; Medlin et al., 2000). Estimates of the number of diatom species range from 10,000 to over 1 million (Mann and Droop, 1996; Van Den Hoek et. al., 1997; Mann, 1999; Mann et al., 2010). Many diatom species have yet to be discovered while others have been studied extensively, particularly those species that are invasive, nuisance, or are toxic and form harmful algae blooms (HABs). Diatoms are estimated to make up more than half of photosynthetic production on Earth in both freshwater and marine environments (Medlin et al., 1991; Mann, 1999; Stoermer and Smol, 1999; Jahn et al., 2007; Zimmermann et al., 2011).

Diatoms and other algae are important contributors to energy flow. They cycle nutrients in surface waters and other water bodies (Rott, 1991; McCormick and Cairns, 1994; Mann and Droop, 1996; Stevenson and Pan, 1999). Diatoms have global ecological significance in the carbon and silicon cycles, making them important to all life on Earth (Sgro and Johansen, 1995; Mann and Droop, 1996; Van Den Hoek, 1997). Because diatoms fix atmospheric CO_2 through photosynthesis, they are at the base of the food web and are essential within aquatic ecosystems. Additionally, diatoms have been used to evaluate other important ecological questions such as global warming, anthropogenic disturbances and are presently being studied for use in the biomedical and nanotechnology fields. With the growing need for alternative energy sources, algae such as diatoms have been studied for their potential use as biofuels. Thus, diatoms are considered one of the most ecologically and economically important group of eukaryotic

microorganisms in the environment (Moniz and Kaczmarska, 2009).

Diatom climatological and paleontological research. Diatoms are employed in paleontology and climatology (McCormick and Cairns, 1994; Srgo and Johansen, 1995; Stoermer and Smol, 1999). Diatom walls (frustules) are made of silicon dioxide (SiO_2), making the remains of diatoms less prone to degradation than the walls of many other algae. With diatoms growing, blooming, and then dying off, the frustules accumulate on the bottom of sea floors, lake and river bottoms, leaving a record wherever they thrive (Stoermer and Smol, 1999; Rühland et al., 2003). The frustules of diatoms remain in fossil deposits for long periods of time and are typically well preserved for further investigation to help determine environmental conditions of the past (Stoermer and Smol, 1999; Rühland et al., 2003; Rühland et al., 2008). The frustules are morphologically complex and under SEM can be better characterized, leading to more accurate species identification. This can help tell us about the past environmental conditions and provide information about anthropogenic influences on the landscape (McCormick and Cairns, 1994; Stevenson et al., 1997; Mann, 1999; Rühland et al., 2003; Rühland et al., 2008). Since diatoms are prolific and are found in every kind of water system including moist soils they can be used as an indicator of the past (Stevenson and Lowe, 1986; Stevenson and Pan, 1999; Mann, 1999; Jahn et al., 2007; Stevenson and Sabater, 2010).

Changes in diatom stratigraphy are correlated to human perturbation by land clearing, agriculture, pollutants, nitrogen and phosphorus loading (Cole, 1979; Stevenson et al., 1997; Smol and Stoermer, 2008; Rühland et al, 2008). Diatom fossils are found in sediments of marine, lakes, wetlands and other water bodies. The core samples of sediments can provide information such as global environmental problems, acidic lake changes, climate change and eutrophication concerns which have impacted these sensitive organisms (McCormick and Cairns, 1994;

Stoermer and Smol, 1999; Rühland et al., 2008). Diatom assemblages respond to nutrient availability, lake or sea level changes, and climatic changes, and can be a reliable tool for assessing long term environmental changes with biodiversity (Rühland et al., 2003; Smol and Stoermer, 2008). Diatoms are sensitive to changes in temperature, water chemistry and desiccation, and can therefore tell us something about their geographical distribution and the past environment that they lived in by presence or absence in a particular ecosystem (Stevenson and Lowe, 1986; McCormick and Cairns, 1994; Stevenson and Pan, 1999; Stoermer and Smol, 1999; Gold et al., 2002; Van Den Hoek et al., 2007).

Diatoms and water quality assessment. Diatoms are distributed in all types of aquatic and terrestrial environments. In the water, diatoms can be planktonic, moving with the current and found in the water column. They may be epiphytic, attaching to submerged plants including macrophytes and larger diatoms, or epilithic, attaching to pebbles, rocks, and other hard surfaces, and as epipelons, which are found in sediments (Stevenson and Pan, 1999). Water systems, lentic and lotic, freshwater and marine, can be described by several features including hydrology, chemistry, biology and physical characteristics (Lowe, 1974; Srgo and Johansen, 1995; Chapman, 1996; Stevenson and Pan, 1999; Stoermer and Smol, 1999). Water quality assessment can be based on appropriate monitoring of these attributes. The physical and chemical characteristics of rivers and streams are influenced by geomorphology and climate within a specific watershed region. These influences affect mineral content, pH, temperature, and nutrient cycling as well as nutrient loading from anthropogenic causes which sequentially will have an effect on the biological communities within the river ecosystem (Chapman, 1996; Stevenson and Pan, 1999; USEPA, 2000; Stevenson et al., 2008).

In the late 1800's, at the start of the industrial revolution, environmental and water

quality degradation was becoming more evident. Governmental agencies began taking action in Europe and in the United States but specifically in London and Paris after many of the population became ill and died from cholera and typhoid outbreaks (Garcier, 2010; Anfinson, 2010). Under the Clean Water Act, the 1972 amendments to the Federal Water Pollution Control Act, all states in the U.S. developed, adopted and currently regulate water quality standards which include biological, chemical and physical parameters (water quality criteria) to support living organisms specific to their habitat and geographical distribution and to access the biological integrity and condition of aquatic life (Stevenson and Pan, 1999; U.S.A Clean Water Act, USC 111251–1387).

Water quality degradation is a major concern since the human population has increased significantly and forested areas, known as our ecosystem service filtration systems, are being stressed and in some cases depleted. There is concern for the increasing agricultural industries that stockpile manure and use pesticides, fertilizers, and herbicides. Nutrient enrichment is considered one of the most problematic issues effecting our rivers and streams, while other anthropogenic activities such as storm water runoff or non-point source pollution, erosion and sedimentation are of serious consequence; contributing to water quality degradation (USEPA, 2000; Potapova and Charles, 2007; USGS, 2010). Diatoms can be used to help us make better land-use decisions, and incorporate beneficial management practices when managing drinking water and fragile water ecosystems, which are all part of surface and groundwater bodies. Using chemical, physical and biological indices to measure the health of an ecosystem collectively can help to protect our environment and natural resources. Employing biological indicators has been shown to be a reliable method for water quality assessments (Mann and Droop, 1996; Stevenson and Pan, 1999; Van Den Hoek et al., 2007). Diatoms have known environmental tolerances of

water quality and to physical conditions. Diatoms are good indicators because they have high reproductive rates, they are found in every environment and are especially abundant where there is water and are found throughout the world, and particular species are tolerant of specific types of pollutants, chemical conditions and physical properties (Lowe 1974; Rott, 1991; Sgro and Johansen, 1995). Temperature, conductivity and nutrient (N and P) levels will affect occurrence and relative abundance of species in a river (Stevenson et al., 2006; Walker and Pan, 2006). Therefore, linking specific diatom species to the chemical and physical properties of a river, stream, or water body is essential.

As early as 1908, living organisms were used to measure ecological health (Kolkwitz and Marson, 1908). These authors introduced the saprobic system, the first biotic index for water quality assessment (Lowe, 1974; Chapman, 1996; Srgo and Johansen, 1995; Stevenson and Pan, 1999). Kolkwitz and Marrson (1909) proposed that biota were sensitive to levels of pollutants and other human disturbances (Srgo and Johansen, 1974; Stevenson and Pan, 1999; Gracier, 2010; Sharma and Sharma, 2010). The saprobic system has been updated by many investigators, and is still used in one form or another (Srgo and Johansen, 1995).

The U.S. Environmental Protection Agency (USEPA) published the first guidelines for a national biological criteria program in 1990 (USEPA, 1990). Presently, most states employ benthic macro-invertebrates (e.g. insect larvae, crustaceans, flat worms, mullosks, and annelids) in their water quality monitoring program because this group is widely used in bio-assessments (Barbour et al., 1999; Wright, et al., 2000). Some states use macroalgae, and only a handful of states use diatoms (micro-algae). Macroalgae have been used by the USGS for many years in their National Water Quality Assessment (NAWQA) standard (USGS, 2006). Presently the USGS NAWQA program is compiling information in order to compare information on diatoms

with and U.S. river nutrient levels (USGS, 2006; Potapova and Charles, 2007). Algae and benthic macro-invertebrates have specific environmental tolerances and sensitivities to pollutants and therefore have been used as compliance tools to keep within the CWA water quality criteria (Gold et al., 2002; Stevenson and Sabater, 2010). The European Union (EU) has water quality standards termed the Water Framework Directive (WFD), which are similar to the United States water quality standards (Blanco and Bécares, 2010; Stevenson and Sabater, 2010). The EU WFD states that by 2015 all rivers, streams, lakes and other water bodies are to have good ecological status, having the same connotation as "high biological condition" described by the United States (European Union WFD, 2000; Stevenson and Sabater, 2010; Blanco and Bécares, 2010). In addition, impoundments, small and large dams, bridges, culverts and other impediments influence diatom populations. Low flow conditions in the summer and winter months and high flow conditions in the spring and fall will influence diatom population and species, giving us a broader view of the impact of cyclical and anthropogenic perturbations (Bormans and Webster, 1999; Stevenson and Pan, 1999; Gold et al., 2002).

Several states, including Connecticut, are investigating the utility of diatoms as biological indicators of water quality. Diatoms have been studied for many years but are often not fully utilized (Srgo and Johansen, 1995; Evans et al., 2007). Recently, the Connecticut Department of Energy and Environmental Protection (CT DEEP) assessed diatom morphology in the Eightmile River (Ernest Pizzuto & Mary Becker, CT DEEP, pers. comm.). The CT DEEP envisions integrating diatoms as biological indicators in their water quality assessments. The United States Geological Survey (USGS) in East Hartford also has explored the use of diatoms as biological indicators in Connecticut's rivers and intends to further employ diatoms as assessment tools (Jonathan Morrison, USGS, pers. comm.).

In 1974, Lowe compiled data on 300 diatom species sensitivities and tolerance levels to pollution and environmental change. He further investigated the autecology of fresh water micro-algae to better serve as biological indicators of water quality (EPA, 1974; Lowe, 1974). In 2002, Stevenson continued to utilize diatoms as biological indicators of water quality for freshwater wetlands (EPA, 2002). The dominant approach used to identify diatom species is morphological characterization with light (LM) and scanning electron microscopy (SEM). However, using morphology alone to distinguish diatom species can be challenging because the phenotype of a species is often influenced by the life cycle stage and the environmental conditions of its habitat.

CONCLUSION

The goals of this study were to 1. Employ morphology based approaches and DNA barcoding to ascertain if these methods have comparable resolving power for diatom diversity and identification, 2. Use morphology and DNA barcoding methods to identify and distinguish two nuisance stalk-forming diatoms in the West Branch of the Farmington River, Connecticut. Currently, there is an increasing use of DNA barcoding for biodiversity studies, although the information provided by DNA barcoding of diatoms has not yet been compared with that from morphology, except from cultured material. This study contrasts the performance of DNA barcoding and morphological methods to distinguish diatom taxa from freshwater samples of the Eightmile River, Connecticut. The study examines the use of DNA barcoding to identify and document the presence of the putative invasive diatom species, *Didymosphenia* sp., and the nuisance diatom species, *Cymbella janischii*, in the West Branch of the Farmington River, Connecticut.

Diatom taxonomy is in a state of flux, however, as more is discovered and understood about the autecology and biology of these organisms their uses will be fully recognized and

accepted as indicators of water quality (Stevenson and Sabater, 2010). This is true for molecular approaches as well. With newer technologies, using DNA barcoding will become more conventional and better established as an accepted method to identify diatoms from environmental samples.

REFERENCES

Anfinson, J. 2010. Environmental History. 15 (4): 781-783. doi: 10.1093/envhis/emq094.

Barbour, M.T., Gerritsen, J., Snyder, B.D. and Stribling, J.B. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, 2nd Ed. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

Blanco, S., and Bécares, E. 2010. Are biotic indices sensitive to river toxicants? A comparison of metrics based on diatoms and macro-invertebrates. Chemosphere 79: 18–25.

Bormans, M., and Webster, I. T. 1999. Modeling the Spatial and Temporal Variability of Diatoms in the River Murray. Journal of Plankton Research 21: 581-598.

Chapman, D. 1996. Water Quality Assessments - A Guide to Use of Biota, Sediments and Water in Environmental Monitoring – 2nd Ed. UNESCO/WHO/UNEP. ISBN 0 419 21590 5 (HB) 0 419 21600 6 (PB).

Cole, G. 1979. Textbook of Limnology 2nd Ed. The C. V. Mosby Company.

Gibbs, S. P. 1981. The chloroplasts of some algal groups may have evolved from endosymbiotic eukaryotic algae. Annals of the New York Academy of Sciences 361: 193–208.

Gold, C., Feurtet-Mazel, A., Coste, M., and Boudou, A. 2002. Field transfer of periphytic diatom communities to assess short-term structural effects of metals (Cd, Zn) in rivers. Water Research 36: 3654–3664.

Jahn, R., Zetzsche, H., Reinhardt, R., and Gemeinholzer, B. 2007. Diatoms and DNA barcoding: a pilot study on an environmental sample. Proceedings of the 1st Central European Diatom Meeting. Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin ISBN 978-3-921800-63-8 BGBM, Berlin 63-68.

Lowe, R. L. 1974. Environmental requirements and pollution tolerance of freshwater diatoms. United States Environmental Protection Agency, Cincinnati, Ohio. EPA-670/474–005.

Mann, D. G., and Droop, S. J. M. 1996. Biodiversity, biogeography and conservation of diatoms. Hydrobiologia 336: 19–32.

Mann, D. G. 1999. Phycological reviews 18. The species concept in diatoms. Phycologia 38: 437-495.

Mann D. G., Sato, S., Trobajo, R., Vanormelingen, P., and Souffreau, C. 2010. DNA barcoding for species identification and discovery in diatoms. Cryptogamie, Algologie 31 (4): 557-577.

McCormick, P. V., and Cairns, J. Jr. 1994. Algae as indicators of environmental change, Journal of Applied Phycology, Kluwer Academic Publishers 6: 509.

Medlin, L. K., Kooistra, W. C., and Schmid, A-MM. 2000. A review of the evolution of the diatoms: a total approach using molecules, morphology and geology. In: Witkowski, A., and Sieminska, J. (eds). The origin and Early Evolution of the Diatoms: Fossil, Molecular and Biogeographical Approaches. W. Szafer Institute of Botany, Polish Academy of Sciences, Cracow, Poland 13–35.

Medlin, L. K., Kooistra, W. C., Gersonde, R., and Wellbrock, U. 1996 Evolution of the Diatoms (Bacillariophyta). II. Nuclear-Encoded Small-Subunit rRNA Sequence Comparisons Confirm a Paraphyletic Origin for the Centric Diatoms. Molecular Biology and Evolution 13 (1): 67-75.

Medlin, L. K., Elwell, H. J., Strickel, S., and Sogin, M. L. 1991. Morphological and genetic variation within the diatom *Skeletonema costatum* (Bacillariophyta): evidence for a new species, *Skeletonema pseudocostatum*. Journal of Phycology 27: 514-524.

Moniz, M. B. J., and Kaczmarska, I. 2009. Barcoding diatoms: Is there a good marker? Molecular Ecology Resources 9: 65-74.

Potapova, M., and Charles, D. F. 2007. Diatom metrics for monitoring eutrophication in rivers of the United States. Ecological Indicators 7: 48–70.

Rott, E. 1991. Methodological aspects and perspectives in the use of periphyton for monitoring and protecting rivers. In: Whitton, B. A., and Rott, E. (eds). Use of Algae in Monitoring Rivers II. Institut für Botanik, Universität Innsbruck. Innsbruck, Austria 9-16.

Rühland, K., Priesnitz, A., and Smol, J. P. 2003. Paleolimnological Evidence from Diatoms for Recent Environmental Changes in 50 Lakes across Canadian Arctic Treeline. Arctic, Antarctic, and Alpine Research 35 (1): 110–123.

Rühland, K., Paterson, A., and Smol, J. P. 2008. Hemispheric-scale patterns of climate-related shifts in planktonic diatoms from North American and European lakes. Global Change Biology 14: 2740–2754

Saunders, G. W., Potter, D., Paskind, M. I., and Andersen, R. A. 1995. Cladistic analyses of combined traditional and molecular data sets reveal an algal lineage. Proceedings of the National Academy of Sciences USA 92: 244-248.

Sgro, G. V., and Johansen, J. R. 1995. Rapid bioassessment of algal periphyton in freshwater streams. In: Butterworth F. M., (ed). Biomonitors and Biomarkers as Indicators of Environmental Change, Plenum Press, New York 291–311.

Sharma, S., and Sharma, P. 2010. Biomonitoring of Aquatic Ecosystem with Concept and Procedures Particular Reference to Aquatic Macro invertebrates. Journal of American Science 6 (12): 1246-1255.

Smol, J. P., and Stoermer, E. F. 2008. The Diatoms. 2nd Ed. Applications for the Environmental and Earth Sciences. Cambridge University Press, Cambridge, UK.

Stevenson, R. J., and Pan, Y. 1999. Assessing ecological conditions in rivers and streams with diatoms. In: Stoermer, E. F. and Smol, J. P. (eds). The Diatoms: Applications to the Environmental and Earth Sciences. Cambridge University Press, Cambridge, UK 11-40.

Stevenson, R. J., and Lowe, R. L. 1986. Sampling and interpretation of algal patterns for water quality assessment. In: Isom, B. G. (ed). Rationale for sampling and interpretation of ecological data in the assessment of freshwater ecosystems. ASTM STP 894. American Society for Testing and Materials, Philadelphia, Pennsylvania, USA 118–149.

Stevenson, R. J., Rier, S. T., Riseng, C.M., Schultz, R. E., and Wiley, M. J. 2006. Comparing effects of nutrients on algal biomass in streams in 2 regions with different disturbance regimes and with applications for developing nutrient criteria. Hydrobiologia 561: 149–165.

Stevenson, R. J., Hill, B. H., Herlihy, A. T., Luan, L. L., and Norton, S. B. 2008. Algae–P relationships, thresholds, and frequency distributions guide nutrient criterion development. Journal of the North American Benthological Society 27(3): 783–799.

Stevenson, R. J., and Sabater, S. 2010. Global Change and River Ecosystems – Implications for Structure, Function and Ecosystem Services. Hydrobiologia 657: 1-2.

Stoermer, E. F., and Smol, J. P. 1999. The Diatoms. Applications for the Environmental and Earth Sciences. Cambridge University Press, Cambridge, UK 488.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 1974. United States

Environmental Protection Agency (USEPA). Environmental Requirements and Pollution Tolerance of Freshwater Diatoms (EPA-670-4-74-005). Office of Research and Development, US Environmental Protection Agency, Washington, DC., USA.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 1990. United States Environmental Protection Agency (USEPA). Biological Criteria: National Program Guidance for Surface Waters (EPA-440-5-90-004). Office of Water, US Environmental Protection Agency, Washington, DC., USA.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000. United States Environmental Protection Agency (USEPA). Nutrient criteria technical guidance manual. Rivers and streams. EPA-822-B-00-002. Office of Water, US Environmental Protection Agency, Washington, DC., USA.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2002. United States Environmental Protection Agency (USEPA). Methods for evaluating wetland condition: using algae to assess environmental conditions in wetlands. EPA-822-R-02-021. Office of Water, US Environmental Protection Agency, Washington, D.C., USA.

United States Geological Survey (USGS) CT. website. 2008. http://ct.water.usgs.gov/nawqa/design.htm

United States Geological Survey (USGS) CT. 2008. Algal Attributes: An Autecological Classification of Algal Taxa Collected by the National Water-Quality Assessment Program.

United States Geological Survey (USGS) CT. website. 2010. Water Quality Information. http://water.usgs.gov/owq/.

Van Den Hoek, C., Mann, D. G., and Johns, H. M. 1997. Algae: An Introduction to Phycology. Cambridge University Press, Cambridge, UK.

Wright, J. F., Sutcliffe, D. W., and Furse, M. T. 2000. Assessing the biological quality of fresh waters: RIVPACS and other techniques. *Freshwater Biological Association,* Ambleside, UK.

Zimmerman, J., Jahn, R., and Gemeinholzer, B. 2011. Barcoding diatoms evaluation of the V4 sub region on the 18S rRNA gene, including new primers and protocols. Organisms, Diversity, & Evolution. DOI 10.1007/s13127-011-0050-6.

BioInvasions Records (2014) Volume 3, Issue 3: 139–149
doi: http://dx.doi.org/10.3391/bir.2014.3.3.02

Open Access

Research Article

Observations of two nuisance stalk-forming diatoms (Bacillariophyta) from a river in Connecticut, Northeastern U.S.A.: *Didymosphenia* sp. and *Cymbella janischii* (A. Schmidt) De Toni

Diba A. **Khan-Bureau**[1,2]*, **Michael S. Beauchene**[3], **Luc Ector**[4], **Eduardo A. Morales**[5] **and Louise A. Lewis**[6]

[1]*Department of Natural Resources and the Environment, University of Connecticut, 75 North Eagleville Road, Storrs Connecticut, 06269, USA*
[2]*Three Rivers Community College, 574 New London Turnpike, Norwich, Connecticut, 06360, USA*
[3]*Inland Fisheries Division, Department of Energy and Environmental Protection, 79 Elm Street, Hartford, Connecticut, 06106-5127, USA*
[4]*Department of Environment and Agro-biotechnologies (EVA), Public Research Center – Gabriel Lippmann, 41 rue du Brill, L-4422 Belvaux, Luxembourg*
[5]*Herbario Criptogámico, Universidad Católica Boliviana San Pablo Calle M. Márquez esq. Plaza Jorge Trigo s/n, P.O. Box 5381, Cochabamba, Bolivia*
[6]*Department of Ecology and Evolutionary Biology, University of Connecticut, 75 North Eagleville Road, Storrs Connecticut, 06269, USA*
**Corresponding author*
E-mail: diba.khan-bureau@uconn.edu

Received: 17 February 2014 / Accepted: 24 July 2014 / Published online: 1 September 2014

Handling editor: Vadim Panov

Abstract

The nuisance taxon *Didymosphenia geminata* was reported in the West Branch of the Farmington River in March, 2011 after a fisherman detected cotton-like tufts attached to rocks. In response, the Connecticut Department of Energy and Environmental Protection (CT DEEP) conducted a comprehensive survey of the river system. After major late summer storms, *Didymosphenia geminata* was not observed again. Surveys in 2012–2013 tracked the spatial and temporal distribution of stalk-forming diatoms at the confluence of the West Branch of the Farmington and Still Rivers, thereby allowing comparison of data from adjacent rivers with distinct water chemistries. Water chemistry and temperature data were collected to characterize nutrient concentrations associated with these diatoms. Surveys showed no evidence of *Didymosphenia geminata* but four native stalk-forming diatom species and a taxon previously unreported in Connecticut, *Cymbella janischii*, were observed throughout the year. Also from November, 2012 through June, 2013, a morphologically distinct diatom in the genus *Didymosphenia* was observed growing prolifically bank to bank with thick mats of long filamentous stalks. Subsequent examination revealed that the taxon previously reported as *Didymosphenia geminata* was instead a different taxon, *Didymosphenia* sp. Furthermore, *Didymosphenia* sp. continued to flourish in the West Branch of the Farmington River, absent from the neighbouring Still River, suggesting that the physiochemical features and in particular higher nutrients may limit the distribution of this diatom. In contrast, *C. janischii* was found growing abundantly further downstream in warmer water and higher nutrient levels.

Key words: *Cymbella janischii*, *Didymosphenia*, extracellular polymeric stalks, invasive diatoms, mucilaginous tufts, nuisance, rock snot

Introduction

Didymosphenia geminata (Lyngbye) M. Schmidt has become a taxon of interest to ecologists, biologists, anglers, and water quality managers throughout the world since its invasion of New Zealand in 2004 (Spaulding and Elwell 2007; Blanco and Ector 2009). *Didymosphenia geminata*, often referred to as "rock snot", is a putatively invasive, and nuisance species (Kilroy 2004; Spaulding and Elwell 2007; Kuhajek et al. 2014). This species produces extracellular polymeric stalks (EPS) that persist even after the cells are no longer viable, forming mats that have the potential to negatively impact aquatic organisms within rivers and streams (Spaulding and Elwell 2007; CT DEEP 2011).

Didymosphenia geminata natively occurs in cold, oligotrophic waters, in mountainous regions and temperate climates with cold winters and warmer summers, although there are conflicting reports of its habitat preference. Its distribution now spans diverse conditions from unpolluted to polluted waters (Krammer and Lange-Bertalot 1988;

Kilroy 2004; Kilroy et al. 2007; Spaulding and Elwell 2007). The geographical range of *D. geminata* has expanded since it was originally described nearly 200 years ago (Lyngbye 1819; Blanco and Ector 2009; Whitton et al. 2009). This expansion may be because its growth has become more abundant in recent years making it more readily observed where it once may not have been detected (Spaulding and Elwell 2007; Blanco and Ector 2009; Kumar et al. 2009; Bothwell et al. 2014). The current rapid growth and geographical expansion of *D. geminata* may in part be due to seasonal changes, climate change, variation of nutrients such as orthophosphate (SRP), light intensity, rainfall patterns and other environmental factors (Ellwood and Whitton 2007; Kilroy et al. 2007; Spaulding and Elwell 2007; Bothwell and Spaulding 2008; Kilroy et al. 2008; Bothwell and Kilroy 2011; Kuhajek et al. 2014).

In the United States, *D. geminata* was potentially transported by anglers boots, fishing gear and other recreational equipment (Kirkwood et al. 2007; Bothwell et al. 2009) from Western into several Southeastern states, including Virginia, West Virginia, Tennessee, North Carolina (Spaulding and Elwell 2007), and more recently the Northeastern states. Significant growth of *D. geminata* was documented in North America only within the last 20 years (Bothwell and Spaulding 2008; Spaulding et al. 2008; Blanco and Ector 2009). Recently, in the Northeastern U.S.A., *D. geminata* was found in the main stem of the Connecticut River and several of its tributaries in Vermont, New Hampshire and purportedly Connecticut (CT DEEP 2011). In May, 2013 the Massachusetts Executive Office of Energy and Environmental Affairs (MASS EOEEA 2013) announced that *D. geminata* was confirmed in the Green River. Some researchers suggest that *D. geminata* may be indigenous, but rare in the Northeastern U.S.A. (Bothwell and Spaulding 2008; Spaulding et al. 2008; Blanco and Ector 2009) but many state agencies consider *D. geminata* as non-indigenous in the Northeastern U.S.A.

The first observation of *Didymosphenia* in Connecticut was reported to the CT DEEP in 2011 with no earlier record to substantiate this taxon as indigenous to Connecticut (CT DEEP 2011). Uncertainties persist whether *D. geminata* is introduced to or native in Connecticut. Terry (1907) provided a partial list of diatoms found in Connecticut, with *Gomphonema geminatum* (Lyngbye) Ehrenberg (Terry mistakenly used Ehrenberg rather than Agardh as authority) listed as a common species, although no illustrations were provided.

The objectives of this study were to document the presence of two previously unrecorded stalk-forming diatoms in the West Branch of the Farmington River in Connecticut, *Cymbella janischii* (A. Schmidt) De Toni, and an unidentified taxon in the genus *Didymosphenia*. We also contribute information about the environmental conditions that are associated with these and other dominant and native stalk-forming diatom species, including *Cymbella affinis* Kützing, *Gomphonema truncatum* Ehrenberg, *Gomphoneis minuta* (J.L. Stone) Kociolek and Stoermer, and *Encyonema* cf. *minutum* (Hilse) D.G.Mann in the West Branch of the Farmington River over the years 2012–2013. The study site spans the confluence of two rivers with distinct management regimes and characteristics, allowing a comparison of the environmental conditions that are associated with abundant stalk growth of these diatom taxa.

Materials and methods

Study location

The confluence of the West Branch of the Farmington River and the Still River in north central Connecticut provides a unique setting to examine the relationship of water chemistry, temperature, and the presence of stalk-forming diatoms. The upstream catchments for both rivers have markedly different land-use, water quality, and flow regulation.

The West Branch Farmington River (616 km^2) is one of several sub-basins within the greater Farmington River regional basin (1,572 km^2), a significant tributary to the lower Connecticut River (29,184 km^2). The land-use is greater than 85% forested, with no significant population density and has less than 3% impervious cover. The West Branch Farmington River begins in Otis, Massachusetts and is impounded twice after entering Connecticut, first at the Colebrook River Reservoir and then the West Branch Reservoir. It has been suggested that *D. geminata* is more likely to occur in rivers that are regulated by dams because of stable stream flows and constant cooler temperatures (Spaulding and Elwell 2007). The MDC operates West Branch Reservoir and has an agreement to maintain a minimum discharge from the nutrient-poor hypolimnion. This very cold water is essential to support a highly managed destination trout fishery for North American and international anglers, with over 116,000 angling

BioInvasions Records (2014) Volume 3, Issue 3: 139–149
doi: http://dx.doi.org/10.3391/bir.2014.3.3.02

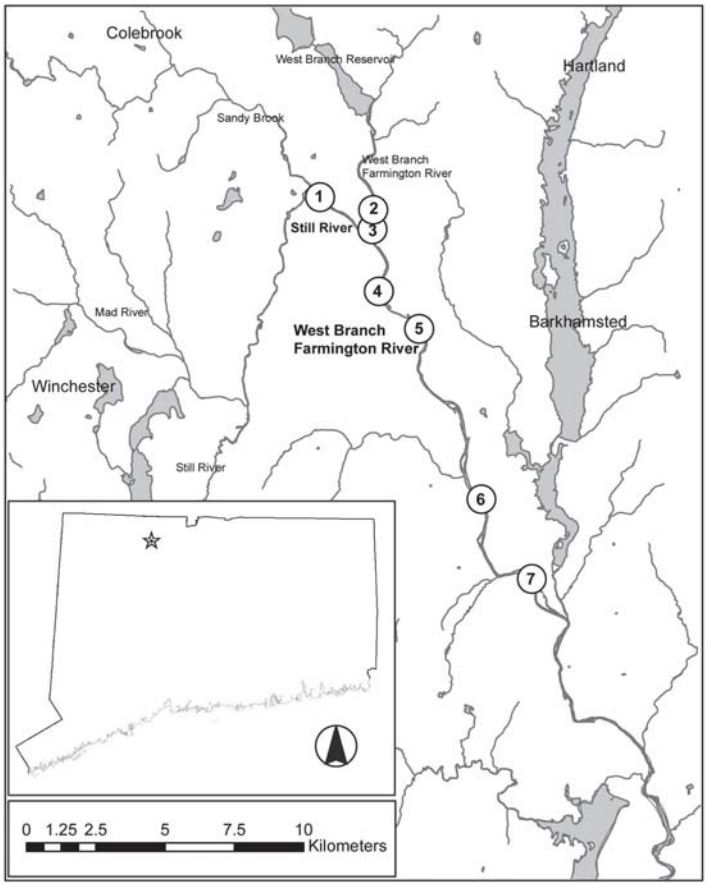


Figure 1. Map of the sampling locations at the Still River and West Branch of the Farmington River in Connecticut (corresponding to sites listed in Tables 1 and 2). Site 1 represents the Still River. Sites 2–7 represent the West Branch of the Farmington River. Site 2 is the location of the mixed stalk-forming diatoms. Site 3 represents the location of the abundant growth of *Didymosphenia* sp. Site 6 represents the location of the abundant growth of *Cymbella janischii*. Site 7 represents the location where *Gomphoneis minuta* was growing.

hours estimated annually (C. Bellucci, CT DEEP, pers. comm.).

Approximately 3 km downstream of the Goodwin dam is the confluence with the Still River (270 km^2). Thirty percent is comprised of the Sandy Brook sub-regional basin, which is over 86% forested and has minimal human disturbance with less than 3.7% impervious surface, minimizing nonpoint source pollution runoff considered a contributing factor for the growth of algae. The remaining 70% comprises the Still River sub-regional basin. At the lower end of this catchment is the city of Winsted, with highly urbanized land use and a waste water treatment facility with a permitted maximum allowable final effluent discharge of 13,250 m^3/d with a 10 year average of 5,678 m^3/d (C. Bellucci CT DEEP, pers. comm., MDC 2013). In addition the permitted maximum daily discharge for total N is 50 mg/l/d.

Water Sampling

Historically, the CT DEEP has monitored the water chemistry of these rivers. At the sites illustrated in Figure 1 and listed in Tables 1 and 2, water samples were taken by grab sampling at depths from 38 – 76.2 cm on 22 May, 2013 and analyzed by the United States Geological Survey (USGS) in East Hartford, Connecticut (Table 2). Also, water samples from the Still and West Branch of the Farmington Rivers were taken and transported to the USGS in East Hartford for analysis of total N and SRP on 12 June, 2013 (Table 2). For the remainder of the sampling dates the authors of this study monitored river water *in situ* for temperature, pH, and conductivity using an YSI® 30 portable hand-held metering probe. Samples were placed in 125 ml sterile wide mouthed poly containers, placed in a cooler and

BioInvasions Records (2014) Volume 3, Issue 3: 139–149
doi: http://dx.doi.org/10.3391/bir.2014.3.3.02

Table 1. Location and characteristics of sites 1–7 of the Still River and the West Branch of the Farmington River in Connecticut, USA.

Site number	Locality	River and site characteristics
1	Still River, Colebrook 41.967° N, 73.033° W	Larger rocks, open river deeper channel, shaded east and west banks, no impoundments, non-regulated free flowing with an average flow rate of 6.145 m^3/s, mid-sized wastewater treatment plant located 4.7 km upstream from the confluence West Branch of the Farmington and Still River
2	West Branch Farmington River Riverton USGS, Barkhamsted 41.962° N, 73.0176° W	Larger rocks, deeper channel, shaded east and west, regulated flow, impoundments, above the confluence of the West Branch of the Farmington and Still River
3	West Branch Farmington River Riverton Cemetery, Barkhamsted 41.960° N, 73.017° W	Open river channel, shaded - western bank, full morning and afternoon sun, regulated flow, impoundments, above the confluence of the West Branch of the Farmington and Still River, cobbles and boulder substrate, riffles and an average flow rate of 8.9 m^3/s
4	West Branch Farmington River 1 km below confluence, Barkhamsted 41.957° N, 73.015° W	Open river channel, shaded - western banks, morning and afternoon sun, below the confluence of the West Branch of the Farmington and Still River
5	West Branch Farmington River Whittemore Recreation Area, Barkhamsted 41.945° N, 73.016° W	Small vegetated islands with rushes, grasses sedges, shaded - western bank, morning and afternoon sun, shallower channel, cobbles and boulder substrate, riffles
6	West Branch Farmington River Pleasant Valley, Barkhamsted 41.897° N, 72.984° W	Small vegetated islands with rushes, grasses, sedges with full sun, shallower wider channel, cobbles and boulder substrate, riffles
7	West Branch Farmington River Black Bridge, New Hartford 41.877° N, 72.965° W	Open river channel, shaded - western bank, morning and afternoon sun, shallower wider channel, cobbles and boulder substrate, riffles

transported to the Center for Environmental Sciences and Engineering Analytical Services (CESE) at the University of Connecticut in Storrs, a Department of Health certified lab. CESE tested river samples for SRP, total N, and pH (Table 2). Statistical analyses were performed using IBM SPSS Statistics version 21. Log transformation was employed after testing the assumption of normality. Analysis of variance (ANOVA) was used to determine if significant differences existed among samples which contained *Didymosphenia* sp. and those that did not for each of the water quality parameters (SRP, total N, and temperature). Box plots were prepared to illustrate the levels of SRP, total N, and the water temperature across the seven sites, with the presence of different diatoms indicated. We also plotted the frequency of *Didymosphenia* sp. against the levels of SRP, for the 19 distinct water samples (as shown in Bothwell et al. 2014, for *D. geminata*).

Diatom sampling

Benthic samples were collected from several locations in the West Branch of the Farmington and Still Rivers (see Table 1 and Figure 1). Sampling took place in the locations where the putative *D. geminata* had been reported from collections made in 2011. Grab samples of mucilaginous tufts were pulled from rocks and taken from vegetation and placed in Whirl-Pac® bags. The latter were placed on ice and transported to the lab for processing. Sampling took place weekly or two times a month during snow and ice cover. All samples were stored at 4ºC until further processing.

Diatom preparation

Diatom samples were simmered on a hot plate in a 1:1 ratio of water and 68% nitric acid to oxidize organic matter, after which the samples were removed from the hotplate to cool. Deionized water was used to rinse the samples of the acid, and then the samples were centrifuged to concentrate the diatom frustules at 600 g to avoid frustule damage. The process of rinsing included the addition of deionized water, centrifuging and the removal of supernatant 4–5 times or until the pH was neutral.

Light microscopy and scanning electron microscopy

Prior to acid washing, samples were placed on a microscope slide with a coverslip overlain and then viewed at ×200 and ×400 magnifications using a BX 60 Olympus microscope. The diatom sample slurry was air dried onto microscope coverslips, then used to make permanent slides with the mounting medium NAPHRAX®. The diatom frustules were examined at ×1000

magnifications with a BX 60 Olympus microscope. Images were captured using an Olympus DP 25 color digital camera (2560 × 1920 pixels) with Olympus cellSens software. The diatoms on these slides were identified based on their morphological characteristics according to Krammer and Lange-Bertalot (1988), Round et al. (1990), and three online databases, the ANSP Algae Image Database (http://diatom.ansp.org/algae_image/), Diatoms of the United States (http://westerndiatoms.colorado.edu/), and the Great Lakes Image Database: (http://www.umich.edu/~phytolab/GreatLakesDiatomHomePage/top.html).

For SEM, single diatom cells were isolated using a microscope at x100 magnification with a micropipette and transferred onto 25 mm, 3 µm pore polycarbonate Millipore filters (Lang and Kaczmarska 2011). The filters were adhered to SEM stubs with double-sided tape. *Cymbella janischii* and other diatom samples were prepared following the methodology of Morales et al. (2001) the stubs were coated for 1 min at 1.8 kV with gold/palladium using a Polaron sputter coater. The stubs were viewed with the field emission Leo/Zeiss DSM 982 and a field emission FEI Nova Nano 450 scanning electron microscope located at the University of Connecticut Electron Microscopy lab. Image plates were created using Adobe® Creative Suite® 6 Photoshop.

Results

The Still River and the West Branch of the Farmington River are distinct river systems (Figure 1). The water quality data and physical attributes presented in Tables 1 and 2 illustrate that the two rivers have diverse water chemistry, temperature, flow regimes, geomorphology, and sunlight availability due to canopy coverage. These rivers also differ in the benthic diatom taxa present (Table 2). For the Still River (site 1, Tables 1, 2), the mean (antilog ± standard deviation) total N concentration was 413.0 ± 242.5 µg/l, the mean SRP concentration was 13.3 ± 7.4 µg/l, and the mean water temperature was 9.12 ± 4.36°C. *Didymosphenia* sp. and *C. janischii* were not observed at this site over the entire sampling period. At site 2 (Tables 1, 2) the mean total N concentration was 258.0 ± 57.5 µg/l, the mean SRP concentration was 4.12 ± 2.9 µg/l, and the mean water temperature was 8.37 ± 0.65°C, with minimal sunlight exposure from bridge and canopy shading. There was a mixture of the stalk-forming species, *Cymbella affinis, Gomphonema truncatum,* and *Gomphoneis minuta* at this site. Stalk material from a mixture of the three common native diatom taxa covered the substrate from November, 2012, through December, 2012 without *Didymosphenia* sp. Visible growth occurred again in May and June, 2013 with *C. janischii* (although limited and patchy) and *Didymosphenia* sp. present (Figure 2), with the mixture of species restricted to an area of approximately 15 m bank to bank. *Didymosphenia geminata*, described and illustrated by Spaulding and Elwell (2007) and Spaulding (2010), was not detected at this site or during this study. Notably, thick mats with 90% coverage of the morphologically different *Didymosphenia* sp. dominated one segment of the West Branch of the Farmington River above the river confluence (site 3). At site 3 the mean total N concentration was 225.0±31.7 µg/l; the mean SRP concentration was 2.88 ± 1.0 µg/l; and the mean water temperature was 6.09 ± 2.9°C. Site 3 had a wider channel,

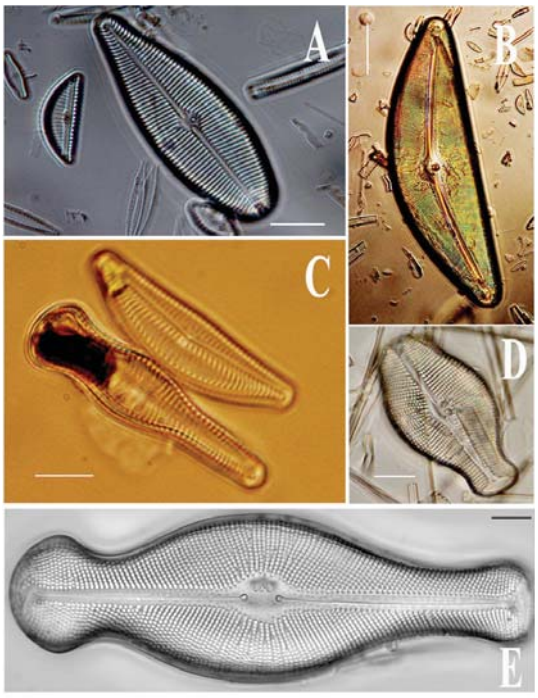


Figure 2. LM images **A**. *Gomphoneis minuta*. **B.** *Cymbella janischii*. **C.** *Cymbella affinis* and *Gomphonema truncatum*. **D.** *Didymosphenia* sp. E. *Didymosphenia geminata* for comparison with **D** (**E** image courtesy of Sarah Spaulding). Scale bars represent 10 µm in **A, C, D**, and **E.** Scale bar represents 20 µm in **B.** Photomicrographs by D. Khan-Bureau.

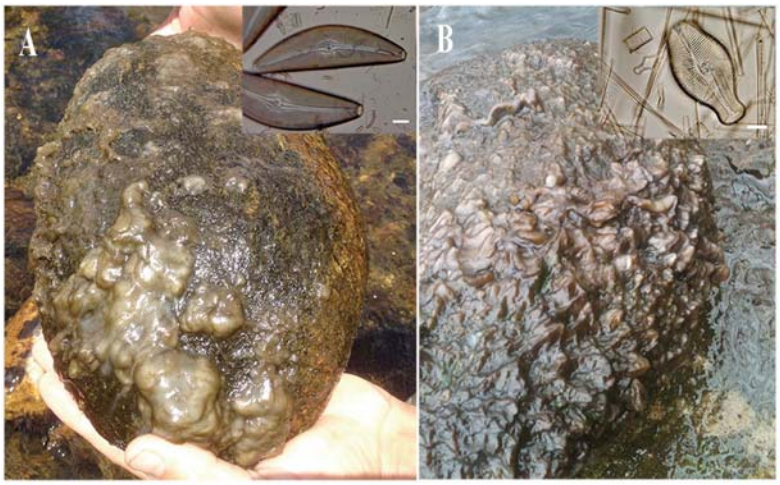


Figure 3. Rocks covered with mucilaginous stalk growth from the West Branch of the Farmington River in Connecticut. **A.** *Cymbella janischii* stalk growth. Inset: cleaned *C. janischii* cells. **B.** *Didymosphenia* sp. stalk growth. Inset: cleaned *Didymosphenia* sp. cell. Scale bar represents 20 µm in image A and 10 µm in image B. Photographs by D. Khan-Bureau.

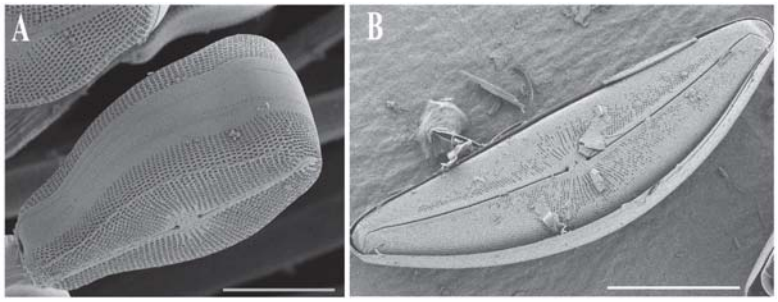


Figure 4. SEM images of **A.** *Didymosphenia* sp. and **B.** *C. janischii*. Scale bar represents 20 µm in A and 50 µm in B. SEM images by D. Khan-Bureau.

is shallower (76 cm) and had abundant sun (Tables 1, 2) unlike sites 1 and 2. The bloom of *Didymosphenia* sp. covered 1 km, 50–60 m bank to bank, with stalked material forming 2.0–5.0 cm thick on the rocky substrate. LM confirmed a combination of other diatoms, benthic macroinvertebrates and river debris within *Didymosphenia* sp. stalked mats. As the river flows further downstream, and particularly after the confluence of the Still and the West Branch of the Farmington Rivers, the nutrient levels and temperatures increase (Table 2). Further downstream, past the confluence (sites 4, 5), *Didymosphenia* sp. was observed in late May, 2013, although the growth was limited to just 1–3 tufts observed. LM and SEM observations confirmed *Cymbella janischii* at site 6 in July, 2012 and stalk growth absent by early October, 2012, as green algae colonized the site. The *C. janischii* mats were thick and covered the substrate bank to bank approximately 40 m wide and 0.5 km each side of the islands (Figure 3). *Encyonema* cf. *minutum* tufts were seen at site 6 further downstream of *C. janischii*, although patchy. Mixing continued below the confluence for approximately 11 km as demonstrated by physicochemical properties (Table 2, site 7). LM confirmed *Gomphoneis minuta* tufts growing at site 7 in early November, 2012, but growth was not observed in December, 2012.

SEM was used to examine the walls of *Didymosphenia* sp. and *C. janischii* for identification purposes because these species have similar mucilaginous growth and both are new records in Connecticut. *Cymbella janischii* cells are asymmetrical whereas *Didymosphenia* sp. cells are not (Figure 4). *Cymbella janischii* cells are large, normally 130–360 µm (Kociolek and Stoermer 1988; Round et al. 1990; Metzeltin and Lange-Bertalot 1995; Bahls 2007; Kociolek 2011). The *C. janischii* cells collected for this study ranged in size from 130–150 µm. The cells of *Didymosphenia* sp. ranged consistently in size from 50–60 µm with few at 38 µm and 68 µm and formed macroscopic mucilaginous strands as long as 18 cm. The size range for *D. geminata* in the U.S.A.

is 87–137 µm (Spaulding 2010, Figure 2E). The *Didymosphenia* sp. cells were also unusual because of their compressed morphology in contrast to *D. geminata* (Figure 2D, E).

The water chemistry data were found to be log-normally distributed thus statistical analyses were performed on log-transformed values. ANOVA showed no significant difference among samples that had *Didymosphenia* sp. and those that did not, for total N, SRP, and temperature. A lack of significant difference could indicate no difference or be due to the small sample size, therefore additional water quality sampling would be required to test the relationship between water chemistry and absence/presence of *Didymosphenia* sp. Box plots of total N, SRP, and the T across the seven sites illustrate the distribution of different diatom taxa across the sites and suggest that chemical characteristics (Figure 5) may be linked to environmental preferences but further regular analysis of site water parameters is needed to determine potential correlations. Lastly, we show the occurrence of *Didymosphenia* sp. at low SRP levels (Figure 6) similar to Bothwell et al. (2014) for *Didymosphenia geminata.*

Discussion

Confusion about Didymosphenia geminata *in the Northeastern U.S.A.*

In March, 2011 tufts collected by the CT DEEP from the West Branch of the Farmington River in Connecticut were sent to the Vermont Department of Conservation (VT. DEC) for identification. Using LM, the sample was identified as *D. geminata*, although the cells were on the low end of the size range for this species, having been roughly estimated at 80–90 µm long. Identification cannot be verified since the lab in which the voucher specimen was stored was destroyed during Tropical Storm Irene. No other samples exist (L. Matthews, VT DEC, pers. comm.). The Connecticut DEEP speculates that the purported *Didymosphenia geminata* was not found again in 2011 and in early 2012 because the combination of Tropical Storms Irene and Lee created significant mechanical scouring that may have contributed to the reduction of the population of this diatom (M. Becker CT DEEP, pers. comm.). During these storms the associated rainfall created historic flows and channel alterations even though the river flow is highly regulated. It is hypothesized that stable river flows

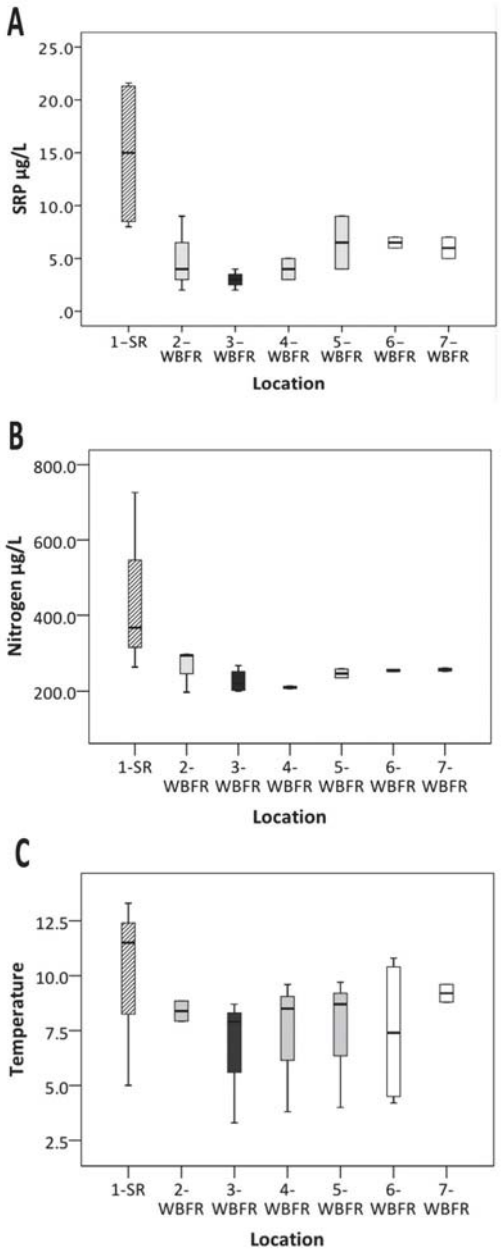


Figure 5. **A**. The distribution relationship of *Didymosphenia* sp. and SRP levels across the 7 sites surveyed. **B.** The distribution relationship of *Didymosphenia* sp. and total nitrogen levels across the 7 sites surveyed. **C.** The distribution relationship of *Didymosphenia* sp. and temperature across the 7 sites surveyed. The hatched lines represent absence of *Didymosphenia* sp. at the Still River site 1. At site 2 light grey boxes represent *Didymosphenia* sp. observed but not abundant and occurred with a mix of stalk-forming diatoms. At site 3 dark grey filled box represent *Didymosphenia* sp. was abundant and blooming 1 km by 50 m. At site 4 and 5 light grey boxes represent only 1–3 tufts observed. At sites 6, 7 unfilled boxes represent native species present, absent of *Didymosphenia* sp. and *C. janischii*.

and secure substrates allow for the establishment of *D. geminata* colonies (Spaulding and Elwell 2007). From the observations made in this study, a stable regulated flow appears to be suitable for this morphologically distinct genus of *Didymosphenia* as well. Many reports of *D. geminata* blooms are from deep, cold, lake-fed, flow restricted and regulated streams (Kilroy 2004; Kumar et al. 2009) and the population of *Didymosphenia* sp. found in this study may share similar environmental tolerances and physio-chemical traits (Figures 5, 6). Further work is underway to understand the morphological and genetic variation of *Didymosphenia* sp. to verify whether this is a new taxon.

Another nuisance stalked diatom, Cymbella janischii

During the summer survey of July 2012, prolific mucilaginous clumps were found (site 6). LM revealed the identification as *Cymbella janischii*. This species, endemic in the Pacific Northwest, was recently reported in Japan and is reported here as present in Connecticut (Bahls 2007; Suzawa et al. 2011). *Cymbella janischii* may have been transported from the Pacific Northwest to the Northeast U.S.A. to as far away as Japan by angler's boots, angler equipment, and by other means (Suzawa et al. 2011). *Cymbella janischii* has similar mucilaginous stalk growth as *D. geminata* (and likewise as *Didymosphenia* sp.) (Pite et al. 2009, Whitton et al. 2009; Suzawa et al. 2011). These taxa are thought to be potentially nuisance, aggressively forming thick gray mats on substrates, even in their native habitats (Spaulding and Elwell 2007). *Didymosphenia geminata* and *C. janischii* have become more problematic in recent years due to expansions of their geographical ranges (Bahls 2007; Kumar et al. 2009; Pite et al. 2009).

Study site habitat preference

The absence of *Didymosphenia* sp. from the Still River, site 1, suggests that environmental factors such as higher levels of SRP and total N with increased water temperatures, and reduced light availability, may limit the growth of this taxon. Discharge from the city of Winsted's waste water treatment plant contributes nutrients to the free-flowing Still River according to the CT DEEP approved permit, CT0101222, for the Town of Winchester 2005 NPDES 2005. Whereas the upper extent (sites 2, 3) of the West Branch of the Farmington is flow-regulated by the MDC Goodwin dam, which discharges very cold oligotrophic waters. Both rivers are heavily visited by anglers throughout the year, and given the close proximity of the two rivers, anglers typically fish in both rivers in one day. Blooms of *Didymosphenia* sp. were recorded only at the upper extent of the West Branch of the Farmington (site 3). Despite the close proximity of the rivers, the spread of *Didymosphenia* sp. by anglers and recreationalists has not occurred in the Still River. Bothwell et al. (2012) reported that the growth of *D. geminata* ceases in river reaches downstream of point source nutrient outfalls. It is possible that *Didymosphenia* sp. does not aggressively grow in higher SRP, total N and warmer waters, as was proposed for *D. geminata* (Bothwell and Spaulding 2011; Kilroy and Bothwell 2011; Bothwell et al. 2012). This hypothesis needs to be tested further.

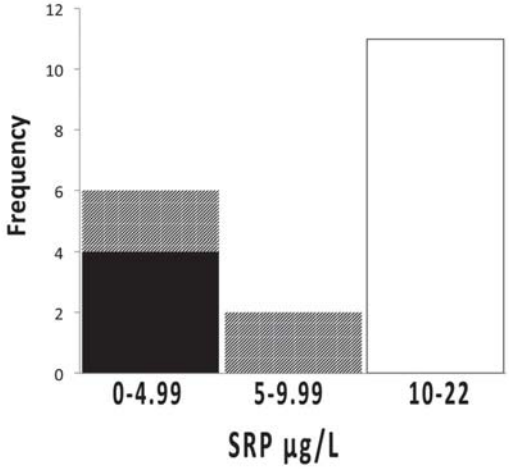


Figure 6. The frequency of occurrence of *Didymosphenia* sp. in 19 river samples as a function of SRP levels. Filled boxes=abundant and blooming, hatched boxes=present but not blooming, unfilled box= not observed.

Analysis of limited grab samples for SRP and total N, and water temperature indicate that the West Branch of the Farmington River may possibly have narrower water chemistry and temperature ranges. Our observations of *Didymosphenia* sp. echoes the recent work by Bothwell et al. (2014), on *D. geminata,* shown to grow prolifically because of SRP limitation. Our location affords a unique opportunity to quantify various physical and chemical variables with blooms of *Didymosphenia* sp. in a natural environmental setting. It may be that the SRP rich waters of the Still River and the mixing at the confluence of the Still River and the West Branch Farmington River

Table 2. Water chemistry and diatom survey data from seven sites the West Branch of the Farmington and Still Rivers in Connecticut, U.S.A. Major stalk-forming diatom taxa were surveyed on all dates shown, and their occurrences are indicated as follows: — = taxon not observed; + = taxon present but at low abundance or of limited distribution; ++ = taxon present and abundant.

Site	Sample Date	SRP µg/L	N µg/L	pH	T °C	Cond. µS/cm	Stalk- forming diatom taxon					
							Didymo-sphenia sp.	*Cymbella janischii*	*Cymbella affinis*	*Gompho-nema truncatum*	*Gomphoneis minuta*	*Encyonema* cf. *minutum*
1	22 Mar. 2013	8	264	6.8	5	n.d.	—	—	—	—	—	—
	14 May 2013	21	367	6.7	11.5	n.d.	—	—	—	—	—	—
	22 May 2013*	21.6	726	n.d.	n.d.	n.d.	—	—	—	—	—	—
	28 May 2013	9	n.d.	6.9	13.3	111	—	—	—	—	—	—
2	29 Nov. 2012	n.d.	n.d.	n.d.	n.d.	n.d.	—	—	++	++	++	—
	22 Mar. 2013	n.d.	n.d.	n.d.	3.3	n.d.	—	—	—	—	—	—
	14 May 2013	2	197	6.8	7.93	n.d.	—	—	+	+	+	—
	22 May 2013*	<4	295	n.d.	n.d.	n.d.	—	—	+	+	+	—
	28 May 2013	9	n.d.	6.8	8.85	86.5	+	+	++	++	++	—
	12 June 2013*	<4	298	n.d.	n.d.	n.d.	+	+	++	++	++	—
3	29 Nov. 2012	n.d.	n.d.	n.d.	n.d.	n.d.	++	—	—	—	—	—
	22 Mar. 2013	3	236	6.6	3.3	n.d.	++	—	—	—	—	—
	14 May 2013	2	204	6.9	7.9	n.d.	++	—	—	—	—	—
	28 May 2013	n.d.	200	6.8	8.7	87.1	++	—	—	—	—	—
	12 June 2013*	<4	268	n.d.	n.d.	n.d.	++	—	—	—	—	—
4	22 Mar. 2013	n.d.	n.d.	n.d.	3.8	n.d.	—	—	—	—	—	—
	14 May 2013	3	213	6.8	8.5	n.d.	—	—	—	—	—	—
	28 May 2013	5	207	6.8	9.6	88.6	+	—	—	—	—	—
5	10 July 2012	n.d.	n.d.	n.d.	n.d.	n.d.	—	—	—	—	—	—
	4 Dec. 2012	n.d.	n.d.	n.d.	n.d.	n.d.	—	—	—	—	—	—
	22 Mar. 2013	n.d.	n.d.	n.d.	4	n.d.	—	—	—	—	—	—
	14 May 2013	9	259	6.8	8.7	n.d.	+	—	—	—	—	—
	28 May 2013	4	235	6.9	9.7	96.4	+	—	—	—	—	—
6	10 July 2012	n.d.	n.d.	n.d.	n.d.	n.d.	—	++	—	—	—	—
	4 Dec. 2012	n.d.	n.d.	n.d.	n.d.	n.d.	—	—	—	—	—	+
	22 Mar. 2013	n.d.	n.d.	n.d.	4.8	n.d.	—	—	—	—	—	—
	14 May 2013	7	258	n.d.	10	n.d.	—	—	—	—	—	—
	28 May 2013	6	252	n.d.	10.8	100.7	—	—	—	—	—	—
7	10 July 2012	n.d.	n.d.	n.d.	n.d.	n.d.	—	—	—	+	+	—
	29 Nov. 2012	n.d.	n.d.	n.d.	n.d.	n.d.	—	—	—	—	+	—
	4 Dec. 2012	n.d.	n.d.	n.d.	n.d.	n.d.	—	—	—	—	+	—
	22 Mar. 2013	n.d.	n.d.	n.d.	4.2	n.d.	—	—	—	—	—	—
	14 May 2013	7	261	6.8	8.8	n.d.	—	—	—	—	—	—
	28 May 2013	5	253	7	9.6	88.6	—	—	—	—	—	—

† *Data analyzed by the USGS for this project; SRP listed as <4 µg/L when not detectable; µg/L=1ppb, n.d = no data
Connecticut DEEP water quality data is available upon request

have limited the downstream expansion of *Didymosphenia* sp. Our preliminary results suggest that *Didymosphenia* sp. blooms are related to SRP limitation but other environmental factors may contribute such as the depth and width of the channel, sunlight availability, flow regulation, and other physical and chemical parameters. Collection of additional chemical, stream flow, and algal community structure, cell density and biomass data will help to test our hypothesis.

Whereas many state regulatory agencies consider *D. geminata* to be invasive in the Northeastern U.S.A., the literature supports a native distribution that is circumpolar in the Northern hemisphere (Krammer and Lange-Bertalot 1988; Blanco and Ector 2009; Kumar et al. 2009). The debate continues as to whether *Didymosphenia geminata* should be classified as invasive to the Northeastern states. We documented two previously unreported taxa, *Didymosphenia* sp. and *C. janischii*. Were these taxa transported via anglers boots and equipment and by other vectors (Kirkwood et al. 2007; Bothwell et al. 2009) or are *D. geminata* and *Didymosphenia* sp. native but rare, now becoming nuisance due to changing environmental conditions throughout their ranges (Valéry et. al 2009; Bothwell 2014)? During our study documentation of *C. janischii*, a diatom not previously found east of the Rocky Mountains (Bahls 2007; Kumar et al. 2009; Suzawa et al. 2011) was

confirmed and its presence was adjacent to a well-travelled footpath to the river, suggestive of an anthropogenic source.

Given that diatoms are important biological indicators and are often used for water quality assessments, it is crucial to identify diatoms accurately (Morales et al. 2001; Mann et al. 2010; Pniewski et al. 2010). Further work is needed to identify this unfamiliar *Didymosphenia* sp. including collection of information on the morphological variation present as well as comparison of this taxon to other species of *Didymosphenia* using genetic data. The presence and excessive stalk growth of *Didymosphenia* sp. in the West Branch of the Farmington River suggests a need for further monitoring of these species to determine which environmental conditions are associated with nuisance growth.

Acknowledgements

We thank K. Fučíková, S. Olm, C. Lo, M. Letsch, H. McManus, M. Dietz and J. Green, and the late F. Trainor for their helpful discussions. Three Rivers Community College's, Dean of IT S. Goetchius, K. Barfield, S. Cohen, D. Jewett and President G.S. Jones for their support of this project. J. Sgro provided verification of species as well as guidance. We thank M. Cantino, B. Goodwin, and S. Daniels of the EM lab at UCONN NSF grant no.1126100 supporting the purchase of the Nova Nano SEM. Thanks to J. Morrison, East Hartford, USGS, C. Bellucci and M. Becker, CT DEEP, L. Matthews, VT DEC. The 2013 International Didymo Conference – Invasive Species Action Network and Sarah Spaulding for allowing us to use her image in this paper. We thank the Reviewers and Editor for their helpful comments. Thanks to The Connecticut Institute of Water Resources USGS for funding this research (grant no. 5606350).

References

Bahls LL (2007) *Cymbella janischii*—Giant endemic diatom of the Pacific Northwest: morphology, ecology and distribution compared to *Cymbella mexicana*. *Northwest Science* 81: 284–292, http://dx.doi.org/10.3955/0029-344X-81.4.284

Blanco S, Ector L (2009) Distribution, ecology and nuisance effects of the freshwater invasive diatom *Didymosphenia geminata* (Lyngbye) M. Schmidt: a literature review. *Nova Hegwigia* 88: 347–422, http://dx.doi.org/10.1127/0029-5035/2009/0088-0347

Bothwell ML, Spaulding SA (2008) Synopsis: the 2007 International Workshop on *Didymosphenia geminata*. In: Bothwell ML, Spaulding SA (eds), Proceedings of the 2007 International Workshop on *Didymosphenia geminata*. *Canadian Technical Report of Fisheries and Aquatic Sciences* 2795

Bothwell ML, Lynch DR, Wright H, Deniseger J (2009) On the boots of fishermen: The history of *Didymo* blooms on Vancouver Island, British Columbia. *Fisheries* 34: 382–388, http://dx.doi.org/10.1577/1548-8446-34.8.382

Bothwell ML, Kilroy C (2011) Phosphorus limitation of the freshwater benthic diatom *Didymosphenia geminata* determined by the frequency of dividing cells. *Freshwater Biology* 56: 565–578, http://dx.doi.org/10.1111/j.1365-2427.2010.02524.x

Bothwell ML, Kilroy C, Taylor B, Ellison E, James D, Gillas CA, Bladon KD, Silins U (2012) Iron is not responsible for *Didymosphenia geminata* bloom formation in phosphorus-poor rivers. *Canadian Journal Fisheries and Aquatic Sciences* 69: 1723–1727, http://dx.doi.org/10.1139/f2012-112

Bothwell ML, Taylor B, Kilroy C (2014) The Didymo story: the role of low dissolved phosphorus in the formation of *Didymosphenia geminata* blooms. *Diatom Research* 29(3): 229–236, http://dx.doi.org/10.1080/0269249X.2014.889041

Connecticut Department of Energy and Environmental Protection (CT DEEP) (2011) DEP Reports Didymo Discovered in the West Branch Farmington River [online]. Available from: http://www.ct.gov/deep/cwp/view.asp?Q=476204&A=4013 (Accessed 12 September 2011)

Connecticut Department of Energy and Environmental Protection (CT DEEP), Bureau of Water Protection and Land Reuse, Municipal Facilities Division, Discharge Monitoring Reports. (2005) Town of Winchester NPDES permit number CT0101222

Ellwood NTW, Whitton BA (2007) Importance of organic phosphate hydrolyzed in stalks of the lotic diatom *Didymosphenia geminata* and the possible impact of atmospheric and climatic changes. *Hydrobiologia* 592: 121–133, http://dx.doi.org/10.1007/s10750-007-0728-0

Kilroy C (2004) A new alien diatom, *Didymosphenia geminata* (Lyngbye) Schmidt: its biology, distribution, effects and potential risks for New Zealand fresh waters [online]. Available from: http://www.biosecurity.govt.nz/files/pests/didymo/didymo-preliminary-org-ia-nov-04.pdf (Accessed 10 June 2012)

Kilroy C, Lagerstedt A, Davey A, Robinson K (2007) Studies on the survivability of the invasive diatom *Didymosphenia geminata* under a range of environmental and chemical conditions. National Institute of Water and Atmospheric Research, New Zealand, Client Report CHC2006-116, NIWA Project MAF06506. Christchurch, New Zealand

Kilroy C, Snelder T, Floerl O, Vieglais CC, Dey KL (2008) A rapid technique for assessing the suitability of areas for invasive species applied to New Zealand's rivers. *Diversity and Distributions* 14: 262–272, http://dx.doi.org/10.1111/j.1472-4642.2007.00406.x

Kilroy C, Bothwell ML (2011) Environmental control of stalk length in the bloom-forming, freshwater benthic diatom *Didymosphenia geminata* (Bacillariophyceae). *Journal of Phycology* 47: 981–989, http://dx.doi.org/10.1111/j.1529-8817.2011.01029.x

Kirkwood AE, Shea T, Jackson LJ, McCauley E (2007) *Didymosphenia geminata* in two Alberta headwater rivers: an emerging invasive species that challenges conventional views on algal bloom development. *Canadian Journal of Fisheries and Aquatic Sciences* 64: 1703–1709, http://dx.doi.org/10.1139/f07-152

Kociolek JP, Stoermer EF (1988) A preliminary investigation of the phylogenetic relationships among the freshwater, apical pore field-bearing Cymbelloid and Gomphonemoid diatoms (Bacillariophyceae). *Journal of Phycology* 24: 377–385, http://dx.doi.org/10.1111/j.1529-8817.1988.tb00187.x

Kociolek P (2011) *Cymbella janischii*. [online]. In Diatoms of the United States. Available from: http://westerndiatoms.colorado.edu/taxa/species/cymbella_janischii (Accessed 07 August 2013)

Krammer K, Lange-Bertalot H (1988) Bacillariophyceae 2.Teil: Bacillariaceae, Epithemiaceae, Surirellaceae. In: Ettl H, Gerloff J, Heynig H, Mollenhauer D (eds), *Süßwasserflora von Mitteleuropa*. Vol. 2/2 Gustav Fisher Verlag, Stuttgart, 596 pp

Kuhajek JM, Lemoine M, Kilroy C, Cary SG, Gerbeaux P, Wood SA (2014) Laboratory study of the survival and attachment of *Didymosphenia geminata* (Bacillariophyceae) in water sourced from rivers throughout New Zealand. *Phycologia* 53:

1–9, http://dx.doi.org/10.2216/13-145.1

Kumar S, Spaulding SA, Stohlgren TJ, Hermann KA, Schmidt TS, Bahls L (2009) Potential habitat distribution for the freshwater diatom *Didymosphenia geminata* in the continental US. *Frontiers in Ecology and Environment* 7: 415–420, http://dx.doi.org/10.1890/080054

Lang I, Kaczmarska I (2011) A protocol for a single-cell PCR of diatoms from fixed samples: method validation using *Ditylum brightwellii* (T. West) Grunow. *Diatom Research* 26: 43–49, http://dx.doi.org/10.1080/0269249X.2011.573703

Lyngbye HC (1819) *Tentamen Hydrophytologiae Danicae Continens omnia Hydrophyta Cryptogama Daniae, Holsatiae, Faeroae, Islandiae, Groenlandiae hucusque cognita, Systematice Disposita, Descripta et iconibus illustrata, Adjectis Simul Speciebus Norvegicis*. Hafniae, Copenhagen, 248 pp, 70 pls

Mann DG, Sato S, Trobajo R, Vanormelingen P, Souffreau C (2010) DNA barcoding for species identification and discovery in diatoms. *Cryptogamie, Algologie* 31: 557–577

Mass Executive Office Energy and Environmental Affairs (Mass EOEEA) (2013) State Environmental Officials Warn Residents About the Discovery of Didymo [online]. Available from: http://www.mass.gov/eea/pr-2013/didymo-discovery.html (Accessed 30 May 2013)

Metropolitan District Commission (MDC) (2013) The Metropolitan District Hartford, CT [online]. Available from: http://www.themdc.com/west_hartfordres.shtml (Accessed 07 August 2013)

Metzeltin D, Lange-Bertalot H (1995) Kritische Wertung der Taxa in *Didymosphenia* (Bacillariophyceae). *Nova Hedwigia* 60: 381–405

Morales EA, Siver PA, Trainor FR (2001) Identification of diatoms (Bacillariophyceae) during ecological assessments: comparison between light microscopy and scanning electron microscopy techniques. *Proceedings of the Academy of Natural Sciences of Philadelphia* 151: 95–103, http://dx.doi.org/10.1635/0097-3157(2001)151[0095:IODBDE]2.0.CO;2

Pite DP, Lane KA, Hermann AK, Spaulding SA, Finney BP (2009) Historical abundance and morphology of *Didymosphenia* species in Naknek Lake, Alaska. *Acta Botanica Croatica* 68: 183–197

Pniewski FF, Friedl T, Latała A (2010) Identification of diatom isolates from the Gulf of Gdańsk: testing of species identifications using morphology, 18S rDNA sequencing and DNA barcodes of strains from the Culture Collection of Baltic Algae (CCBA). *Institute of Oceanography* ISSN 1730-413X

Round FE, Crawford RM, Mann DG (1990) The diatoms. Biology and morphology of the genera. Cambridge University Press, Cambridge, 747 pp

Spaulding SA (2010) *Didymosphenia geminata* (online). In: Diatoms of the United States. Available from: http://westerndiatoms.colorado.edu/taxa/species/didymosphenia_geminata (Accessed 07 August 2013)

Spaulding SA, Elwell L (2007) Increase in nuisance blooms and geographic expansion of the freshwater diatom *Didymosphenia geminata*. U.S. Geological Survey Open-File Report 2007–1425, 38 pp

Spaulding SA, Hermann KA, Johnson T (2008) Confirmed distribution of *Didymosphenia geminata* (Lyngbye) Schmidt in North America. In: Bothwell ML, Spaulding SA (eds), Proceedings of the 2007 International Workshop on *Didymosphenia geminata*, Montréal, Québec. Canadian Technical Report of Fisheries and Aquatic Sciences 2795, pp 26–29

Suzawa T, Seino S, Mayama S (2011) Blooms of *Cymbella janischii* (A.W.F.Schmidt) De Toni accompanied by *Gomphoneis minuta* (Stone) Kociolek and Stoermer from the upper stream of the Chikugo River, Kyushu, Japan: possibility of new alien diatom species. *Diatom* 27: 58–64 (in Japanese)

Terry WA (1907) A partial list of Connecticut diatoms with some account of their distribution in certain parts of the state. *Rhodora* 104: 125–140

Valéry L, Fritz H, Lefeuvre J-C, Simberloff D (2009) Invasive species can also be native. *Trends in Ecology & Evolution* 24: 585, http://dx.doi.org/10.1016/j.tree.2009.07.003

Whitton BA, Ellwood NTW, Kawecka B (2009) Biology of the freshwater diatom *Didymosphenia*: a review. *Hydrobiologia* 630: 1–37, http://dx.doi.org/10.1007/s10750-009-9753-5

Morphological and molecular characterization of a new species in the genus *Didymosphenia* and of *Cymbella janischii* (Bacillariophyta) from Connecticut, U.S.A.

DIBA A. KHAN-BUREAU[1,2*], EDUARDO A. MORALES[3], LUC ECTOR[4], MICHAEL S. BEAUCHENE[5] AND LOUISE A. LEWIS[6]

[1] *Department of Natural* Resources *and the Environment, University of Connecticut, 75 North Eagleville Road, Storrs Connecticut, 06269, U.S.A*

[2] *Three Rivers Community College, 574 New London Turnpike, Norwich, Connecticut, 06360, U.S.A.*

[3] *Herbario Criptogámico, Universidad Católica Boliviana "San Pablo", Calle M. Márquez esq. Plaza Jorge Trigo s/n, P.O. Box 5381, Cochabamba, Bolivia*

[4] *Department of Environment and Agro-biotechnologies (EVA), Public Research Center – Gabriel Lippmann, 41 rue du Brill, L-4422 Belvaux, Luxembourg*

[5] *Inland Fisheries Division, Department of Energy and Environmental Protection, 79 Elm Street, Hartford, Connecticut, 06106-5127, U.S.A.*

[6] *Department of Ecology and Evolutionary Biology, University of Connecticut, 75 North Eagleville Road, Storrs Connecticut, 06269, U.S.A.*

*Corresponding author: (diba.khan-bureau@uconn.edu)

Short running title: A new species of *Didymosphenia* from Connecticut, U.S.A.

ABSTRACT

Two non-indigenous stalk-forming diatoms that were recently observed in the West Branch of the Farmington River, a tributary of the Connecticut River in Connecticut (U.S.A.), are characterized morphologically and through analysis of DNA sequence data. *Cymbella janischii*, the dominant stalk-forming species in this river during the summer of 2012, previously had not been found in the northeastern U.S.A. Samples of *C. janischii* were examined microscopically and used for molecular analysis of the V4 region of the 18S rDNA gene, providing the first DNA sequence for this species. Phylogenetic analysis indicated that the four independent sequences of *C. janischii* from Connecticut were distinct from, but related to, published sequences of *C. proxima, C. tumida*, and *Didymosphenia geminata*. A second stalk-producing diatom new to this region, resembling *D. geminata*, was found from November 2012 to June 2013. This new taxon was first reported as *Didymosphenia* sp. Over this time period, the observed cells had an unusual compressed morphology and small size compared to *D. geminata*. Sequences of the V4 region obtained from independent direct polymerase chain reactions (PCR) of single cells isolated from the Connecticut samples indicated a close relationship to two published sequences of *D. geminata* from Italy and New Zealand. Frustules of the cells used in the PCR reactions were recovered and examined using scanning electron microscopy (SEM) thus providing a direct link between the observed morphology and sequence data. The morphology of the unusual Connecticut *Didymosphenia* taxon was compared with that of other *Didymosphenia* taxa, being most similar to *D. pumila* and *D. sibirica*. The *Didymosphenia* taxon from Connecticut had a triundulate frustule morphology with a length of 50–60 µm. Given the unique combination of morphological features of this diatom, including its size, striae density, areolae structure and number of stigmata, it is hereby proposed as a new species, *Didymosphenia hullii* Khan-Bureau sp. nov.

Keywords: 18S V4 rDNA, Benthic diatoms, Cymbelloid diatoms, Gomphonemoid diatoms, Invasive species, Nuisance species, U.S.A.

INTRODUCTION

The freshwater stalk-forming diatom *Didymosphenia geminata* (Lyngbye) M. Schmidt is a well known invasive and nuisance species with an ability to form copious extracellular polymeric substances that form the stalks (Blanco & Ector 2009, Aboal et al. 2012). *Cymbella janischii* (A.W.F. Schmidt) De Toni, another stalk-forming diatom with abundant stalk growth is often mistaken for *D. geminata* at the macroscopic level with tufts that are similar in appearance (Pite 2009, Whitton et al. 2009). Both species are commonly referred to as rock snot. Under certain environmental conditions these species grow prolifically, forming thick mats that cover sections of the river substrate, negatively impacting other aquatic organisms (Kilroy 2004, Spaulding & Elwell 2007, Kumar et al. 2009, Morales et al. 2012, Zgłobicka 2013, Kuhajek & Wood 2014). Unlike *C. janischii*, *D. geminata* prefers oligotrophic, cold, and low soluble reactive phosphorous environments, which may in part cause unusual overgrowth conditions (Krammer & Lange-Bertalot 1986, Kilroy & Bothwell 2011, Bothwell et al. 2012, 2014). In addition, *D. geminata* establishment is influenced by a structurally suitable substrate, the development of a pad which adheres to the substrate, and the orientation of the cell upon descending (Kilroy & Bothwell 2014, Kuhajek & Wood 2014, Kuhajek et al. 2014).

Many states throughout the U.S.A. are monitoring their waterways for *D. geminata* because of its expanding geographical range (Kuhajek & Wood 2014). In the U.S.A., *D. geminata* was transported from the western states into several southeastern states, and more recently to northeastern states (Bothwell & Spaulding 2008, Blanco & Ector 2009, Spaulding 2010). In May 2013, Massachusetts first recorded and confirmed an occurrence of *D. geminata* with growth lasting two months (personal comm. A. Madden MA. Div. Fisheries and Wildlife).

The Connecticut Department of Energy and Environmental Protection (CT DEEP) started monitoring the West Branch of the Farmington River after purported *D. geminata* tufts were observed in 2011. In July 2012 reports of mucilaginous tufts occurring downstream of the original location in 2011 were later confirmed to be substantial growth of *C. janischii*. In addition, an unusual morphological population of *Didymosphenia* sp. was found in November 2012 (Khan-Bureau et al. in review). The present study characterizes the morphology, taxonomy, and phylogeny of these two diatoms from the West Branch of the Farmington River in Connecticut. We show that *Didymosphenia* sp. is distinct from other species of *Didymosphenia* and propose a new species to accommodate this taxon.

MATERIAL AND METHODS

The West Branch of the Farmington River (WBFR) is located in Northwestern Connecticut, U.S.A. It is one of several sub-basins within the greater Farmington River regional basin, a significant tributary to the lower Connecticut River (CT DEEP 2011). The WBFR is impounded twice, first at the Colebrook River Reservoir and then at the West Branch Reservoir. The Metropolitan District Commission (MDC) operates West Branch Reservoir and has a contractual agreement to maintain a minimum discharge from the nutrient-poor hypolimnion (personal comm. C. Bellucci, CT DEEP, MDC 2013). The WBFR has stable flow regimes and substrate stability because discharge from this reservoir is managed. It also has very cold and nutrient-poor water, making it conducive for the growth of *Didymosphenia* sp., as reported by Kilroy (2004), Spaulding & Elwell (2007), and Bothwell & Kilroy (2011) for *D. geminata*. The river is a destination trout fishery for national and international anglers (personal comm. C. Bellucci, CT DEEP, MDC 2013). Consequently, the risk of nuisance algae is a concern for state environmental agencies as well as the communities living in close proximity of the river.

Benthic samples were collected from the West Branch of the Farmington River in July

2012 – June 2013. Samples of the extracellular polysaccharide stalks were taken from rock substrate, submerged vegetation, and overhanging tree branches, and placed in Whirlpac bags. The latter were placed on ice, and transported to the lab for processing.

***DNA extraction, PCR, and cloning of* Cymbella janischii**

One water sample from July 2012 was used for molecular characterization of *C. janischii*. This sample was centrifuged, rinsed with deionized water, and then split into three replicate 50 µL microtubes. DNA extraction was accomplished using the MoBio PowerLyser™ Soil Extraction kit. PCR amplification of the V4 region of the 18S rDNA gene was achieved using primers D512 F and D978 R (Zimmerman et al. 2011). The PCR temperature profile included an initial denaturation step at 94°C (2 min), then five cycles consisting of denaturation at 94°C (45 s), annealing at 52°C (45 s) and elongation at 72°C (1 min), followed by 35 cycles in which the annealing temperature was lowered to 50°C, and a final elongation at 72°C (10 min). Resulting PCR products were visualized on a Syber Safe stained agarose gel, then quantified with a Nano Drop spectrophotometer. Cloning of PCR products was performed using Invitrogen TOPO® TA Cloning® Kit. Plasmid Prep followed using the QIAprep® Spin Miniprep Kit.

***Direct PCR of single cells of* Didymosphenia**

To best match DNA sequences with a specific morphotype, direct PCR was performed on cells that later were used for SEM imaging. Initially several cells from fresh samples, were isolated using a micropipette and placed in a 0.2 mL PCR tube. From these tubes, one individual cell of *Didymosphenia* sp. was placed in three replicate PCR tubes and washed 3–5 times (Lang & Kaczmarska 2011). After the final wash and centrifugation, the supernatant was removed and replaced by 1 µL of sterile water. The samples were then heated at 95°C on a thermocycler for 10 min prior to PCR to open the frustules for DNA extraction (Lang & Kaczmarska 2011). PCR amplification of the V4 region of the 18S rDNA gene was achieved using the primers D512F and

D978R of Zimmerman et al. (2011). The PCR mix consisted of 10 μL GoTaq® Green Master Mix, 0.5 μL of each primer (Zimmerman et al. 2011), and sterile deionized water for a final volume of 20 μL in the PCR tubes, each containing a single *Didymosphenia* sp. cell. The PCR temperature profile used for the amplification of *C. janischii* DNA was also employed here.

***Sequencing of* Didymosphenia *sp. and* C. janischii**

The cloned fragments and cleaned PCR fragments were directly sequenced using the amplification primers of Zimmerman et al. (2011). The sequencing cycle comprised 27 cycles of denaturing at 96°C for 30 s, annealing at 50°C for 15 s, and extension at 60°C for 4 min, using the Big Dye™ Terminator Cycle Sequencing Reaction Kit (Applied Biosystems, Foster City, CA, USA). Products of cycle sequencing were cleaned using ethanol precipitation and analyzed on ABI 3100 DNA Sequencer™ (Applied Biosystems, Foster City, CA, USA). Contigs of individual reads were assembled in Geneious© (Geneious 2013), to produce consensus sequences. These were compared to published data in GenBank, using the BLASTn tool, to obtain information on the closest matches. The sequences were aligned against a sampling of the sequences presented in Kermarrec et al. (2011) and the sequence of *Cocconeis stauroneiformis* (W. Smith) Okuno AB430614 (Sato et al. 2008). Sequences were aligned manually using Geneious© (Geneious 2013). Confidence of branch support was assessed using MrBayes (Huelsenbeck & Ronquist 2001, Ronquist & Huelsenbeck 2003). Bayesian analyses were run for 10^6 x 3 generations with one cold chain and three heated chains, and sampling every 300 generations. Burn-in length was 300,000. Alignments will be available from www.treebase.org.

Light microscopy (LM)

Prior to acid washing the samples, live samples were placed on a microscope slide with a coverslip overlain and then viewed at 200 and 400X magnifications using a BX 60 Olympus microscope. Images were digitally captured using an Olympus DP 25 camera and cellSens

software then viewed to identify the taxa according to Krammer & Lange-Bertalot (1986), Round et al. (1990), and three online databases, the ANSP Algae Image Database (http://diatom.ansp.org/algae_image/), Diatoms of the United States (http://westerndiatoms.colorado.edu/), and the Great Lakes Image Database (http://www.umich.edu/~phytolab/GreatLakesDiatomHomePage/top.html). For permanent slide preparation the river samples were centrifuged to concentrate the diatom cells to the bottom of the microtube. The supernatant was poured off and distilled water was added. Samples were then simmered on a hot plate in a 1:1 ratio of water and 68 % nitric acid to oxidize organic matter, then taken off the hotplate and allowed to cool for several minutes. Deionized water was used to rinse the samples of the acid, rinsed 4–5 times to neutralize samples, and then centrifuged to concentrate the diatom frustules (following the protocol of R. Lowe pers. comm.). After air-drying the diatom samples overnight on coverslips, frustules were mounted on glass microscope slides in the mounting medium NAPHRAX®, heated on a hot plate and then cooled to produce permanent vouchers. The diatom frustules were examined at 600 and 1000X magnifications with a BX 60 Olympus microscope. One hundred and twenty five valves were measured. Images were captured using an Olympus DP 25 color camera (2560 x 1920 pixels). Image plates were created using Adobe® Creative Suite® 6 Photoshop.

Scanning electron microscopy (SEM)

The 0.2 mL Eppendorf® tube residual from the original PCR product was saved for SEM verification of the isolated cell for identification purposes. The single frustule was retrieved from several individual PCR tubes, washed in deionized water, and centrifuged to ensure that the cell was not discarded and could later be found at the bottom of the tube for SEM verification. The supernatant was removed and replaced with 25 µl of deionized water and transferred onto a 25 mm, 3µm pore size polycarbonate Millipore filter (Lang & Kaczmarska 2011). The filter was

placed on a SEM stub with double-sided tape. For *C. janischii* we followed the methodology of Morales et al. (2001), the stubs were coated for 1 min at 1.8 kV with gold/palladium using a Polaron sputter coater.

A mixture of glutaraldehyde and Bold Basal Medium (BBM) was used to prepare the stalk material for SEM preparation of the *Didymosphenia* sp. tufts. The tufts were placed in the mixture and centrifuged at 600 x g low speed to guard against stalk material damage. The supernatant was discarded and the tufts were re-suspended in BBM combined with 4% glutaraldehyde overnight. The samples were placed into 2% osmium tetroxide for 2 h. The samples were dehydrated through a graded ethanol (EtOH) series of 30, 50, 70, 95 and 100% with the sample remaining in each EtOH shallow glass petri dish bath for 15 m on ice. Critical point drying using a Tousimis 931.GL apparatus was utilized followed by sputter coating. The stubs were viewed with the field emission Leo/Zeiss DSM 982 and a field emission FEI Nova Nano 450 scanning electron microscope.

RESULTS

Phylogenetic analysis

Three independent V4 rDNA sequences of *Didymosphenia* sp., resulting from the three single cell isolations, plus four sequences from the cloned material of *C. janischii* were obtained. The V4 sequences were 334–410 bp in length, and these were complied into a final alignment with selected published diatom sequences to produce an alignment of 1816 nucleotides in length. In the Bayesian phylogenetic tree (Fig. 1) sequences of *C. janischii* from Connecticut were closely related to those of *C. tumida* (Brébisson) Van Heurck and *C. proxima* Reimer and *D. geminata*. Currently there are no sequence data for *C. janischii* in the NCBI sequence database with which to compare the newly acquired sequences. The newly sequenced V4 fragments from *Didymosphenia* sp. were identified as close matches to two published sequences of *D. geminata*

from Italy (JN790293.1) and New Zealand (JN680079), and *C. tumida* (JN790274) and the sequences of *C. janischii* from this study. Phylogenetic analysis of sequences from this study and from Kermarrec et al. (2011) suggests that *D. geminata* (currently classified in the Gomphonemoid clade) and *C. janischii* (the Cymbelloid clade) are closely related to each other. The tree topology (Fig. 1) illustrates that *Didymosphenia* sp. and *C. janischii* are sister taxa indicating that *Cymbella* is not monophyletic as originally described (Kociolek & Stoermer 1988, Kermarrec et al. 2011).

LM and SEM analysis

Cells of *C. janischii* were an average of 130 µm long, consistent with the published range of 105–383 µm (Bahls 2007) (Figs 2–5). Several LM images were taken of live cells of *Didymosphenia* sp. prior to preparing the samples for acid cleaning. Figures 6–11 show recently divided cells. Sexual reproduction was not observed. The valve and girdle views of *Didymosphenia* sp. illustrate the cell size variation of this population (Figs 12–26). *Didymosphenia* sp. cells from the PCR reactions were successfully retrieved from the original PCR tubes, placed on a millipore filter and stub and viewed with SEM. The recovered cells are shown in Figures 27–29 as fractured, but they were adequate for identification. These sequences from the single cell isolations were used to establish a phylogenetic tree (see Fig. 1). The SEM images of the *Didymosphenia* sp. cells attached to their stalks further demonstrate that the average cell size was 50-60 µm and the valve morphology compressed (Figs 30–38). This is in contrast to cells of *D. geminata* found in other North American sites, including Massachusetts, which are more robust and up to 137 µm (Kilroy 2004, Spaulding 2010). SEM images of the apical pore field, stigmata, striae, shape, areolae and length of the frustule demonstrate that this taxon differs from other members of *Didymosphenia* (Figs 39–44, Table 1). *Didymosphenia* sp. is commonly 50–60 µm in length, the cells of 38 µm and 68 µm observed occasionally. The

width is 26.5–30.5 µm with 1–4 stigmata and 9–11 striae in 10 µm. The areolae have deep inclined walls within the valves surrounded by spine-like projections (spines) with dendritic slits below the spines, and are most similar to the areolae of *D. geminata* and *D. clavaherculis* (Ehrenberg) Metzeltin & Lange-Bertalot as described by Metzeltin & Lange-Bertalot (1995).

The *D. geminata*-like diatom that occurs in the West Branch of the Farmington River is unusual. Unlike the diatom that was found in Massachusetts, it is morphologically distinct from *D. geminata* and other species in this genus, warranting unique species-level status.

***Didymosphenia hullii* Khan-Bureau** sp. nov. (Figs 6–44)

Description: *Didymosphenia hullii* Khan-Bureau sp. nov. can be motile or form colonies of cells on long stalks. Heteropolar, headpole is capitate with a rotund shorter compressed headpole and footpole than *D. geminata*. The footpole is slightly capitate though blunt. The footpole has an apical pore field of very small spherical perforations that are present where the stalk growth originates. The frustule holds together the epivalve and hypovalve through four girdle bands patterned with raised pustules. From the girdle view the headpole is broad and tapers to the footpole similar to a wedge or a V shape. The valve has a length generally in the range of 50–60 µm with cells slightly smaller at 38 µm and larger to 68 µm occassionally observed, and the width range of 26.5–30.5 µm. There are 1–4 stigmata present although the majority typically had 2–3 stigmata. The central area is inflated and elliptical with 9–11 striae in 10 µm that are radiate and have irregular short and long lengths. Larger pentagon and square shaped depressions with pores (areolae) are present throughout the valve in complex deep wells that are surrounded by spines. The distal raphe ends quickly to a tight curve or hook shape, but does not go through the apical pore field. This taxon has both an asymmetrical apical and transpical axis; nuclear encoded rRNA sequence = GenBank accessions KJ160170, KJ160171, KJ160172.

HOLOTYPE: CONN00178537, collected 29 November 2012 (University of Connecticut Herbarium, Storrs Connecticut U.S.A.).

ISOTYPES: Voucher number pending (Academy of Natural Sciences, Philadelphia, PA). HCUCB D-00791 (*Herbario Criptogámico, Universidad Católica, San Pablo, Cochabamba, Bolivia*). Voucher number pending (Royal Botanic Garden Edinburgh 20A Inverleith Row Edinburgh EH3 5LR United Kingdom).

TYPE LOCALITY: The West Branch of the Farmington River, a tributary of the Connecticut River in Barkhamsted, Connecticut U.S.A. (41.960° N 73.017° W). Collected 29 November 2012, by Diba Khan-Bureau. Samples were taken from the epilithon.

HABITAT: Abundant growth occurred on a wide range of large cobbles and boulders covering the river substrate, bank to bank.

ETYMOLOGY: The species is named in honor of the late David Hull MD, Director of Transplant at Hartford Hospital in Connecticut. He enjoyed nature and aspired to understand the many facets of science.

Didymosphenia hullii, a visually attractive diatom, resembles *D. geminata*, *D. sibirica* (Grunow) M. Schmidt, *D. pumila* Metzeltin & Lange-Bertalot, *D. curvata* (Skvortsov & K.I. Meyer) Metzeltin & Lange-Bertalot, and *D. tatrensis* Mrozińska, Czerwik-Marcinkowska & Gradziński, but differs in length, striae density, number of stigmata, and areolae (Metzeltin & Lange-Bertalot 1995, Mrozińska et al. 2006) (Table 1). The areolae of *D. hullii* are similar to that of *D. geminata* and *D. clavaherculis.* Unlike *D. tatrensis, D. sibirica* and *D. pumila, D. hullii* has deep inclined walls within the valves surrounded by spines with dendritic slits below the spines of the interior walls (Mrozińska et al. 2006, Metzeltin & Lange-Bertalot in press). *Didymosphenia hullii* morphology is most similar to *D. sibirica* and *D. pumila*, which are known from Russia (Metzeltin & Lange-Bertalot 1995, Mrozińska et al. 2006). These three species are the smaller of

the *Didymosphenia* taxa. They are compressed and their basal poles are much less elongated than *D. geminata*, and they differ in valve length and width, striae density, and number of stigmata (Table 1). *Didymosphenia sibirica* was reported by Dawson (1973) and Stoermer et al. (1986) as having only one isolated stigma internally and referred to this as a raised convolution. Subsequently Metzeltin & Lange-Bertalot (1995) described *D. sibirica* with 1–3 stigmata and *D. pumila* having 1–2 stigmata, whereas *D. hullii* has 1–4 stigmata internally. *Didymosphenia sibirica* has much finer striations and has a strong degree of longitudinal asymmetry (Stoermer et al. 1986) not seen in *D. hullii*.

DISCUSSION

In this study we characterize two nuisance stalk-forming diatoms in Connecticut contributing information on the morphology, variation and phylogenetic relationships of cymbelloid diatoms. The first sequence data for *Cymbella janischii* are provided indicating that *C. janischii*, *D. hullii*, and *D. geminata* are closely related, and we describe the new taxon, *Didymosphenia hullii*. In the U.S.A., *C. janischii* is described as endemic to the Pacific Northwest. Outside of the Pacific Northwest, *C. janischii* had only been observed in four other states, Arizona, Colorado, New York and Oklahoma (Bahls 2007), until recently reported in Connecticut in July 2012 (Khan-Bureau et al. in review). Suzawa et al. (2011) confirmed blooms of the non-indigenous *C. janischii* in Japan and reported that it was introduced from North America. Phylogenetic analysis of four independent V4 sequences of *C. janischii* from Connecticut indicate that this species is distinct from, but related to published sequences of *Cymbella proxima*, *C. tumida* (Cymbellaceae), and *D. geminata* (Gomphonemataceae). The *C. janischii* sequences also indicate that this species is distantly related to most sequences from the family Cymbellaceae, such as *Cymbella affinis* Kützing and *C. cymbiformis* (Ehrenberg) Grunow. These data provide additional evidence that the taxonomy of the families Gomphonemataceae and Cymbellaceae

require re-evaluation (Kociolek & Stoermer 1988, Nakov & Theriot 2009, Kermarrec et al. 2011, Graeff & Kociolek 2013).

Ten species (and several varieties of *D. geminata*) are currently described within the genus *Didymosphenia*: *D. clavaherculis* (Ehrenberg) Metzeltin & Lange-Bertalot 1995, *D. curvata* (Skvortzow & C.I. Meyer) Metzeltin & Lange-Bertalot 1995, *D. curvirostrum* (Tempère & Brun in Brun & Tempère) M. Schmidt 1899, *D. dentata* (Dorogostaïsky) Skvortzow & C.I. Meyer 1928, *D. fossilis* Horikawa & Okuno in Okuno 1944, *D. geminata* (Lyngbye) M. Schmidt in Schmidt et al. 1899, *D. lineata* Skabichevskij 1983, *D. pumila* Metzeltin & Lange-Bertalot 1995, *D. sibirica* (Grunow) M. Schmidt in Schmidt et al. 1899, and *D. tatrensis* Mrozińska, Czerwik-Marcinkowska & Gradzinski 2006. Additionally twelve new *Didymosphenia* taxa have been described (personal comm. D. Metzeltin). The range of this species has expanded, with *D. geminata* reported from midwestern and eastern states (Kilroy 2004, Spaulding & Elwell 2007, Blanco & Ector 2009, Kumar et al. 2009, Kilroy & Unwin 2011, Bothwell et al. 2012), and in new locations in Canada (Kirkwood et al. 2007, Gillis & Chalifour 2009, Lavery et al. 2014) and South America (Kilroy & Unwin 2011, Segura 2011, Morales et al. 2012, Rivera et al. 2013, Sastre et al. 2013, Reid & Torres 2014). Only one species of *Didymosphenia*, *D. geminata*, occurs within the geographic boundaries of the continental U.S.A., specifically in the western states, although *D. clavaherculis* was documented in Alaska (Spaulding 2010).

The type material of *D. geminata* had not been available until recently (personal comm. D. Metzeltin) and morphological data limited even though it has been almost 200 years since Lyngbye first described *Didymosphenia geminata* as *Echinella geminata* in 1819 (Lyngbye 1819, Whitton et al. 2009, Blanco & Ector 2013). Most reports were based on drawings and light micrographs (Blanco & Ector 2013).

Despite the economic and ecological importance of *Didymosphenia* there are only two accessions of this genus currently in GenBank, both reporting sequence of the 18S gene of *D. geminata*. We targeted the V4 region of the 18S gene to make the new data from *D. hullii* comparable to the published sequences of *D. geminata* and related taxa. The results of the present study indicate that *D. geminata* and *D. hullii* are more closely related to *C. janischii* than to *Gomphonema* and other genera (*Reimeria*) in Gomphonemataceae, further illustrating the paraphyly of Cymbellaceae and Gomphonemataceae (Kociolek & Stoermer 1988, Nakov & Theriot 2009, Kermarrec et al. 2011, Graeff & Kociolek 2013). A more variable marker is needed to better differentiate among taxa and facilitate identification of new or cryptic species of algae. This is especially true for the species of *Didymosphenia* considering that it can grow prolifically when favorable conditions exist while its geographical range has expanded. However, finding a suitable species level marker has proven challenging for some algae and specifically diatoms (Evans et al. 2007, Hall et al. 2010, Hamsher et al. 2011, Zimmerman et al. 2011, Luddington et al. 2012, Kermarrec et al. 2014). We attempted to use a more variable marker, the plastid *rbcL*, although we were unsuccessful.

In contrast to the variation among sequences of *Cymbella*, data from the V4 region were unable to separate *D. hullii* and *D. geminata* due to very low sequence variation, indicating that these two taxa are very closely related and maybe evolutionarily young. The lack of variation coupled with a lack of published sequences from the other *Didymosphenia* species means that we are currently unable to trace the origin of the new species, be it from *D. geminata*, *D. pumila* or other species. Clearly, more studies are needed that relate the morphology of a diatom to a corresponding molecular signature if we are to elucidate relationships among the species of *Didymosphenia*, to better understand the spread of these nuisance taxa, and to connect physiological preferences and tolerances to particular diatom species.

ACKNOWLEDGEMENTS

We thank the Connecticut Institute of Water Resources (CTIWR) for funding this research. We thank K. Fučíková, S. Olm, C. Lo, M. Letsch, H. McManus and acknowledge the late F. Trainor for helpful discussion. This work made use of the computer cluster maintained by the Bioinformatics Facility (Biotechnology/Bioservices Center) at the University of Connecticut. We are grateful for D. Metzeltin's expert opinion and support. We thank Three Rivers Community College's Dean of IT S. Goetchius, K. Barfield, S. Cohen, D. Jewett, and President G. S. Jones for providing continual support for this project, J. Sgro, John Carroll University, for providing verification of species, M. Cantino and S. Daniels of the EM lab at UCONN, NSF 1126100 supporting the purchase of the Nova Nano SEM, V. Kask scientific illustrator in Biological Central Services at UCONN, J. Morrison, East Hartford, USGS, L. Matthews VT DEC, M. Becker and C. Bellucci, CT DEEP for their assistance.

REFERENCES

ABOAL M., MARCO S., CHAVES E., MULERO I. & GARCÍA-AYALA A. 2012. Ultrastructure and function of stalks of the diatom *Didymosphenia geminata*. *Hydrobiologia* 695: 17–24.

BAHLS L.L. 2007. *Cymbella janischii*—Giant endemic diatom of the Pacific Northwest: morphology, ecology and distribution compared to *Cymbella mexicana*. *Northwest Science* 81: 284–292.

BLANCO S. & ECTOR L. 2009. Distribution, ecology and nuisance effects of the freshwater invasive diatom *Didymosphenia geminata* (Lyngbye) M. Schmidt: a literature review. *Nova Hedwigia* 88: 347–422.

BLANCO S. & ECTOR L. 2013. The nomenclatural history of *Didymosphenia geminata* (Lyngbye) M. Schmidt in Schmidt *et al.* (Bacillariophyta) and related taxa. *Acta Nova* 6: 157–164.

BOTHWELL M.L. & SPAULDING S.A. 2008. Synopsis of the 2007 International Workshop on

Didymosphenia geminata. In: *Proceedings of the 2007 International Workshop on Didymosphenia geminata, Montréal, Québec* (Ed. by M.L. Bothwell & S.A. Spaulding), pp. xiii–xxxi. *Canadian Technical Report of Fisheries and Aquatic Sciences 2795.*

BOTHWELL M.L. & KILROY C. 2011. Phosphorous limitation of the freshwater benthic diatom *Didymosphenia geminata* determined by the frequency of dividing cells. *Freshwater Biology* 56: 565–578.

BOTHWELL M.L., KILROY C., TAYLOR B.W., ELLISON E.T., JAMES D.A., GILLIS C.-A., BLADON K.D. & SILINS U. 2012. Iron is not responsible for *Didymosphenia geminata* bloom formation in phosphorus-poor rivers. *Canadian Journal of Fisheries and Aquatic Sciences* 69: 1723–1727.

BOTHWEL M.L., TAYLOR B.W., KILROY C. 2014. The Didymo story: the role of low dissolved phosphorus in the formation of *Didymosphenia geminata* blooms. *Diatom Research* DOI: 10.1080/0269249X.2014.889041.

CT DEEP - CONNECTICUT DEPARTMENT OF ENERGY AND ENVIRONMENTAL PROTECTION 2011. http://www.ct.gov/deep/cwp/view.asp?Q=476204&A=4013

DAWSON P.A. 1973. Further observations on the genus *Didymosphenia* M. Schmidt–*D. sibirica* (Grun.) M. Schm. *British Phycological Journal* 8: 197–201.

EVANS K.M., WORTLEY A.H. & MANN D.G. 2007. An assessment of potential diatom "barcode" genes (*cox*1, *rbc*L, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist* 158: 349–364.

GENEIOUS 2013 (version 7.0.5) created by Biomatters. Available from http://www.geneious.com/

GILLIS C.-A. & CHALIFOUR M. 2009. Changes in the macrobenthic community structure following the introduction of the invasive algae *Didymosphenia geminata* in the Matapedia River (Québec, Canada). *Hydrobiologia* 647: 63–70.

GRAEFF C.L. & KOCIOLEK J.P. 2013. New or rare species of cymbelloid diatoms

(Bacillariophyceae) from Colorado (USA). *Nova Hedwigia* 97: 87–116.

HALL J.D., FUČÍKOVÁ K., LO C., LEWIS L.A. & KAROL K.G. 2010. An assessment of proposed DNA barcodes in freshwater green algae. *Cryptogamie, Algologie* 31: 529–555.

HAMSHER S.E., EVANS K.M., MANN D.G., POULÍČKOVÁ A. & SAUNDERS G.W. 2011. Barcoding diatoms: exploring alternatives to COI-5P. *Protist* 162: 405–422.

HUELSENBECK J.P. & RONQUIST F. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* 17: 754–755.

KERMARREC L., ECTOR L., BOUCHEZ A., RIMET F. & HOFFMANN L. 2011. A preliminary phylogenetic analysis of the Cymbellales based on 18S rDNA gene sequencing. *Diatom Research* 26: 305–315.

KERMARREC L., FRANC A., RIMET F., CHAUMEIL P., FRIGERIO J.M., HUMBERT J.F. & BOUCHEZ A. 2014. A next-generation sequencing approach to river biomonitoring using benthic diatoms. *Freshwater Science* 33: 349–363.

KHAN-BUREAU D.A., BEAUCHENE M.S., ECTOR L., MORALES E.A. & LEWIS L.A. In review. Observations of two nuisance stalk-forming diatoms (Bacillariophyta) from a river in Connecticut, Northeastern U.S.A.: *Didymosphenia* sp. and *Cymbella janischii*. *BioInvasions Records.*

KILROY C. 2004. A new alien diatom, *Didymosphenia geminata* (Lyngbye) Schmidt: its biology, distribution, effects and potential risks for New Zealand fresh waters. Available at: http://www.biosecurity.govt.nz/files/pests/didymo/didymo-preliminary-org-ianov-04.pdf.

KILROY C. & BOTHWELL M. 2011. Environmental control of stalk length in the bloom-forming, freshwater benthic diatom *Didymosphenia geminata* (Bacillariophyceae). *Journal of Phycology* 47: 981–989.

KILROY C. & UNWIN M. 2011. The arrival and spread of the bloom-forming, freshwater diatom, *Didymosphenia geminata*, in New Zealand. *Aquatic Invasions* 3: 249–262.

KILROY C. & BOTHWELL M. 2014. Attachment and short-term stalk development of *Didymosphenia geminata*: effects of light, temperature and nutrients. *Diatom Research*. DOI: 10.1080/0269249X.2014.889043

KIRKWOOD A.E., SHEA T., JACKSON L.J. & MCCAULEY E. 2007. *Didymosphenia geminata* in two Alberta headwater rivers: an emerging invasive species that challenges conventional views on algal bloom development. *Canadian Journal of Fisheries and Aquatic Sciences* 64: 1703–1709.

KOCIOLEK J.P. & STOERMER E.F. 1988. A preliminary investigation of the phylogenetic relationships among the freshwater, apical pore field-bearing Cymbelloid and Gomphonemoid diatoms (Bacillariophyceae). *Journal of Phycology* 24: 377–385.

KRAMMER K. & LANGE-BERTALOT H. 1986. Bacillariophyceae 1. Teil: Naviculaceae. In: *Süßwasserflora von Mitteleuropa* (Ed. by H. Ettl, J. Gerloff, H. Heynig & D. Mollenhauer), Band 2/1. Gustav Fisher Verlag, Stuttgart & New York. 876 pp.

KUHAJEK J.M. & WOOD S.A. 2014. Novel techniques for the short-term culture and laboratory study of *Didymosphenia geminata*. *Diatom Research* DOI: 10.1080/0269249X.2014.889042

KUHAJEK J.M., LEMOINE M., KILROY C., CARY S.C., GERBEAUX P. & WOOD S.A. 2014. Laboratory study of the survival and attachment of *Didymosphenia geminata* (Bacillariophyceae) in water sourced from rivers throughout New Zealand. *Phycologia* 53: 1–9.

KUMAR S., SPAULDING S.A., STOHLGREN T.J., HERMANN K.A., SCHMIDT T.S. & BAHLS L.L. 2009. Potential habitat distribution for the freshwater diatom *Didymosphenia geminata* in the continental US. *Frontiers in Ecology and Environment* 7: 415–420.

LANG I. & KACZMARSKA I. 2011. A protocol for a single-cell PCR of diatoms from fixed samples: method validation using *Ditylum brightwellii* (T. West) Grunow. *Diatom Research* 26: 43–49.

LAVERY J.M., KUREK J., RÜHLAND K.M., GILLIS C.A., PISARIC M.F.J. & SMOL J.P. 2014. Exploring the environmental context of recent *Didymosphenia geminata* proliferation in Gaspésie, Quebec, using paleolimnology. *Canadian Journal of Fisheries and Aquatic Sciences* 71: 1–11.

LUDDINGTON I.A., KACZMARSKA I. & LOVEJOY C. 2012. Distance and character-based evaluation of the V4 region of the 18S rRNA gene for the identification of diatoms (Bacillariophyceae). *PLoS ONE* 7: e45664 (11 pp.).

LYNGBYE H.C. 1819. *Tentamen Hydrophytologiae Danicae Continens omnia Hydrophyta Cryptogama Daniae, Holsatiae, Faeroae, Islandiae, Groenlandiae hucusque cognita, Systematice Disposita, Descripta et iconibus illustrata, Adjectis Simul Speciebus Norvegicis*. Hafniae. 248 pp., 70 pls.

MDC - METROPOLITAN DISTRICT COMMISSION 2013. http://www.themdc.com/

METZELTIN D. & LANGE-BERTALOT H. 1995. Kritische Wertung der Taxa in *Didymosphenia* (Bacillariophyceae). *Nova Hedwigia* 60: 381–405.

METZELTIN D. & LANGE-BERTALOT H. in press. The genus *Didymosphenia* M. Schmidt. A critical evaluation of established and description of 12 new taxa. *Iconographia Diatomologica* 25.

MORALES E.A., SIVER P.A. & TRAINOR F.R. 2001. Identification of diatoms (Bacillariophyceae) during ecological assessments: comparison between light microscopy and scanning electron microscopy techniques. *Proceedings of the Academy of Natural Sciences of Philadelphia* 151: 95–103.

MORALES E.A., RIVERA S.F., VEIZAGA A. & FIORINI R. 2012. *Didymosphenia geminata* (Lyngbye) M. Schmidt (Bacillariophyta), una especie invasora y potencial amenaza para ecosistemas acuáticos bolivianos. *Acta Nova* 5: 327–343.

MROZIŃSKA T., CZERWIK-MARCINKOWSKA J. & GRADZIŃSKI M. 2006. A new species of *Didymosphenia* (Bacillariophyceae) from the Western Carpathian Mountains of Poland and Slovakia. *Nova Hedwigia* 83: 499–510.

NAKOV T. & THERIOT E. 2009. Preliminary molecular phylogeny of the Cymbellales (Bacillaryophyceae). Program and Abstracts. 20th North American Diatom Symposium, Iowa Lakeside Lab, September 23-27, 2009, Milford Iowa, pp. 28–29.

PITE D.P., LANE K.A., HERMANN A.K., SPAULDING S.A. & FINNEY B.P. 2009. Historical abundance and morphology of *Didymosphenia* species in Naknek Lake, Alaska. *Acta Botanica Croatica* 68: 183–197.

REID B. & TORRES R. 2014. *Didymosphenia geminata* invasion in South America: Ecosystem impacts and potential biogeochemical state change in Patagonian rivers. *Acta Oecologica* 54: 101–109.

RIVERA P., BASUALTO S. & CRUCES F. 2013. Acerca de la diatomea *Didymosphenia geminata* (Lyngbye) M. Schmidt: su morfología y distribución en Chile. *Gayana Botánica* 70: 154–158.

RONQUIST, F. & HUELSENBECK J.P. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572–1574.

ROUND F.E., CRAWFORD R.M. & MANN D.G. 1990. The Diatoms: Biology and Morphology of the Genera. 747 pp., Cambridge University Press, Cambridge.

SASTRE A.V., SANTINELLI N.H., BAUER G.A., AYESTARÁN M.G. & UYUA N.M. 2013. First record of the invasive diatom *Didymosphenia geminata* (Lyngbye) Schmidt in a Patagonian Andean river of Argentina. *BioInvasions Records* 2: 11-17.

SATO S., MANN D.G., MATSUMOTO S. & MEDLIN L.K. 2008. *Pseudostriatella* (Bacillariophyta): a description of a new araphid diatom genus based on observations of frustule and auxospore structure and 18S rDNA phylogeny. *Phycologia* 47: 371–391.

SEGURA P. 2011. A slimy invader blooms in the rivers of Patagonia. *Science* 331: 18.

SKVORTZOW B.W. & MEYER C.I. 1928. A contribution to the diatoms of Baikal Lake. *Proceedings of the Sungaree River Biological Station* 1: 1–55.

SPAULDING S. 2010. *Didymosphenia geminata*. In: Diatoms of the United States. from http://westerndiatoms.colorado.edu/taxa/species/didymosphenia_geminata. Accessed on 19 December, 2013.

SPAULDING S.A. & ELWELL L. 2007. Increase in nuisance blooms and geographic expansion of the freshwater diatom *Didymosphenia geminata.* White Paper. Recommendations for response. Livingston, Montana: Federation of Fly Fishers and USA Environmental Protection Agency, USA. 33 pp.

STOERMER E.F, YU-ZAO Q. & LADEWSKI T.B. 1986. A quantitative investigation of shape variation in *Didymosphenia* (Lyngbye) M. Schmidt (Bacillariophyta). *Phycologia* 25: 494–502.

SUZAWA T., SEINO S. & MAYAMA S. 2011. Blooms of *Cymbella janischii* (A.W.F.Schmidt) De Toni accompanied by *Gomphoneis minuta* (Stone) Kociolek & Stoermer from the upper stream of the Chikugo River, Kyushu, Japan: possibility of new alien diatom species. *Diatom* 27: 58–64.

WHITTON B.A., ELLWOOD N.T.W. & KAWECKA B. 2009. Biology of the freshwater diatom *Didymosphenia*: a review. *Hydrobiologia* 630: 1–37.

ZGŁOBICKA I. 2013. Aspects of structural biology of *Didymosphenia geminata* (Lyngb.) M. Schmidt (Bacillariophyta). *International Journal on Algae* 15: 293–312.

ZIMMERMAN J., JAHN R. & GEMEINHOLZER B. 2011. Barcoding diatoms evaluation of the V4 sub region on the 18S rRNA gene, including new primers and protocols. *Organisms, Diversity & Evolution* 11: 173–192.

Figure Legends

Fig. 1: Bayesian phylogenetic tree based the V4 region of the 18S rDNA of selected published sequences of Cymbellales, plus newly-obtained sequences from three isolated single cells of *Didymosphenia hullii* and four clones of *Cymbella janischii* (GenBank Accession numbers are included in the taxon labels). The families of Gomphonemataceae and Cymbellaceae are indicated. Bayesian posterior probabilities (BPP) values indicate node support. Scale bar = expected number of substitutions/site.

Figs 2–5: LM images of field collected *C. janischii* cells showing their size distribution. Fig. 4: Central valve illustrates the distinct radiated striae with striae more closely together and longer pointed areolae in the central area, differentiating between other larger *Cymbella* taxa and *C. janischii*. Figs 2–4: Scale bars = 10 μm. Fig. 5: LM image of cell on its stalk prior to acid cleaning. Scale bar = 20 μm.

Figs 6–11: LM images of field collected *D. hullii.* Figs 6, 10, 11: Cells attached to their stalks illustrate cell division. Fig. 8: Acid washed single cell with two stigmata. Figs 6, 7, 10: Scale bars = 20 μm, Fig. 8: Scale bar = 10 μm, Fig. 9: Scale bar = 50 μm, Fig. 11: Scale bar = 40 μm.

Figs 12–21: LM images of *D. hullii* cells in valve view showing size variation. Scale bar = 10 μm.

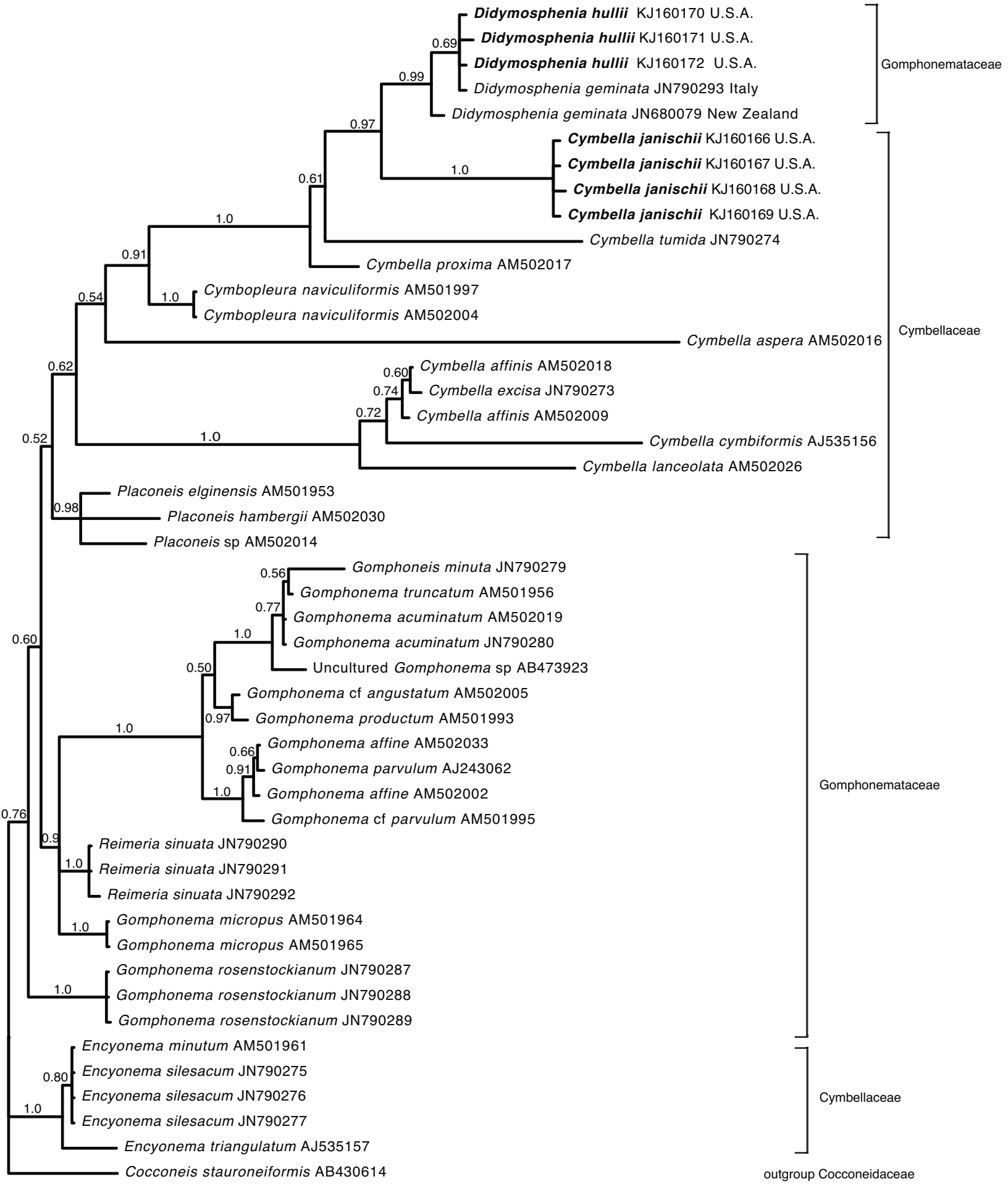
Figs 22–26: LM images of *D. hullii* cells in girdle view showing size variation. Scale bar = 10 μm.

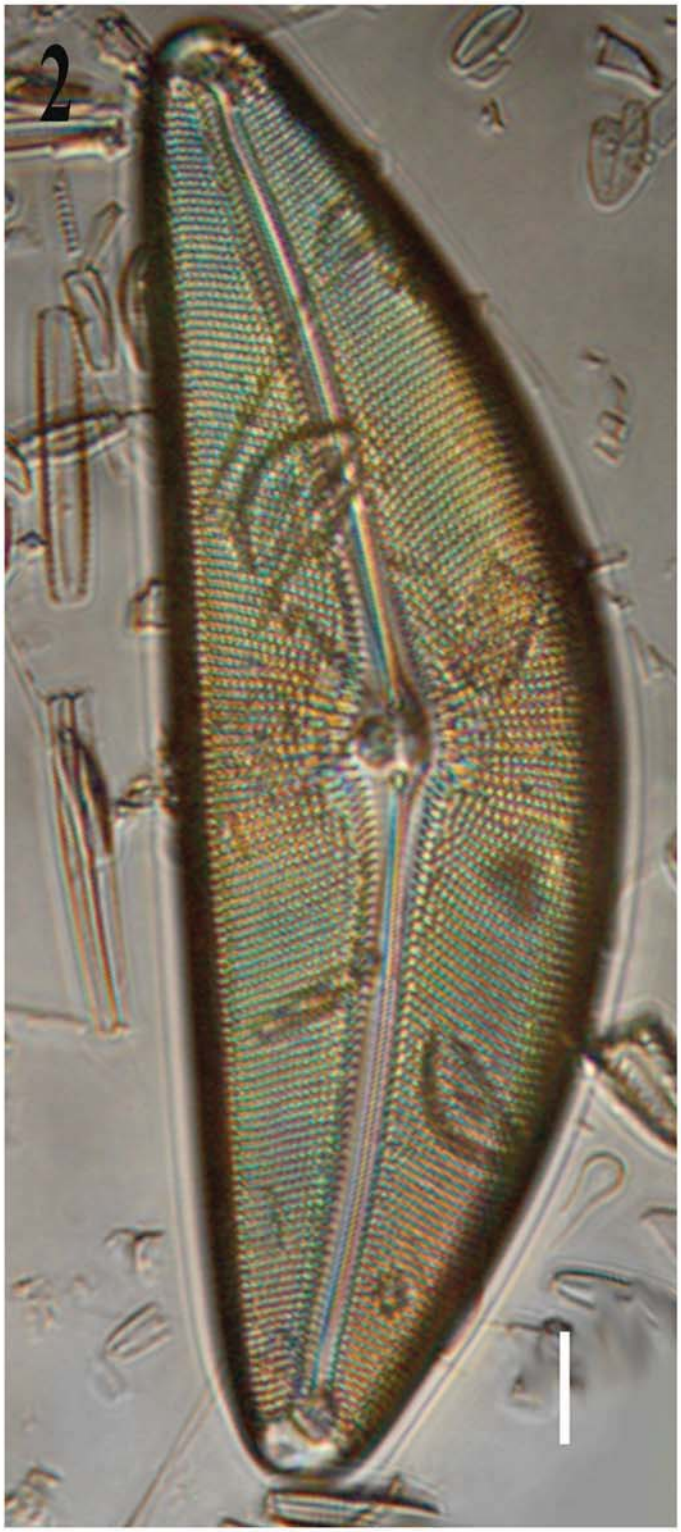
Figs 27–29: SEM images of single valves of *D. hullii* retrieved after PCR reactions. Fig. 27: The valve is fractured but still identifiable. Figs 27, 29: Scale bar = 20 μm. Fig. 28: High magnification image of a cell with two stigmata. Fig. 28: Scale bar = 5 μm.

Figs 30–38: SEM images of *D. hullii* on stalks from the West Branch of the Farmington River showing bifurcated cells attached to stalks. Figs 30, 31, 34: Scale bar = 20 μm, Figs 32, 33, 37:

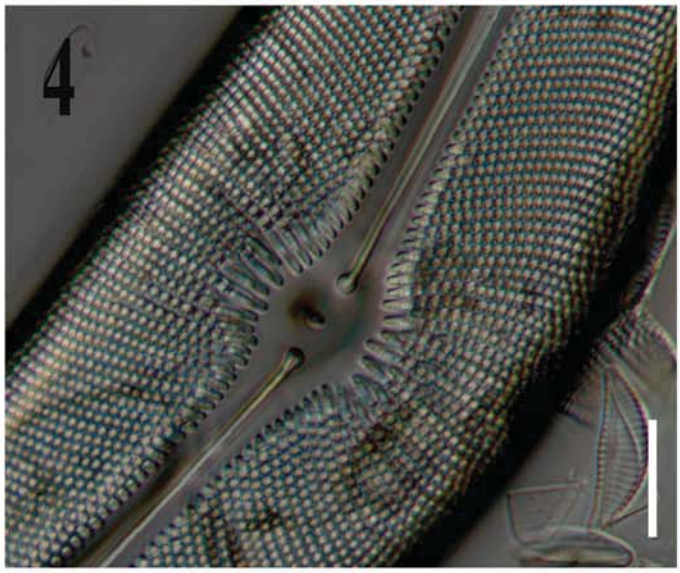
Scale bar = 50 μm, Fig. 35: Scale bar = 40 μm, Fig. 36: Scale bar = 5 μm, Fig. 38: Scale bar = 200 μm.

Figs 39–44: SEM images of *D. hullii*. Fig. 39: Internal view of the internal valve displaying three stigmata. Scale bar = 20 μm. Fig. 40: View of apical pore field, footpole. Scale bar = 5 μm. Figs 41, 42: Central valve view with stigmata, 1 stigma and 4 stigmata respectively and striae. Scale bar = 10 μm. Fig. 43: Internal view of central valve showing two stigmata and uniseriate striae. Scale bar = 5 μm. Fig. 44: External views of frustule girdle and valve morphology. Scale bar = 30 μm.


Didymosphenia hullii KJ160170 U.S.A.
Didymosphenia hullii KJ160171 U.S.A.
Didymosphenia hullii KJ160172 U.S.A.
Didymosphenia geminata JN790293 Italy
Didymosphenia geminata JN680079 New Zealand
Gomphonemataceae
Cymbella janischii KJ160166 U.S.A.
Cymbella janischii KJ160167 U.S.A.
Cymbella janischii KJ160168 U.S.A.
Cymbella janischii KJ160169 U.S.A.
Cymbella tumida JN790274
Cymbella proxima AM502017
Cymbopleura naviculiformis AM501997
Cymbopleura naviculiformis AM502004
Cymbella aspera AM502016
Cymbella affinis AM502018
Cymbella excisa JN790273
Cymbella affinis AM502009
Cymbella cymbiformis AJ535156
Cymbella lanceolata AM502026
Cymbellaceae
Placoneis elginensis AM501953
Placoneis hambergii AM502030
Placoneis sp AM502014
Gomphoneis minuta JN790279
Gomphonema truncatum AM501956
Gomphonema acuminatum AM502019
Gomphonema acuminatum JN790280
Uncultured Gomphonema sp AB473923
Gomphonema cf angustatum AM502005
Gomphonema productum AM501993
Gomphonema affine AM502033
Gomphonema parvulum AJ243062
Gomphonema affine AM502002
Gomphonema cf parvulum AM501995
Reimeria sinuata JN790290
Reimeria sinuata JN790291
Reimeria sinuata JN790292
Gomphonema micropus AM501964
Gomphonema micropus AM501965
Gomphonema rosenstockianum JN790287
Gomphonema rosenstockianum JN790288
Gomphonema rosenstockianum JN790289
Gomphonemataceae
Encyonema minutum AM501961
Encyonema silesacum JN790275
Encyonema silesacum JN790276
Encyonema silesacum JN790277
Encyonema triangulatum AJ535157
Cymbellaceae
Cocconeis stauroneiformis AB430614
outgroup Cocconeidaceae
0.03


2

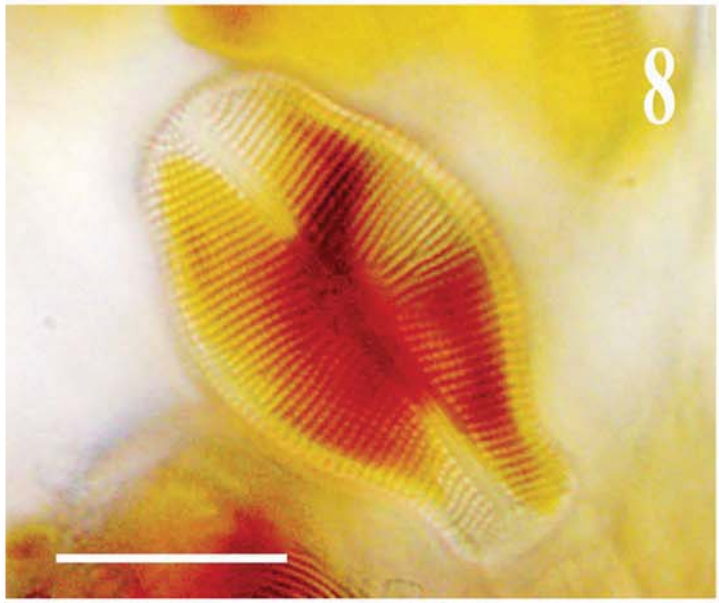

3

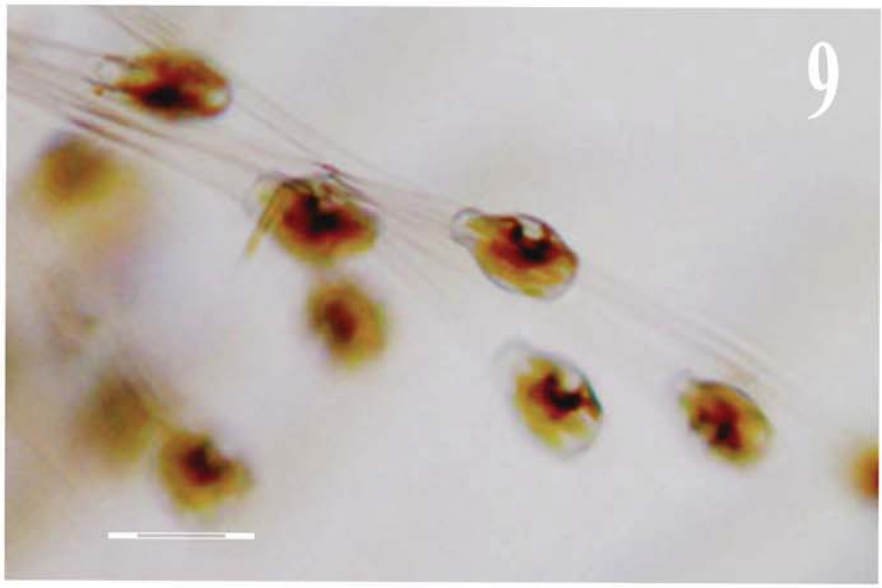

4

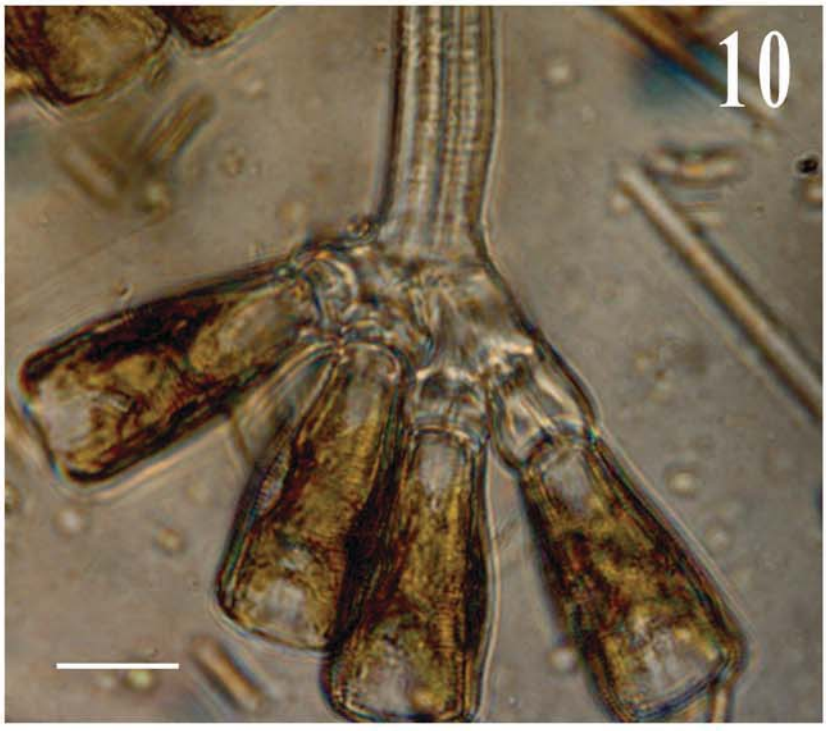

5

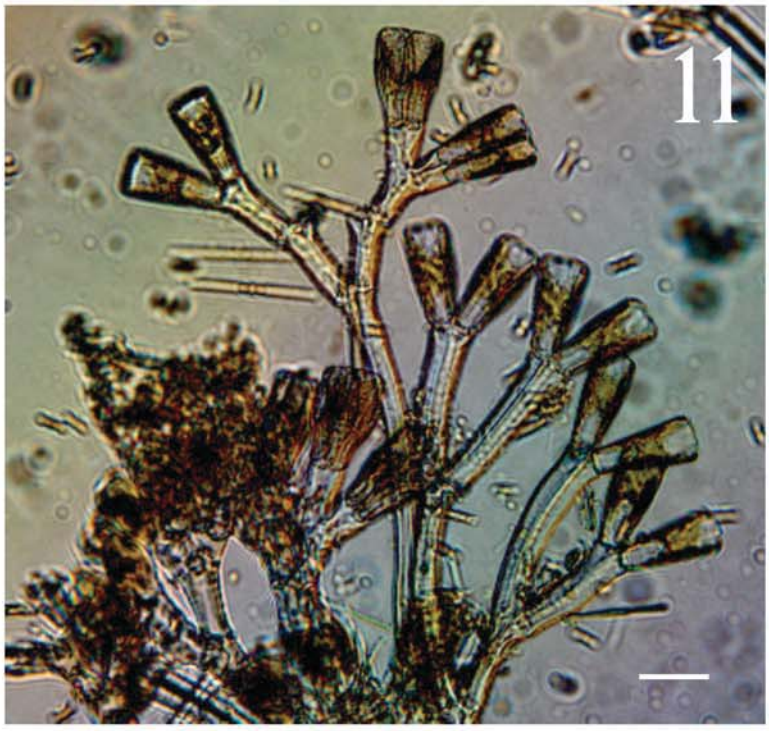

6

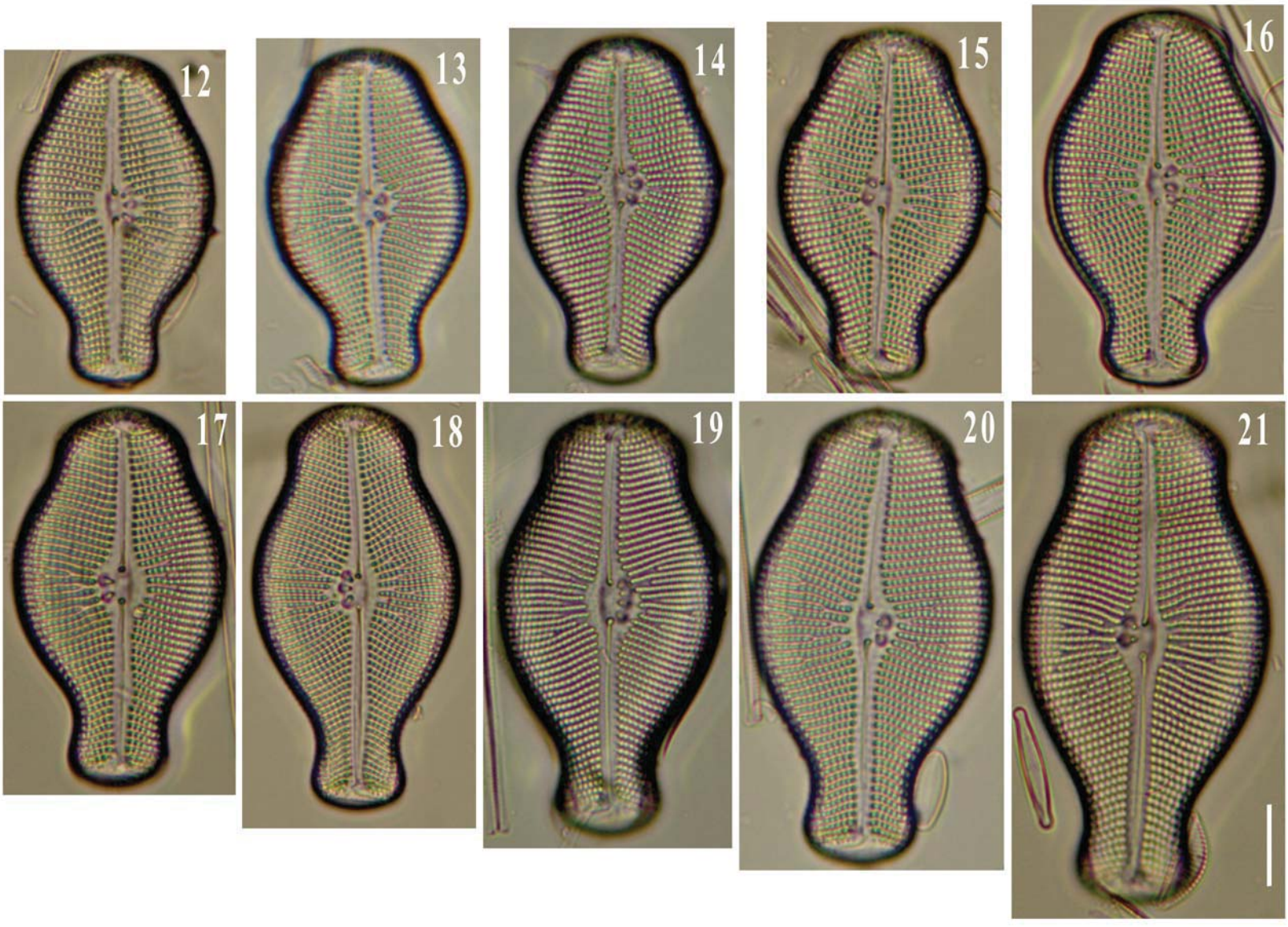

7

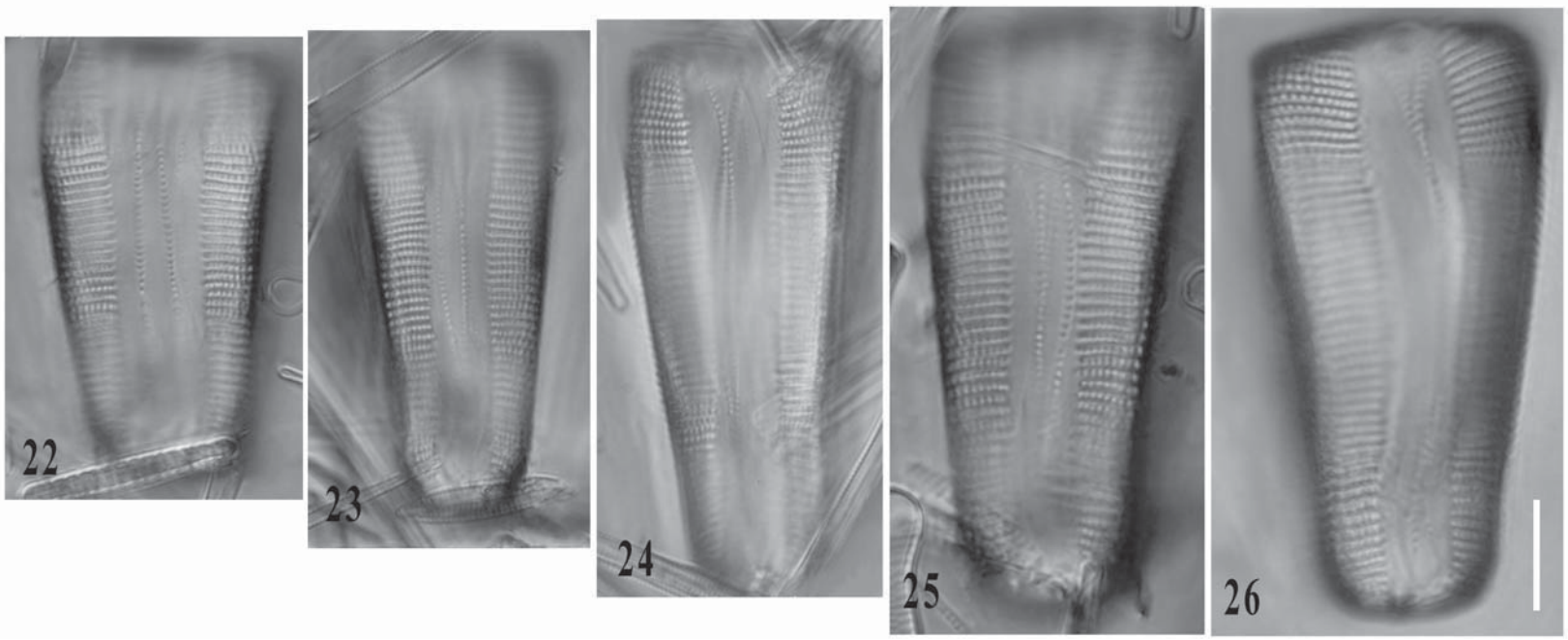

8

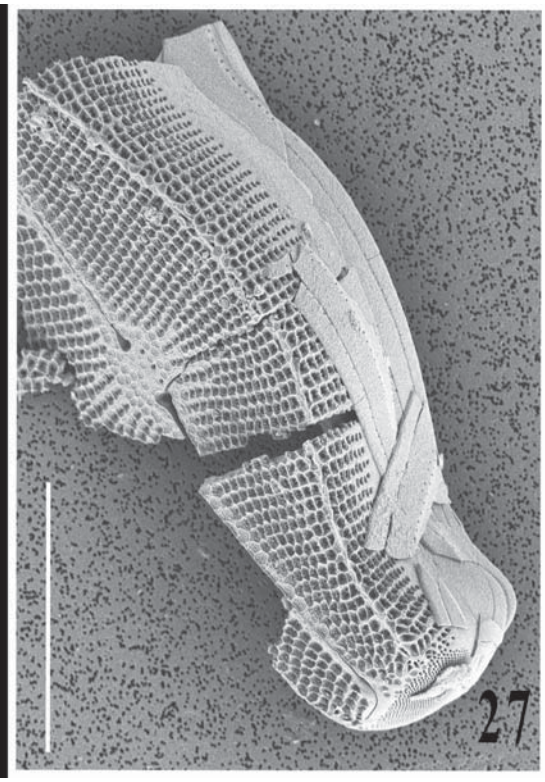

9

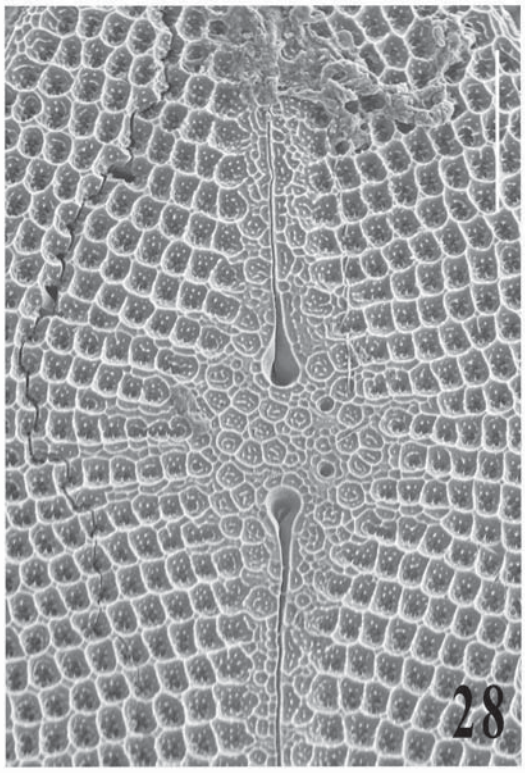

10

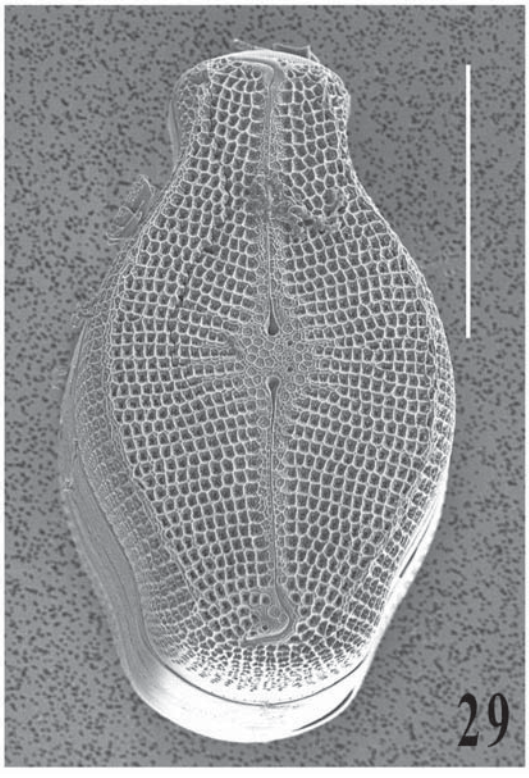

11


12
13
14
15
16
17
18
19
20
21

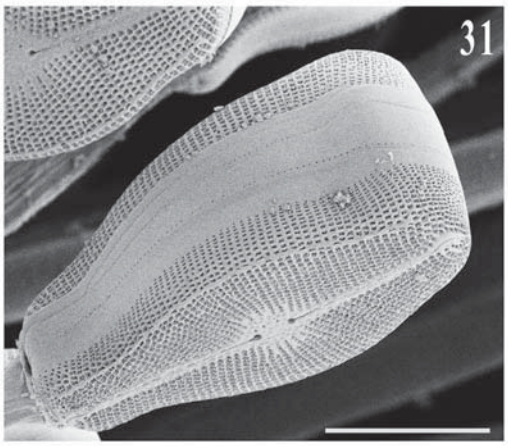

22
23
24
25
26

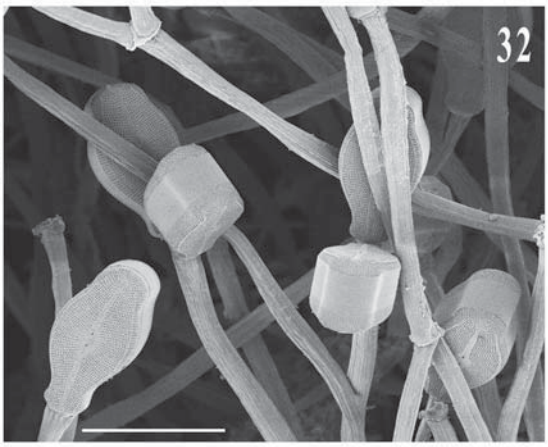

27

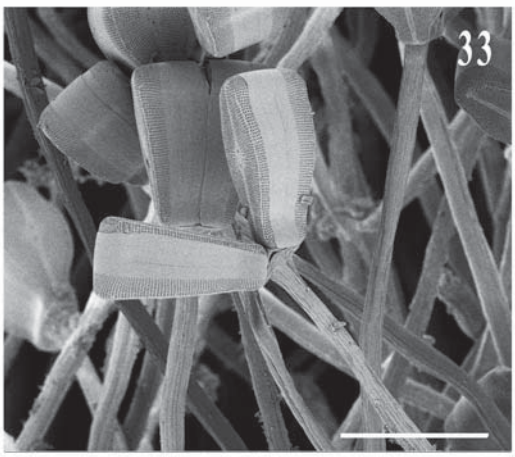

28

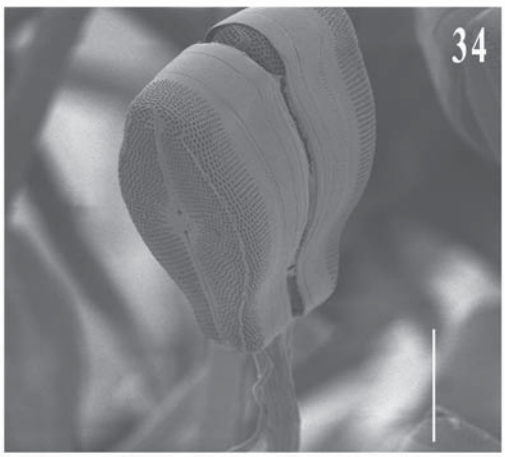

29

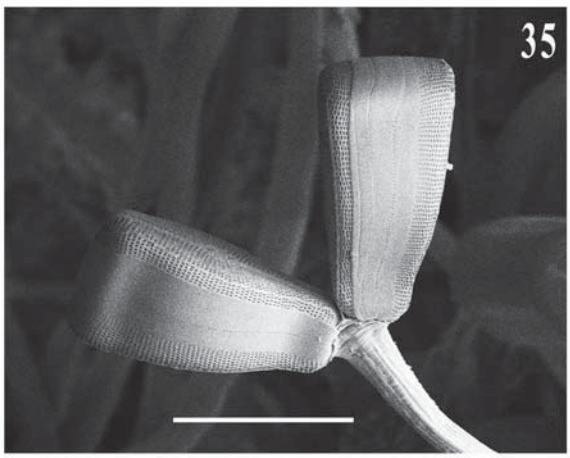

30

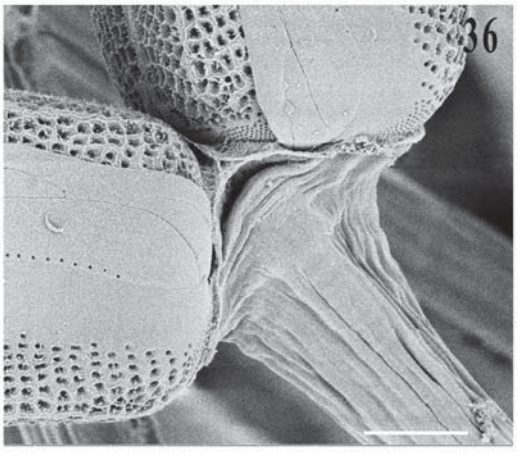

31


32

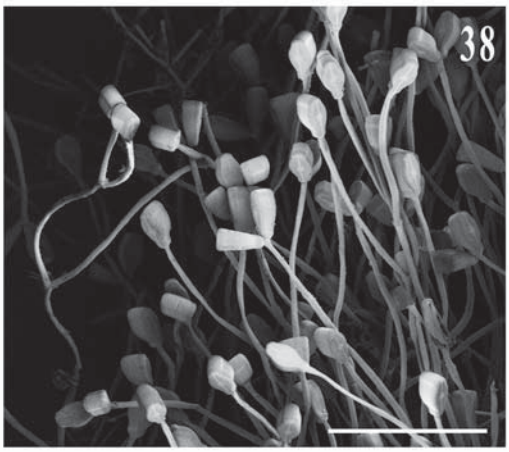

33

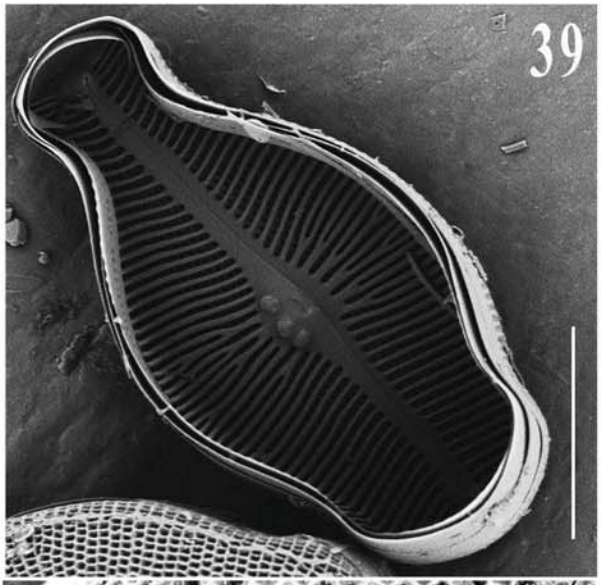

34

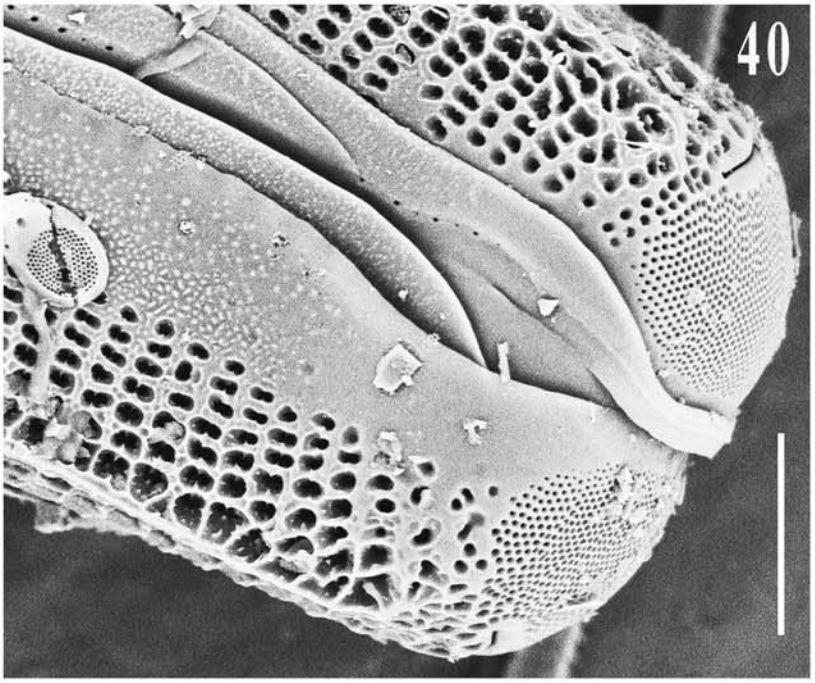

35

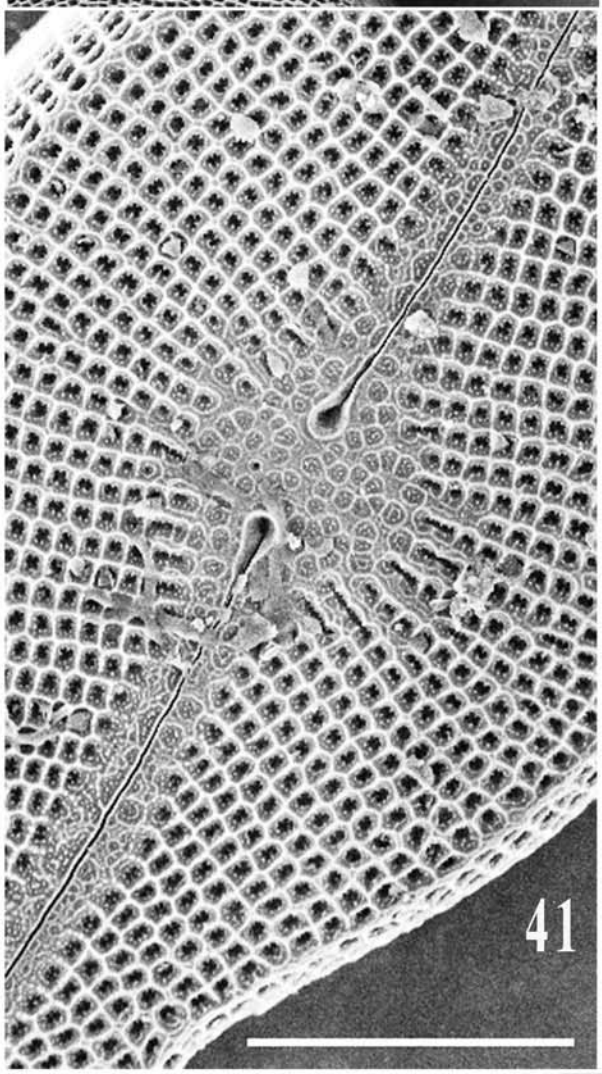

36

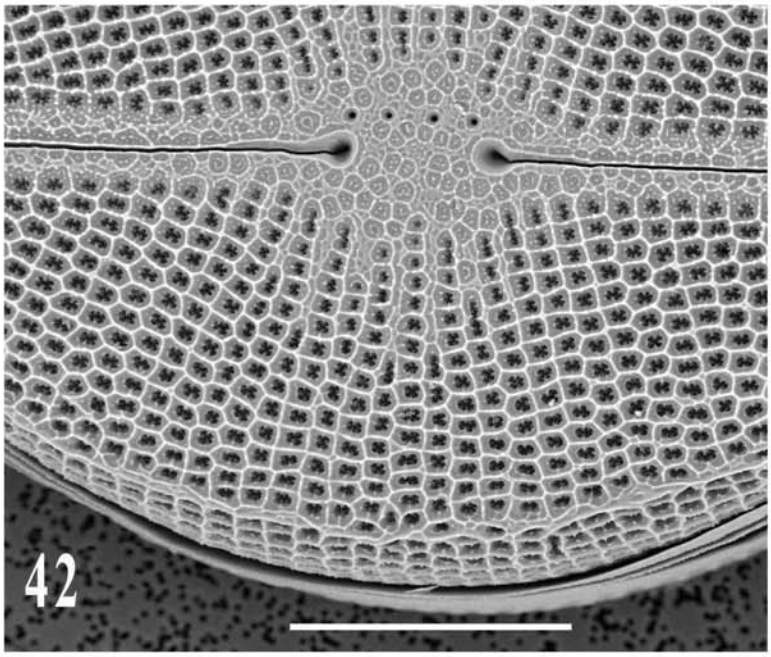

37

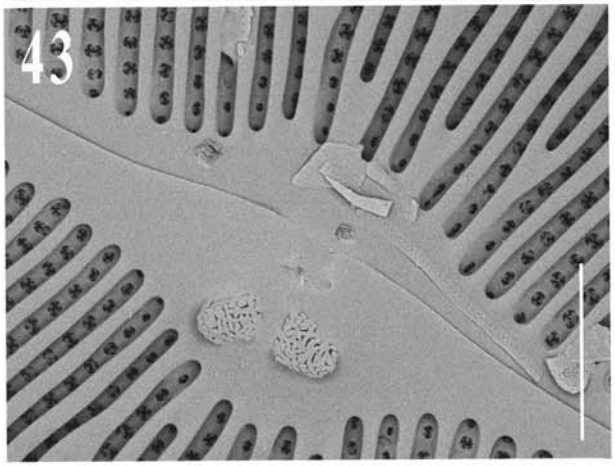

38


39


40


41


42


43

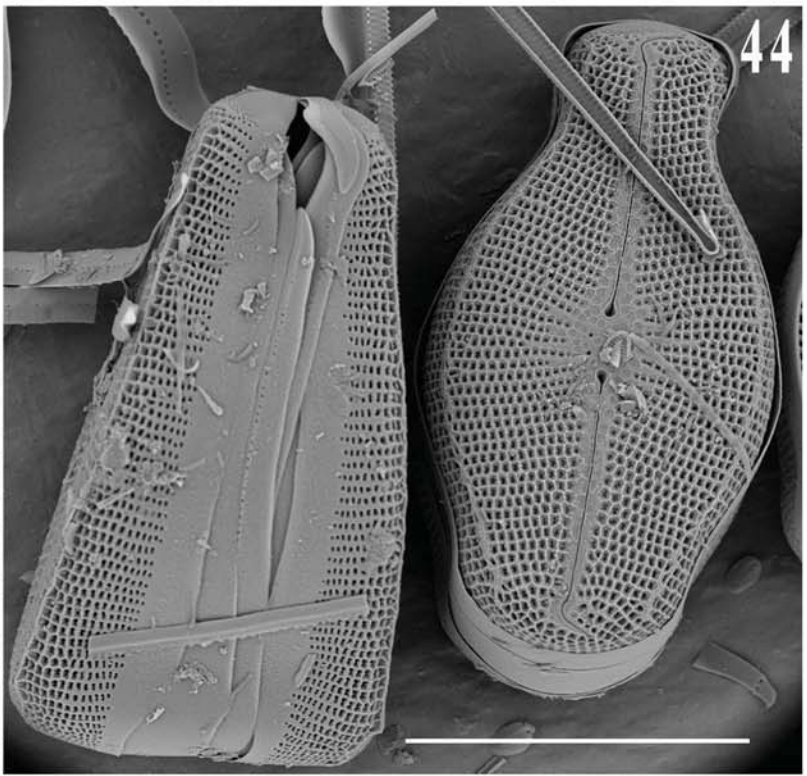

44

Table 1. Comparison of selected traits of *Didymosphenia* taxa and their localities. Morphology of areolae determined using images in Metzeltin & Lange-Bertalot (1995) and Mrozińska et al. (2006).

Species	Cell Length (μm)	Cell Width (μm)	Striae (in 10 μm)	Stigmata number	Areolae Morphology	Locality	References
D. curvata (Skvortsov & K.I. Meyer) Metzeltin & Lange-Bertalot	60–146	28–40	0–10	0–2	Lacking deep inclined walls, lacking spine-like projections	Russia - Siberia	Metzeltin & Lange-Bertalot (1995)
D. geminata (Lyngbye) M. Schmidt	48–132 87–137	25–45 36–41	8–10 7–9	1–6 2–5	Deep inclined walls surrounded by spine-like projections, dendritic slits below spines	Europe, Asia, and North America	Metzeltin & Lange-Bertalot (1995) Spaulding (2010)
D. geminata var. *dorogostaiskii* Skvortzow & K.I. Meyer	98–134	45–51	6	2–5		Russia - Siberia Lake Baikal	Skvortzow & Meyer (1928)

***D. hullii* Khan-Bureau**	**50–60**	**26.5–30.5**	**9–11**	**1–4**	Deep inclined walls surrounded by spine-like projections, dendritic slits below spines	U.S.A., Connecticut, West Branch Farmington River	This study
D. pumila Metzeltin & Lange-Bertalot	40–50	28–40	9–10	1–2	Lacking deep inclined walls, lacking spine-like projections	Russia - Siberia, reservoir in Irkutsk	Metzeltin & Lange-Bertalot (1995)
D. siberica (Grunow) M. Schmidt	70–85	34–40	10–12	1	Lacking deep inclined walls, lacking spine-like projections	Russia - Siberia, Jenissey River	Dawson (1973), Stoerm et al. (1986)
	55–85	28–40	10–13	1– 3			Metzeltin & Lange-Bertalot (1995)
D. tatrensis Mrozińska, Czerwik-Marcinkowska & Gradziński	60–99	28–41	13–14.5	1	Deep inclined walls, no spine-like projections, dendritic slits	Poland, Slovakia	Mrozińska et al. (2006)

Contrasting Morphological and DNA Barcoding Methods for Diatom (Bacillariophyta) Identification from Environmental Samples in the Eightmile River in Connecticut U.S.A.

ABSTRACT

The dominant and traditional approach used to identify diatom species is morphological characterization with light (LM) and scanning electron microscopy (SEM). However, using morphology alone to distinguish diatom species can be challenging because the phenotype of a species is often influenced by the life cycle stage and the environment. There is an increasing use of DNA barcoding for biodiversity studies and water quality monitoring, although the information provided by DNA barcoding of diatoms has not been compared comprehensively with that from morphology, except from cultured material. This study contrasted the performance of morphology and molecular data to distinguish diatom taxa from a single sample of the Eightmile River in Connecticut. Using a portion of the sample for morphological analysis with LM and SEM, the number of species, genera, and their taxonomic identities were evaluated. Three approaches for analysis of barcode data were compared. In total, the morphological approach yielded 59 taxa, and the molecular approaches yielded from 23 to 40 taxa. Some morphologically detected taxa were not detected by molecular means and some molecularly detected species were not detected morphologically. Using DNA barcoding and morphological methods simultaneously would provide more information on species diversity within an environmental sample.

INTRODUCTION

Diatoms are ubiquitous and ecologically important eukaryotic microalgae (Saunders et al. 1995, Medlin et al. 2000). Because many diatom species have been shown to be associated with particular environmental conditions, these taxa are accepted as biological indicators for assessing watercourses (Sgro & Johansen 1995, Saunders et al. 1995). Additionally, diatoms are used in climatological, paleontological, and biological diversity studies and are being tested for use in alternative energy, biomedical and nanotechnology fields. In order to address biodiversity assessment, water quality and other applications using diatoms, accurate taxonomic identification is essential. Estimates of the number of diatom species range from 10,000 to over 1 million (Van den Hoek et al. 1995, Mann & Droop 1996, Mann et al. 2010, Kermarrec et al. 2011). Many diatom species are poorly known while others have been studied extensively, particularly those species that are invasive, nuisance, or are toxic and form harmful algae blooms (HABs) (Stevenson & Pan 1999, Stoermer & Smol 1999). Diatoms and other algae are important contributors to energy flow in both aquatic and terrestrial ecosystems. Diatoms cycle nutrients in fresh and marine environments (Medlin et al. 1991, Sgro & Johansen 1995, Mann & Droop 1996) and have global ecological significance in the carbon and silicon cycles, making them important to all life on Earth (Sgro & Johansen 1995, Mann & Droop 1996, Van den Hoek et al. 1995, Stoermer & Smol 1999, Zimmermann et al. 2011). Additionally, diatoms, because their cell walls are made of silica, can fossilize, and are important recorders of past climates. With the growing need for alternative energy sources, algae such as diatoms have been investigated for their potential use as biofuels (Stoermer & Smol 1999). Thus, diatoms are one of the most ecologically and economically important eukaryotic microorganisms in the environment (Moniz &

Kaczmarska 2009).

Water quality degradation is a major concern as the human population continues to expand. There is concern because of increasing agricultural industries that stockpile manure and use pesticides, fertilizers, and herbicides. Nutrient enrichment is considered one of the most problematic issues affecting our rivers and streams, while other anthropogenic activities such as storm water runoff or non-point source pollution, erosion and sedimentation are of serious consequence, contributing to water quality degradation (USEPA 2000, Potapova & Charles 2007, USGS 2010). Information from diatoms can provide useful land-use assessments, and incorporate beneficial management practices and decision making when managing drinking water and fragile water ecosystems.

Employing biological indicators has been shown to be a reliable method for water quality assessments (Stevenson & Lowe 1986, Van den Hoek et al. 1995, Mann & Droop 1996, Stevenson & Pan 1999, Ector & Rimet 2005). Diatoms are good indicators because they have high reproductive rates, they are found in nearly every environment, are especially abundant where there is water and are found throughout the world, and particular species are tolerant of specific conditions and physical properties (Lowe 1974, Rott 1991, Sgro & Johansen 1995, Ector & Rimet 2005).

Presently, many states employ benthic macro-invertebrates (e.g., insect larvae, crustaceans, flat worms, mollusks, and annelids) in their water quality-monitoring program (Barbour et al. 1999, Wright et al. 2000, Blanco & Bécares 2010). Diatom data can complement information from macroinvertebrates because diatoms reproduce more quickly than benthic macro-invertebrates and respond more rapidly to changing environmental conditions. Diatoms respond to specific physical and chemical factors, such as high or low

nutrient conditions, causing visible growth (i.e., blooming) unlike that of benthic macroinvertebrates. The cell walls (frustules) of diatoms can persist indefinitely in water or in soil and can provide historical information (Sgro & Johansen 1995, Michels 1998, Stevenson & Sabater 2010). Algae, fish, and benthic macro-invertebrates have been used as compliance tools within the U.S.A. Clean Water Act (CWA) water quality criteria (Gold et al. 2002, Blanco & Bécares 2010, Stevenson & Sabater 2010).

Diatom taxa are identified morphologically using LM and SEM (Fig. 1). Light microscopy is useful because of its nearly universal access and because it can be used in determining the abundance of specific morphotypes. SEM provides a three-dimensional view with great detail and clarity, and is useful for distinguishing morphologically similar species. Lower magnification results in taxonomic inaccuracies that may change our understanding of the true biodiversity of rare or smaller diatoms, which are more difficult to distinguish (Medlin 1991, Medlin 1996, Morales et al. 2001, Mann et al. 2010, Zimmermann at al. 2011). Accurate identification is necessary for accurate assessments, but can be problematic when using morphology alone for taxonomy. Morphological identification frequently requires trained taxonomists and often, genetic divergence is not recognized morphologically because of phenotypic plasticity or cryptic species. Since morphological alterations in microalgae are associated with changes in physiological and/or ecological parameters, incorporating molecular analysis can provide useful data to help distinguish similar species (Bartual et al. 2008, Zimmermann et al. 2011, 2014b).

The use of DNA barcoding analysis in conjunction with LM and SEM can provide useful and new information about the diversity and the identity of diatoms found in a river sample since environmental factors contribute to considerable morphological plasticity

(Trainor et al. 1971, 1976, Morales et al. 2001). DNA-based methods combined with SEM have led to revisions of the LM morphology-based taxonomy, including descriptions of new genera and species that are morphologically cryptic (Morales et al. 2001, Mann et al. 2010). DNA based methods also have been used to examine the accuracy of diatom identification using morphology alone (Kermarrec et al. 2011, Zimmermann at al. 2011, 2014a). Many researchers think that DNA barcoding would be a valuable tool for water quality studies, to provide a consistent identification of diatoms and make the data from different studies directly comparable, even if taxonomy changes (Evans et al. 2007, Jahn et al. 2007, Kermarrec et al. 2011, 2014, Zimmermann 2014b).

OBJECTIVES

My goal was to address the following questions:

1) How does DNA barcoding compare with morphology for biodiversity assessment from environmental samples?

2) Will DNA barcoding of environmental samples help to reveal diatom taxa not seen using morphological approaches?

3) Will morphological approaches reveal diatoms that are not detected using molecular approaches?

MATERIAL AND METHODS

Study Location

The study took place in the Eightmile River 150 m downstream of the convergence on the Main Branch of the Eightmile River and the East Branch of the Eightmile River (41.43N 72.34W), a tributary of the Connecticut River located in Lyme, Connecticut (Fig. 2). The width of the river was approximately 105 m in a sunny area with brush and deciduous trees along the riparian zone. The Eightmile River Watershed is 160.5 km^2 drainage area for East

Haddam, Lyme, and Salem and smaller areas in East Lyme and Colchester. It drains into Hamburg Cove in Lyme, which then flows into the Connecticut River.

Environmental sampling

A location below the confluence was chosen where the substrate was no more than 33 cm deep. The sampling location has continuous flow with a stable cobble substrate since epilithic diatoms are found in this habitat and are the desired specimen for the study organism. Random sets of five-six small stones of 3-4 cm in diameter were collected and placed into Whirl-Pak® plastic bags. The cobbles were scrubbed with a clean toothbrush and rinsed with dH_2O into a 1L container and taken back to the lab for analysis. Each sample was centrifuged, the supernatant poured off and then split for DNA barcoding and morphology analysis. The morphological sample was stored at 4°C until further processing. The DNA barcoding sample was processed immediately or placed in the -20°C until processing.

Diatom preparation for microscopy

Diatom samples were simmered on a hot plate in a 1:1 ratio of water and 68% nitric acid to oxidize organic matter, after which the samples were removed from the hotplate to cool. Deionized water was used to rinse the samples of the acid, and then the samples were centrifuged to concentrate the diatom frustules at 600 g to avoid frustule damage. The process of rinsing included the addition of deionized water, centrifuging and the removal of supernatant 4–5 times or until the pH was neutral.

Light microscopy and SEM

Prior to acid washing, samples were placed on a microscope slide with a coverslip overlain and then viewed at x200 and x400 magnifications using a BX 60 Olympus microscope. The diatom sample slurry was air dried onto microscope coverslips, then used to make permanent

slides with the mounting medium NAPHRAX®. The diatom frustules were examined at ×1000 magnifications with a BX 60 Olympus microscope. Images were captured using an Olympus DP 25 color digital camera (2560 × 1920 pixels) with Olympus cellSens software. The diatoms on these slides were identified based on their morphological characteristics according to Krammer and Lange-Bertalot (1988), Round et al. (1990), and three online databases, the ANSP Algae Image Database (http://diatom.ansp.org/algae_image/), Diatoms of the United States (http://westerndiatoms.colorado.edu/), and the Great Lakes Image Database: (http://www.umich.edu/~phytolab/GreatLakesDiatomHomePage/top.html).

Aliquots of each cleaned sample were dried onto aluminum foil. The aluminum foil was adhered to SEM stubs with double-sided tape. Diatom samples were prepared following the methodology of Morales et al. (2001) the stubs were coated for 1 min at 1.8 kV with gold/palladium using a Polaron sputter coater. The stubs were viewed with a field emission FEI Nova Nano 450 scanning electron microscope located at the University of Connecticut Electron Microscopy lab. Image plates were created using Adobe® Creative Suite® 6 Photoshop.

LM species accumulation curves

Species accumulation curves (SAC) are frequently used to determine species richness, compare the similarity of species and evaluate sampling effort (Colwell & Coddington 1994). For this study I graphed the cumulative number of species with the number of individuals to compare the sampling effort and diatom diversity of LM using SAC until a plateau of species was reached. Diatom frustules were counted and identified on each of the three replicate slides. The statistical computation program, EstimateS (Colwell et al. 2014) was used as the analytical tool to graph the SAC.

DNA extraction, PCR, and cloning of diatoms

Samples were centrifuged, rinsed with deionized water, and then split into three replicate 50 μL microtubes for DNA extractions. Three replicates were compared to assess if a single replicate would represent the sample and assess possible heterogeneity across extractions. DNA extraction was accomplished using the MoBio PowerLyser™ Soil Extraction kit. Each extraction had 2 PCR replicates. PCR amplification of the V4 region of the 18S rDNA gene was achieved using diatom-specific primers D512 F and D978 R (Zimmermann et al. 2011). These primers amplify a region of the rDNA that is variable among eukaryotes and provides an appropriate level of signal to differentiate among diverse diatom taxa except for cryptic or very closely related species (Zimmermann et al. 2011). Two separate PCR reactions were performed, and then pooled into a single PCR product to minimize PCR biases in early cycles as shown by the schematic flow chart (Fig. 3). Any remaining product was stored at -80°C to archive the sample for future use. The PCR temperature profile included an initial denaturation step at 94°C (2 min), then five cycles consisting of denaturation at 94°C (45 s), annealing at 52°C (45 s) and elongation at 72°C (1 min), followed by 35 cycles in which the annealing temperature was lowered to 50°C, and a final elongation at 72°C (10 min). Resulting PCR products were visualized on a Syber Safe stained agarose gel, then quantified with a Nano Drop spectrophotometer. Cloning of PCR products was performed in two ways, both using Invitrogen TOPO® TA Cloning® Kit with an average cloning efficiency of 95%. The first method used a standard plasmid prep following Qiagen mini prep procedures using the QIAprep® Spin Miniprep Kit. The second method directly sequenced the colonies produced from cloning with the GoTaq®Green Master Mix, which was somewhat less labor intensive than the first method.

Sequencing

Cleaned PCR fragments were directly sequenced using the amplification primers of Zimmermann et al. (2011). The sequencing cycle comprised 27 cycles of denaturing at 96°C for 30 s, annealing at 50°C for 15 s, and extension at 60°C for 4 min, using the Big Dye™ Terminator Cycle Sequencing Reaction Kit (Applied Biosystems, Foster City, CA, USA). Products of cycle sequencing were cleaned using ethanol precipitation and analyzed on ABI 3100 DNA Sequencer™ (Applied Biosystems, Foster City, CA, USA). Contigs of individual reads were assembled in Geneious© (Geneious 2013), to produce consensus sequences. These sequences were trimmed to correspond to the partial 18S V4 sequences.

Species accumulation curves for cloned sequences

Using the BLAST algorithm to identify clone sequences to species, I graphed the number of sampled individuals against the cumulative number of species to compare the sampling effort and diatom diversity. The 3 replicates of this method were evaluated using EstimateS (Colwell et al. 2014), until a plateau of the species was reached indicating that the maximum amount of unique species in the sample was accomplished.

Phylogenetic tree building

I compared two different tree-building methods, maximum likelihood and Bayesian analyses, to determine phylogenetic relationships among the sequences. The optimal model of evolution selected by jModelTest (Darriba et al. 2012) was the General Time Reversible (GTR+I+G). I constructed a maximum likelihood (ML) tree with 1000 bootstrap replicates in GARLI (Genetic Algorithm for Rapid Likelihood Inference; Zwickl 2006). Bayesian analyses were conducted in MrBayes (Huelsenbeck & Ronquist 2001, Ronquist & Huelsenbeck 2003) and were run for 10 million generations with one cold chain and three

heated chains, and trees sampled every 1000 generations. The first 10% of the trees were discarded as burn-in.

p-Distances analysis

Uncorrected *p*-distances were calculated using MEGA version 6 (Tamura et al. 2013). In previous research, a threshold of $p = 0.02$ was shown to separate diatom species (Zimmermann et al. 2011, Luddington et al. 2012). Analyses for computing pairwise distances, within group mean distances, and among group mean distances included the diatom sequences from this study and public sequences downloaded from the NCBI website.

Species accumulation curves for p-Distances

I graphed the cumulative number of species with the number of individuals to compare the sampling effort and diatom diversity using the BLAST algorithm for the *p*-distances analyses. The 3 replicates of this method were evaluated using EstimateS (Colwell et al. 2014), until a plateau of the species was reached indicating that the maximum amount of unique species in the sample was accomplished.

GMYC analysis

BEAST version 1.7.5 (Bayesian Evolutionary Analysis Sampling Trees; Drummond et al. 2012) was run under the HKY+I+G model for 10 million generations. BEAST analyses were run with a strict clock (node height will be in units of mutations per site thus there is rate consistency across lineages of the phylogeny) and coalescent with constant size tree priors (uniform rates performs better on larger trees) resulted in an ultrametric tree (Drummond 2012). All other priors were left at the default values (base frequencies=uniform, HKY kappa= lognormal mean 1, Gamma - 4 rate categories exponential alpha with a mean of 0.5, Invariant sites = uniform). Ten percent of the trees were discarded as burn-in. The BEAST

trees were imported into R, a statistical software package downloaded from ProjectR (cran.r-project.org). The best tree was used to assess the threshold between interspecific and intraspecific diversification using the generalized mixed Yule-coalescent (GMYC) model of Pons et al (2006) implemented in the R package SPLITS (Fontaneto et al. 2007, Monaghan et al. 2009). The GMYC model uses phylogenetic trees to distinguish taxon boundaries from branching points in the tree and estimates species divergence (Reid & Carstens 2012, Zimmermann 2014). Using the GMYC approach with the BEAST tree I estimated the number of distinct clusters or species. GMYC was run with an interval of 0,10 (the default). Intervals are the upper and lower limits of the scaling parameters. Gamma distribution has two parameters; one is the shape parameter and the other a rate parameter. The scale parameter is the inverse of the rate parameter. An interval range of 0-1000 was tested to determine its impact on the estimated number of species.

RESULTS

Morphology

A range of 270-314 diatom frustules were counted on each of the 3 replicate microscope slides. SEM was used to confirm diatom identity of the LM images and was able to identify nine species that were not seen using LM. Together these approaches resulted in 59 taxa, of which 56 could be identified to species and three to the genus level (Table 1). Using LM and SEM I was able to identify taxa, illustrate diatom diversity and use the images to make comparisons of taxa found with molecular means (Figs 4 & 5).

Using EstimateS, a plateau of the species accumulation curve was reached and indicated that the amount of unique species in the sample was accomplished (Fig. 6). The SACs obtained from morphological data indicated that there were enough data to demonstrate diatom

diversity and sampling effort. Each replicate was similar demonstrating that replicates were representative of the sample.

Sequencing and initial identification

Cloned PCR fragments of 390-410 bp in length were sequenced. These were compared to published sequences from GenBank, using the BLASTn tool, to obtain information on the closest matches and to compare taxa that were detected by LM (see Table 1). A total of 36 published sequences were retained for phylogenetic analysis with the new data. Three hundred fifty seven clones where processed, of which 167 were identified as containing diatom specific sequences by the NCBI database BLASTn tool. Some of the cloned sequences were closely related to known species while others were not. I used two different thresholds with the BLAST algorithm. The cloned sequences matched 23 distinct taxa in the 100-99% similarity range using the BLAST tool, and 40 different species in the 98% similarity range. Any non-diatom or non-specific sequences were eliminated from consideration. Many of the BLASTn hits to the cloned sequences were listed as stramenopile or undetermined eukaryote sequences and were not specific to diatoms therefore they were not used.

Species accumulation curves of cloned sequences

Using EstimateS, a plateau of the species accumulation curve was reached indicating that the amount of unique species in the sample was accomplished (Fig. 8). Each replicate was similar with some exception, indicating that the sampling effort and species diversity was adequately represented for each replicate.

Phylogenetic trees

The resulting Bayesian and GARLI phylogenetic trees were similar in groupings and their support, therefore the Bayesian tree was presented as a summary (Fig. S1). A simplified version of this Bayesian tree (Fig. 7) shows several large distinct clades and the phylogenetic distribution of the sequences from the environmental sample related to known taxa. The three replicates are color-coded in the tree; their fairly even distribution among the well sampled clades indicates that the DNA replicates produced similar results. The more rarely sampled taxa were represented by just one or two replicates. Several of the cloned sequences that were originally identified to a given taxon using BLAST were found to be more closely related to a different species in the phylogenetic tree.

p-distances analysis

Uncorrected pairwise *p*-distances were computed and a threshold *p*-distance was set at 0.02 divergence for estimation of within versus among species variation. Overall mean distance was $p = 0.086$. Distances among taxa ranged between $p=0.000$ and 0.297 indicating high intrageneric variation. The average within group evolutionary divergence of *Cymbella* taxa was $p=0.015$ for clones and $p=0.163$ for the published sequences. Estimates of evolutionary divergence between groups mean was $p=0.194$ for published sequences of *Cymbella* taxa. The estimated average evolutionary divergence for the within mean group distance was $p=0.032$ for clone sequences identified as *Gomphonema* taxa and $p=0.067$ for the published sequences of *Gomphonema* taxa. Estimates of evolutionary divergence between groups, the mean was $p=0.050$ for published sequences of *Gomphonema* taxa. This indicates that the genetic variation between these groups is high. Comparing clones of *Eunotia* to published sequences of *Eunotia* taxa, the estimated average divergence within groups was $p=0.050$ for

published *Eunotia* and p=0.025 for the *Eunotia* clones. The divergence among groups is 0.001 for the *Eunotia* clones found in this study. The *p*-distances are lower indicating less genetic diversity for this group. The uncorrected *p*-distances for cloned sequences are shown in Table S2. Any number greater than 2 % were interpreted as different species; in this study 23 species were detected.

p-distances species accumulation curves

Using EstimateS, a plateau of the species accumulation curve was reached indicating that the amount of unique species in the sample was accomplished (Fig. 9). Each replicate was similar demonstrating that the sampling effort and species diversity was adequately represented for each replicate.

GMYC analysis

The GMYC method, a discovery tool, relies on the prediction that independent evolution leads to the presence of diverse genetic clusters (Pons et al. 2006). The GMYC method identified 92 distinct taxa, having the highest diversity of all the methods used (Fig. S3). The number of estimated species was the same across the different tested scaling parameters.

DISCUSSION

In May 2008, the Eightmile River received Wild and Scenic Federal recognition after 10 years of study (EIGHTMILERIVER.ORG). The Eightmile River and its watershed are known for its pristine water quality and biodiversity of plant and animal life. The Eightmile River is an ideal study site because of its history of assessment using benthic macro-invertebrates as bioindicators.

The United States Geological Survey (USGS) has real-time monitoring stations along the river from which data can be obtained for physical properties (i.e. gage height,

precipitation, and discharge rate). Water flow, temperature and depth information can be obtained at the USGS website http://waterdata.usgs.gov/ct/nwis/. The Connecticut Department of Energy and Environmental Protection Bureau of Water Management provided several years of water chemistry measurements for the Eightmile River (available upon request). Using these databases, comparisons can be made of the diatom species collected and the physicochemical conditions. The long-term utility of these methods can help future researchers to establish a protocol to compare the relationship of the species of diatoms found in a river using DNA barcoding and morphology with diatom autecology and correlate diatom species absence/presence with physical and chemical variables (Potapova & Charles 2007).

In this study I compared different methods for detecting diatom species. I used DNA barcoding methods to assess a broad taxonomic spectrum of diatoms from environmental samples and contrasted results obtained using morphology. DNA barcoding methods were comparable to morphological methods when abundant species were the focus. The rare or smaller species could be underestimated when using PCR-based methods. However as Kermarrec et al. (2014) noted rare species, although ecologically important, are not used in biomonitoring assessments. Morphology based methods can also be somewhat problematic because of taxonomic uncertainties. Both molecular and morphological methods used have some ambiguities, which may lead to incorrect analyses. Having both or more methods provides more information pertaining to diatom diversity, absence, presence and abundance.

Such comparisons have been made previously. For example, Zimmermann et al. (2011) and Luddington et al. (2012) constructed artificial communities of diatoms using cultures of known identities. They found that the V4 region of the18S was useful to

distinguish diatom taxa. However the 18S has been considered insufficient for recognition of closely related evolutionarily young taxa for a barcoding marker, although the V4 region can be a good genetic marker for environmental sampling for water quality assessments because of its ease and ability to distinguish many taxa (Zimmermann et al. 2011, Luddington et al. 2012). Zimmermann et al. (2011) suggested using two or more diatom specific markers to help identify those species that are evolutionarily young or closely related. Here I used diatom specific primers from the V4 region of the 18S. The use of rbcL diatom specific primers that have been successful in recent studies (Nakov et al. 2014) should be considered.

Both morphological and DNA-based approaches have their place in diatom diversity analyses. For instance, morphological methods estimated more taxa than did molecular methods when using the NCBI BLASTn. In some cases the NCBI published sequences were identified only to the division level in GenBank, which led to several unidentifiable taxa for this study and many sequences that were identified by BLASTn were identified to stramenopile rather than to a diatom taxon, which was uninformative and those sequences not used. The BLAST tool is a first step towards identification that should be used with other methods. BLASTn can provide initial information that is valuable. Cloning and PCR artifacts may be problematic, however in this study I used PCR replicates and pooled my products to minimize PCR biases. Cloning and sequencing provided useful data for those sequences that were diatom specific, demonstrating diatom diversity even if only to the genus level. Nonetheless, morphology and taxonomy can be costly and for an inexperienced person an exorbitant amount of time to process. Cleaning the sample, preparing the slide and counting 300-500 frustules accurately is time consuming. LM and SEM training requires a university setting for experience in diatom taxonomy and equipment usage. All too often diatomists do

not agree with identifications because of the many taxonomic descriptions that are available and interpretation can be diverse leading to misidentification (Morales et al. 2001, Mann et al. 2010, Rimet et al. 2012, Zimmermann et al. 2014b). If universal diatom primers are utilized and the amount of published and valid sequences on the NCBI website improved, DNA barcoding will provide useful information (Jahn et al. 2007). DNA barcoding can distinguish those taxa that are difficult to discriminate during their life cycle stages, cryptic and have phenotypic plasticity characteristics when morphology may not (Jahn et al. 2007, Mann et al. 2010, Zimmermann et al. 2014b).

The GMYC method currently is used widely in biodiversity assessments and phylogenetic community ecology, especially where only DNA sequence data are available such as that in environmental sampling (Fujisawa & Barraclough 2013, Zimmerman 2014a). GMYC can be problematic for several reasons including, it only uses data from a single locus, trees must be fully resolved (polytomies are not allowed) and this method does not take tree uncertainty into account. Therefore nodes with low and high probability are treated the same, causing errors in species delimitation by overestimating the taxa diversity (Pons et al. 2006, Reis & Carstan 2012, Fujisawa & Barraclough 2013). The GMYC model may split clades and overestimate species. In my study the GMYC estimated 92 species suggesting twice as many species then other methods. The *p*-distance analyses indicated the diversity of 23 species without phylogenetic inference using only genetic divergence and the *p*-distance SACs were able to inform us that my work was sufficient with adequate diatom abundance. It appears advantageous to use these methods simultaneously to evaluate and help distinguish taxa identity, presence, or absence in an environmental sample. Further work can help to develop a rigorous molecular database that can provide useful information for those who

want to differentiate diatom taxa in an environmental sample without having to be an expert in taxonomy. Newer molecular methods to obtain and quantify diatoms in environmental samples, based on quantitative real-time polymerase chain reaction (qRT-PCR) or Next Generation Sequencing (NGS), are being tested and utilized (Kermarrec et al. 2014, Nakov et al. 2014, Zimmermann et al. 2014a,b). Those testing qRT-PCR and NGS have found that these methods are useful however PCR biases can lead to underestimation of rare taxa and as with the traditional molecular methods the challenge is the lack of reference genomes available for most taxa (Kermarrec et al. 2014, Zimmerman et al. 2014 a, b).

The NCBI BLASTn tool allows DNA sequencing analyses to be possible although to date there are limited diatom sequences available for accurate comparisons and many of the diatom sequences in the database may not be efficiently verified to provide reliable identification of diatom species. Furthermore, a link between the DNA sequence databases and a diatom image database should be considered. Linking images with the sequence may be an intensive undertaking but a powerful tool. Providing an image linked with a sequence can aid in the evaluation of what you see via the microscope and matching your sequence. If these molecular methods 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 biological indicators. It can also facilitate biological diversity studies as well as provide information on non-native, nuisance and invasive species. Long-term biodiversity studies should include both microscopy and molecular methods. Presently, molecular methods should not take the place of morphological methods but rather be used in conjunction with morphological approaches.

Rimet et al. (2012) tested and determined that using diatoms to genus level was

sufficient for pollution assessment however species level resolution was important and required for ecoregion classification studies. Insomuch as diatoms are important bioindicators and are often used for routine water quality and biodiversity assessments it is crucial to identify diatoms accurately (Morales et al. 2001, Pniewski et al. 2010, Mann et al. 2010), but it is also important that a cost effective and efficient method be developed (Jahn et al. 2007, Kaczmarska 2007). The cost of sequencing has dropped to levels that are more affordable than morphology based methods (Mann et al. 2010). Even so 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 in GenBank and the refinement of laboratory protocols for optimal extraction and amplification methods (Evans et al. 2007, Jahn et al. 2007, Kermarrec et al. 2014, Zimmermann et al. 2014b).

Establishing a reliable diatom DNA barcoding protocol with accurate sequence data may provide a more uniform identification process. The resulting nucleotide data could be compared with existing, publicly available data that are archived through the NCBI database in order to compare sequences. 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.

If diatoms are to be used as water quality and biodiversity assessment tools it is important to create a diatom specific database to be developed for future researchers. This powerful tool could be used by anyone wanting to understand diatom diversity. The development of a quality control regime to ensure that deposited sequences meet specified criteria with high quality data and images could be ideal. Having a separate search engine for taxonomic levels divided by ecoregion could help those interested, to obtain more accurate

and concise information. In conclusion many state agencies want to employ a more comprehensive and holistic approach for monitoring rivers, utilizing many organisms, which would include benthic macro invertebrates, macro algae, and diatoms. In addition, employing morphological and DNA barcoding methods would provide a wide-ranging view of the health of the river ecosystem.

ACKNOWLEDGEMENTS

We thank the Connecticut Institute of Water Resources (CTIWR) for funding this research. I thank K. Fučíková, S. Olm, C. Lo, M. Letsch, H. McManus and acknowledge the late F.R.Trainor for helpful discussion. I am grateful to J. Zimmermann for his thoughtful correspondence. This work used the computer cluster maintained by the Bioinformatics Facility (Biotechnology/Bioservices Center) at the University of Connecticut. I thank Three Rivers Community College's Dean of IT S. Goetchius, K. Barfield, S. Cohen, D. Jewett, and President G. S. Jones for providing continual support for this project, J. Sgro, John Carroll University, for providing verification of species, M. Cantino, S. Daniels, and X. Sun of the EM lab at UCONN, NSF 1126100 supporting the purchase of the Nova Nano SEM, R. Colwell for his EstimateS support, V. Kask scientific illustrator in Biological Central Services at UCONN, J. Morrison, East Hartford, USGS for their assistance.

REFERENCES

ANSP Algae Image Database (http://diatom.acnatsci.org/AlgaeImage/).

Barbour M.T., Gerritsen J., Snyder B.D. & Stribling J.B. 1999. Rapid Bioassessment Protocols for Use in Streams and Wadeable Rivers: Periphyton, Benthic Macroinvertebrates and Fish, 2nd Ed. EPA 841-B-99-002. U.S. Environmental Protection Agency; Office of Water; Washington, D.C.

Bartual A., Galvez J. A. & Ojeda F. 2008. Phenotypic response of the diatom *Phaeodactylum tricornutum* Bohlin to experimental changes in the inorganic carbon system Botanica Marina 51: 350–359.

Blanco S. & Bécares E. 2010. Are biotic indices sensitive to river toxicants? A comparison of metrics based on diatoms and macro-invertebrates. Chemosphere 79: 18–25.

Colwell R. K. & Coddington J. A. 1994. Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions Royal Society London B. 345: 101–118.

Colwell R. K., Chao A., Gotelli N. J., Lin S.-Y., Mao C. X., Chazdon R. L., and Longino J. T.. 2012. Models and estimators linking individual-based and sample-based rarefaction, extrapolation, and comparison of assemblages. Journal of Plant Ecology 5:3–21.

Connecticut River Watershed Council. 2011. http://www.ctriver.org/our_region_and_rivers/river_facts/index.html.

Darriba D., Taboada G. L., Doallo R. & Posada D. 2012. jModelTest 2: more models, new heuristics and parallel computing. Nature Methods 9: 772.

Ector L. & Rimet F. 2005. Using bioindicators to assess rivers in Europe: An overview. In: Lek S., Scardi M., Verdonschot P. F. M., Descy J. P. & Park Y. S. (eds). Modeling Community Structure in Freshwater Ecosystems. Berlin: Springer Verlag, Heidelberg. p. 7–19.

Evans K. M., Wortley A. H. & Mann D. G. 2007. An assessment of potential diatom "barcode" genes (cox1, rbcL, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). Protist 158: 349–364.

Fontaneto D., Herniou E. A., Boschetti C., Caprioli M., Melone G., Ricci C. & Barraclough T. G. 2007. Independently evolving species in asexual bdelloid rotifers. PLoS Biology 5:e87

Fontaneto D., Herniou E.A., Boschetti C., Caprioli M., Melone G., Ricci

C., Barraclough T.G. 2007. Independently evolving species in asexual bdelloid rotifers. PLoS Biol. 5:e87

Fujisawa T. & Barraclough T. G. 2013. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: A revised method and evaluation on simulated data sets. Systematic Biology 62: 707–724.

Gold C., Feurtet-Mazel A., Coste M., & Boudou A. 2002. Field transfer of periphytic diatom communities to assess short-term structural effects of metals (Cd, Zn) in rivers. Water Research 36: 3654–3664.

Jahn R., Zetzsche H., Reinhardt R. & Gemeinholzer B. 2007. Diatoms and DNA barcoding: a pilot study on an environmental sample. Proceedings of the 1st Central European Diatom Meeting. Botanic Garden and Botanical Museum Berlin-Dahlem, Freie Universität Berlin ISBN 978-3-921800-63-8 BGBM, Berlin, 63–68.

Kaczmarska I., Reid C. & Moniz M. 2007. Diatom taxonomy: morphology, molecules and barcodes. In: Kusber W. H. & Jahn R. (eds). Proceedings of the 1st Central-European Diatom Meeting 2007. Berlin-Dahlem, Germany, 69–72.

Kermarrec L., Ector L., Bouchez A., Rimet F. & Hoffmann L. 2011. A preliminary phylogenetic analysis of the Cymbellales based on 18S rDNA gene sequencing, Diatom Research 26: 305–315.

Kermarrec L., Franc A., Rimet F., Chaumeil P., Frigerio J. M., Humbert J. F. & Bouchez A. 2014. A next-generation sequencing approach to river biomonitoring using benthic diatoms. Freshwater Science 33: 349–363.

Krammer K. & Lange-Bertalot H. 1988. Bacillariophyceae 2. Teil, Epithemiaceae, Bacillariceae, Surirellaceae. In: Ettl H., Gerloff J., Heynig H. & Mollenhauer D. (eds). Süßwasserflora von Mitteleuropa. Stuttgart & New York 2 (2): 596 p.

Lowe R. L. 1974. Environmental requirements and pollution tolerance of freshwater diatoms. United States Environmental Protection Agency, Cincinnati, Ohio. EPA-670/474–005.

Luddington I. A., Kaczmarska I. & Lovejoy C. 2012. Distance and character-based evaluation of the V4 region of the 18S rRNA gene for the identification of diatoms (Bacillariophyceae). PLos ONE 9: e45664.

Mann D. G. & Droop S. J. M. 1996. Biodiversity, biogeography and conservation of diatoms. Hydrobiologia 336: 19–32.

Mann D. G., Sato S., Trobajo R., Vanormelingen P. & Souffreau C. 2010. DNA barcoding for species identification and discovery in diatoms. Cryptogamie, Algologie 31: 557–577.

Medlin L. K., Kooistra W. C. & Schmid A.-M. M. 2000. A review of the evolution of the diatoms: a total approach using molecules, morphology and geology. In: Witkowski A. & Sieminska J. (eds). The origin and Early Evolution of the Diatoms: Fossil, Molecular and Biogeographical Approaches. W. Szafer Institute of Botany, Polish Academy of Sciences, Cracow, Poland, 13–35.

Medlin L. K., Kooistra W. C., Gersonde R. & Wellbrock U. 1996. Evolution of the diatoms (Bacillariophyta). II. Nuclear-encoded small-subunit rRNA sequence comparisons confirm a paraphyletic origin for the centric diatoms. Molecular Biology and Evolution 13: 67–75.

Medlin L. K., Elwell H. J., Strickel S. & Sogin M. L. 1991. Morphological and genetic variation within the diatom *Skeletonema costatum* (Bacillariophyta): evidence for a new species, *Skeletonema pseudocostatum*. Journal of Phycology 27: 514–524.

Michels A. 1998. Effects of sewage water on diatoms (Bacillariophycae) and water quality in two tropical streams in Costa Rica. Revista de Biologia Tropical 46, Suppl. 6: 153–175.

Monaghan M. T., Wild R., Elliot M., Fujisawa T., Balke M., Inward D. J. G., Lees D. C., Ranaivosolo R., Eggleton P., Barraclough T. G. & Vogler A. P. 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. Systematic Biology 58: 298–311.

Moniz M. B. J. & Kaczmarska I. 2009. Barcoding diatoms: Is there a good marker? Molecular Ecology Resources 9: 65–74.

Morales E. A., Siver P. A. & Trainor F. R. 2001. Identification of diatoms (Bacillariophyceae) during ecological assessments: Comparison between light microscopy and scanning electron microscopy techniques. Proceedings of the Academy of Natural Sciences of Philadelphia 151: 95–103.

Nakov T., Ruck E. C., Galachyants Y., Spaulding S. A. & Theriot E. C. 2014. Molecular phylogeny of the Cymbellales (Bacillariophyceae, Heterokontophyta) with a comparison of models for accommodating rate variation across sites. Phycologia 53: 359–373.

Pons J, Barraclough T. G., Gomez-Zurita J., Cardoso A., Duran D. P., Hazell S., Kamoun S., Sumlin W. D. & Vogler A. P. 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. Systematic Biology 55: 595–609.

Potapova M., & Charles D. F. 2007. Diatom metrics for monitoring eutrophication in rivers of the United States. Ecological Indicators 7: 48–70.

Pniewski F. F., Friedl F. & Latała A. 2010. Identification of diatom isolates from the Gulf of Gdańsk: testing of species identifications using morphology, 18S rDNA sequencing and DNA barcodes of strains from the Culture Collection of Baltic Algae (CCBA). Institute of Oceanography, ISSN 1730-413X.

Reid N. M. & Carstans B. C. 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. BMC Evolutionary Biology 12:196.

Rott E. 1991. Methodological aspects and perspectives in the use of periphyton for monitoring and protecting rivers. In: Whitton B. A. & Rott E. (eds). Use of Algae in Monitoring Rivers II. Institut für Botanik, Universität Innsbruck. Innsbruck, Austria, 9–16.

Round F. E., Crawford R. M. & Mann D. G. 1990. The Diatoms - Biology and morphology of the genera. Cambridge University Press, Cambridge, 747 p.

Saunders G. W., Potter D., Paskind M. I. & Andersen R. A. 1995. Cladistic analyses of combined traditional and molecular data sets reveal an algal lineage. Proceedings of the National Academy of Sciences USA 92: 244–248.

Sgro G. V. & Johansen, J. R. 1995. Rapid bioassessment of algal periphyton in freshwater streams. In: Butterworth F. M. (ed.). Biomonitors and Biomarkers as Indicators of Environmental Change, Plenum Press, New York, 291–311.

Stevenson R. J. & Pan Y. 1999. Assessing ecological conditions in rivers and streams with diatoms. In: Stoermer E. F. & Smol J. P. (eds). The Diatoms: Applications to the Environmental and Earth Sciences. Cambridge University Press, Cambridge, UK, 11–40.

Stevenson R. J. & Lowe R. L. 1986. Sampling and interpretation of algal patterns for water quality assessment. In: Isom B. G. (ed.). Rationale for sampling and interpretation of ecological data in the assessment of freshwater ecosystems. ASTM STP 894. American Society for Testing and Materials, Philadelphia, Pennsylvania, USA, 118–149.

Stevenson R. J. & Sabater S. 2010. Global change and river ecosystems – Implications for structure, function and ecosystem services. Hydrobiologia 657: 1–2.

Stoermer E. F. & Smol J. P. 1999. The Diatoms. Applications for the Environmental and Earth Sciences. Cambridge University Press, Cambridge, UK, 488 p.

Tamura K., Stecher G., Peterson D., Filipski A. & Kumar S. 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30: 2725–2729.

Trainor F. R., Rowland H. L., Lylis J .C., Winter P. A. & Bonanami P. L.1971. Some examples of polymorphism in algae. Phycologia 10: 113–119.

Trainor F. R., Cain J. R. & Shubert L. E. 1976. Morphology and nutrition of the colonial green alga *Scenedesmus*: 80 years later. Botanical Reviews 42: 5–25.

USEPA (US ENVIRONMENTAL PROTECTION AGENCY). 2000. United States Environmental Protection Agency (USEPA). Nutrient criteria technical guidance manual. Rivers and streams. EPA-822-B-00-002. Office of Water, US Environmental Protection Agency, Washington, DC., USA.

United States Geological Survey (USGS) CT. 2008. http://ct.water.usgs.gov/nawqa/design.htm

United States Geological Survey (USGS) CT. 2008. Algal Attributes: An Autecological Classification of Algal Taxa Collected by the National Water-Quality Assessment Program.

United States Geological Survey (USGS) CT. 2010. Water Quality Information http://water.usgs.gov/owq/

Van den Hoek C., Mann D. G. & Johns H. M. 1995. Algae: An Introduction to Phycology. Cambridge University Press, Cambridge, UK.

Wright J. F., Sutcliffe D. W. & Furse M. T. 2000. Assessing the biological quality of fresh waters: RIVPACS and other techniques. Freshwater Biological Association, Ambleside, UK.

Zimmermann J., Jahn R. & Gemeinholzer B. 2011. Barcoding diatoms evaluation of the V4 sub region on the 18S rRNA gene, including new primers and protocols. Organisms, Diversity & Evolution 11: 173–192.

Zimmermann J., Glöckner G., Jahn R., Enke N. & Gemeinholzer B. 2014a. Metabarcoding vs. morphological identification to assess diatom diversity in environmental studies. Molecular Ecology Resources. doi: 10.1111/1755-0998.12336

Zimmermann J., Abarca N., Enk N., Skibbe O., Kusber W.-H., & Jahn R. 2014b. Taxonomic reference libraries for environmental barcoding: a best practice example from diatom research. PLoS ONE 9: e108793.

Zwickl D. J. 2006. Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, The University of Texas at Austin.

Figures and Legends.

Figure 1. A. LM micrograph of diatom diversity showing the morphology of frustules. Scale bar=10 μm

B. SEM 3 dimensional view of the ultrastructure of the most abundant diatom taxa found in the Eightmile River, *Achnanthidium minutissimum*. Scale bar= 2 μm

Figure 2. A.USGS map of the Eightmile River Sampling site. B. Confluence of the Eightmile River sampling site.

Figure 3. Schematic showing the strategy used for collection of samples, LM, SEM, extraction, PCR, cloning, and sequencing to identify taxa.

Figure 4. LM images of diatom diversity in the Eightmile River **A.** *Synedra ulna*, **B.** *Frustulia crassinervia*, **C.** *Eunotia* cf. *praerupta* , **D.** *Synedra rumpens*, **E.** *Brachysira microcephala*, **F.** *Eunotia pectinalis* var. *undulata*, **G.** *Cymbella tumida*, **H.** *Gomphonema* sp. *angustatum*, **I.** *Gomphonema truncatum*, **J.** *Discotella stelligera*, K. *Cocconeis placentula*, **L.** *Meridian circulare* var. *constrictum,* **M.** *Encyonema silesiacum*, **N.** *Tabellaria flocculosa*, **O.** *Gomphonema angustatum*, **P.** *Gomphonema acuminatum*, **Q.** *Karayevia oblongella*. **R.** *Karayevia oblongella*. Scale bar 10 μm.

Figure 5. SEM images of diatom diversity in the Eightmile River. **A.** *Gomphonema acuminatum*, **B**. *Frustulia crassinervia*, **C.** *Nitzschia hantzschiana*, **D.** *Planothidium apiculatum,* **E.** *Gomphonema* sp. *angustatum*, **F.** *Navicula notha*, **G.** *Encyonema silesiacum*, **H.** *Brachysira microcephala*, **I.** *Cavinula cocconeformis*, **J.** *Tabellaria flocculosa*, **K.** *Eunotia* cf. *praerupta*, **L**. *Meridian circulare* var. *constrictum*, **M.** *Karayevia oblongella*, **N.** *Achnanthidium minutissimum*, **O.** *Discotella stelligera*. Scale bar 10 μm.

Figure 6. Species Accumulation Curve (SAC) representing number of individuals and species seen with the LM for morphological identification. The SAC indicates sampling effort and species diversity for each replicate.

Figure 7. Bayesian collapsed phylogenetic tree based on the V4 region of the 18S rDNA of selected published and cloned sequences from this study (published sequences are denoted with the GenBank Accession numbers). Replicates are highlighted. A. The first section of the tree. B. The midsection of the tree. C. The base section of the tree.

Figure 8. Species Accumulation Curve (SAC) representing number of individuals and species sequences and compared to the NCBI BLASTn tool database. The SAC indicates sampling effort and species diversity for each replicate.

Figure 9. Species Accumulation Curve (SAC) representing the number of individuals and species sequences generated from *p*-distances. The SAC indicates sampling effort and species diversity for each replicate.

Table 1. Compilation of all diatom taxa detected using DNA sequencing and morphology. Sequences were assigned to a given taxon when represented on the NCBI site using the BLASTn tool with a 100-98% threshold.

Supplemental materials

Figure S1. Bayesian phylogenetic tree based on the V4 region of the 18S rDNA of clone sequences from this study and selected published sequences (published sequences are denoted with the GenBank Accession numbers). Bayesian posterior probabilities (BPP) values indicate node support. Scale bar =0.5 expected number of substitutions/site. Replicates are highlighted. Blue represents replicate 1, salmon represents replicate 2, and purple represents replicate 3.

Table S2. A partial MEGA *p*-distance table based on diatom V4 18S rDNA sequences at a 2% divergence.

Figure S3. BEAST generated ultrametric tree used as the input tree for the Generalized Mixed Yule Coalescent (GMYC) Method. GMYC was used to identify taxon boundaries and branching points in the tree to represent divergence between taxa.

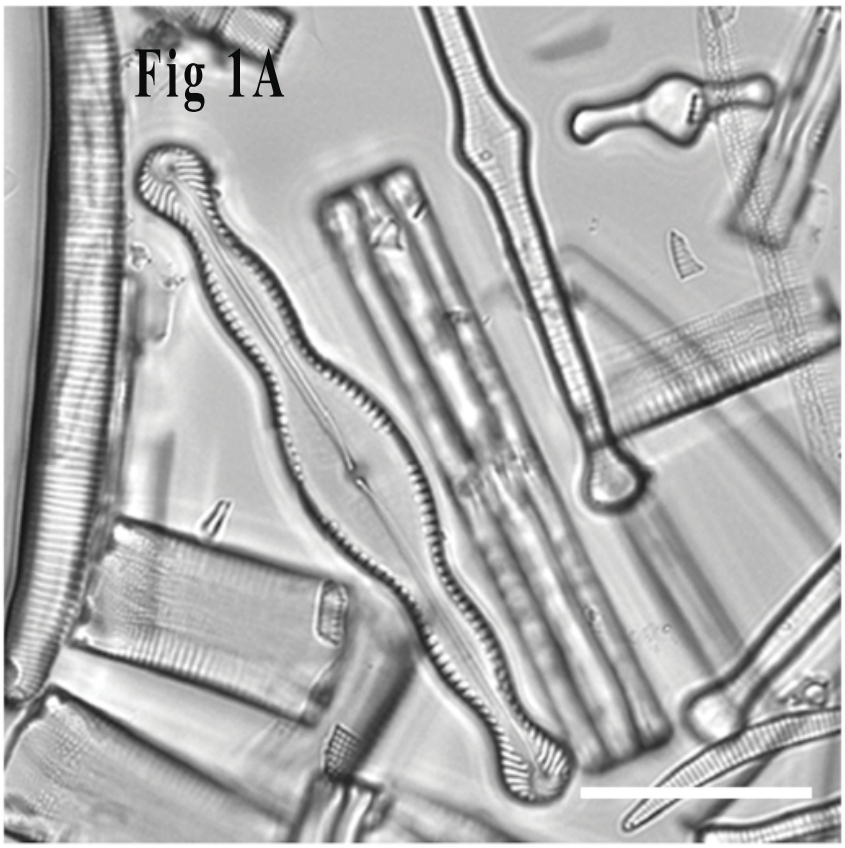
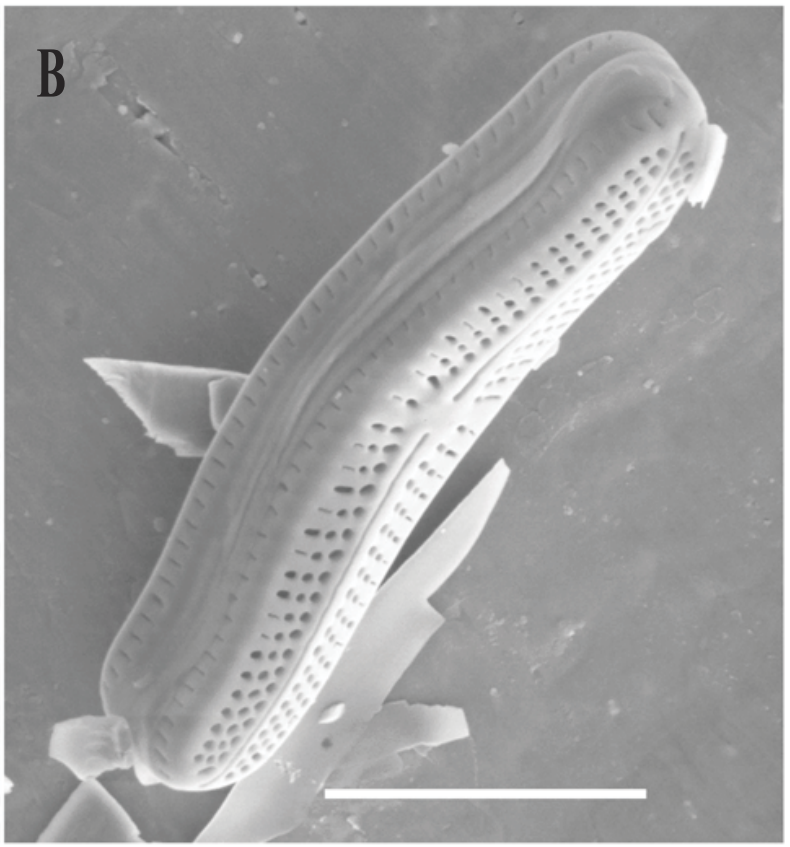


Fig 1A


B

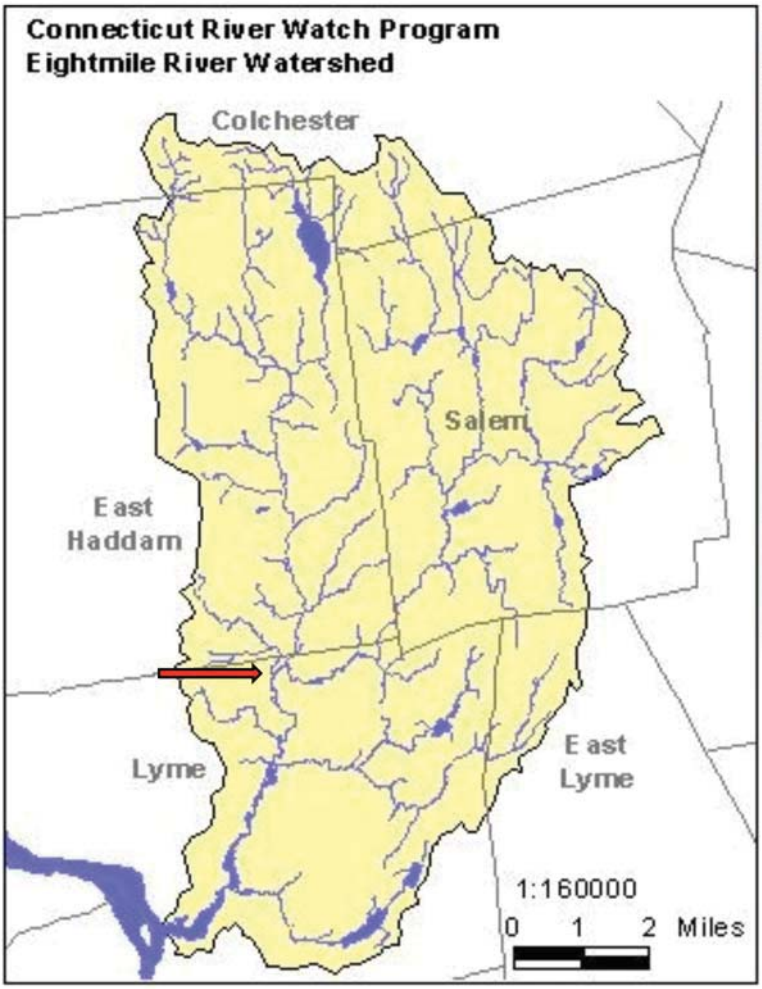
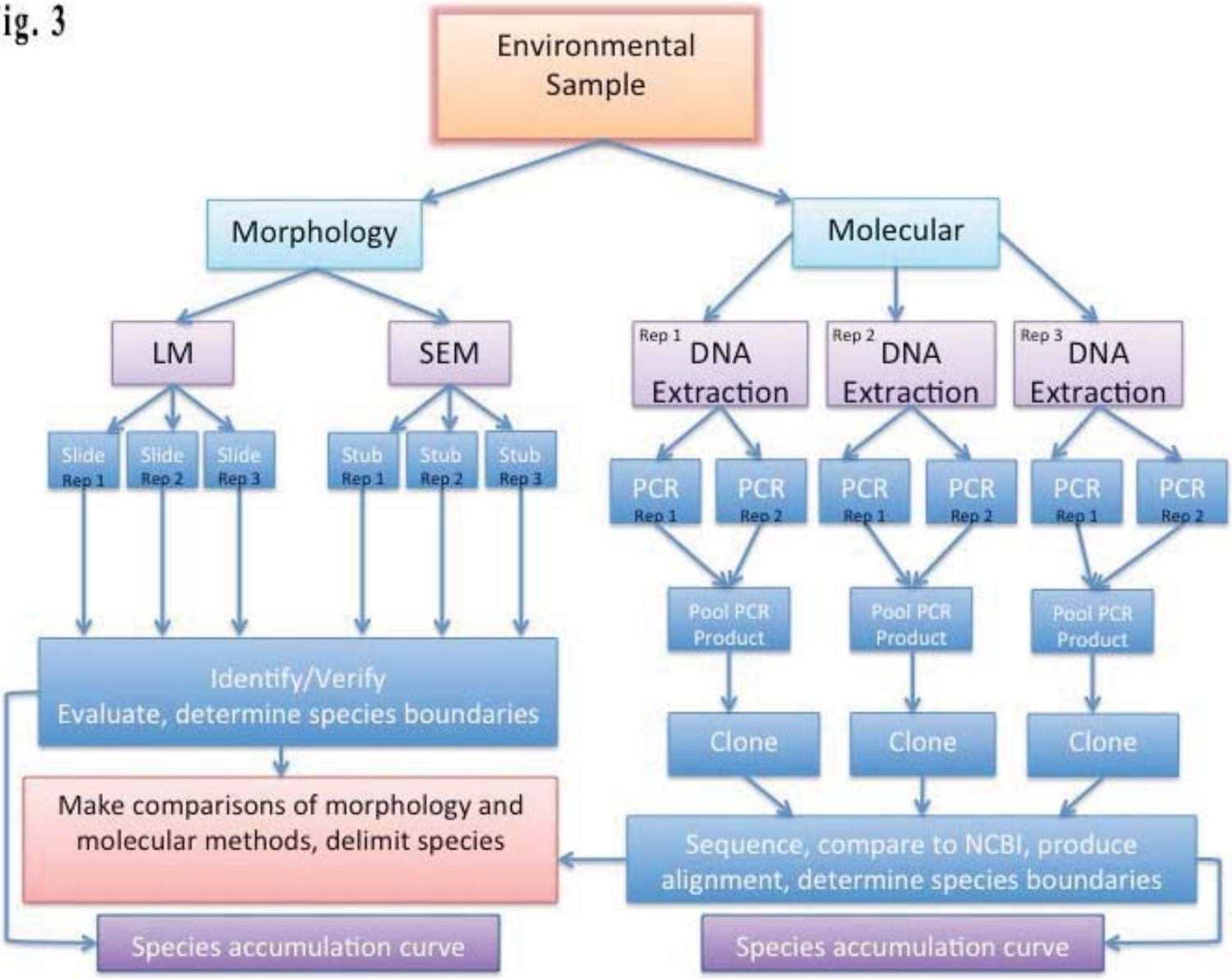


Figure 2. The Eightmile River confluence in Lyme Connecticut.

Fig. 3


Environmental Sample
Morphology
Molecular
LM
SEM
Rep 1 DNA Extraction
Rep 2 DNA Extraction
Rep 3 DNA Extraction
Slide Rep 1
Slide Rep 2
Slide Rep 3
Stub Rep 1
Stub Rep 2
Stub Rep 3
PCR Rep 1
PCR Rep 2
PCR Rep 1
PCR Rep 2
PCR Rep 1
PCR Rep 2
Pool PCR Product
Pool PCR Product
Pool PCR Product
Identify/Verify
Evaluate, determine species boundaries
Clone
Clone
Clone
Make comparisons of morphology and molecular methods, delimit species
Sequence, compare to NCBI, produce alignment, determine species boundaries
Species accumulation curve
Species accumulation curve

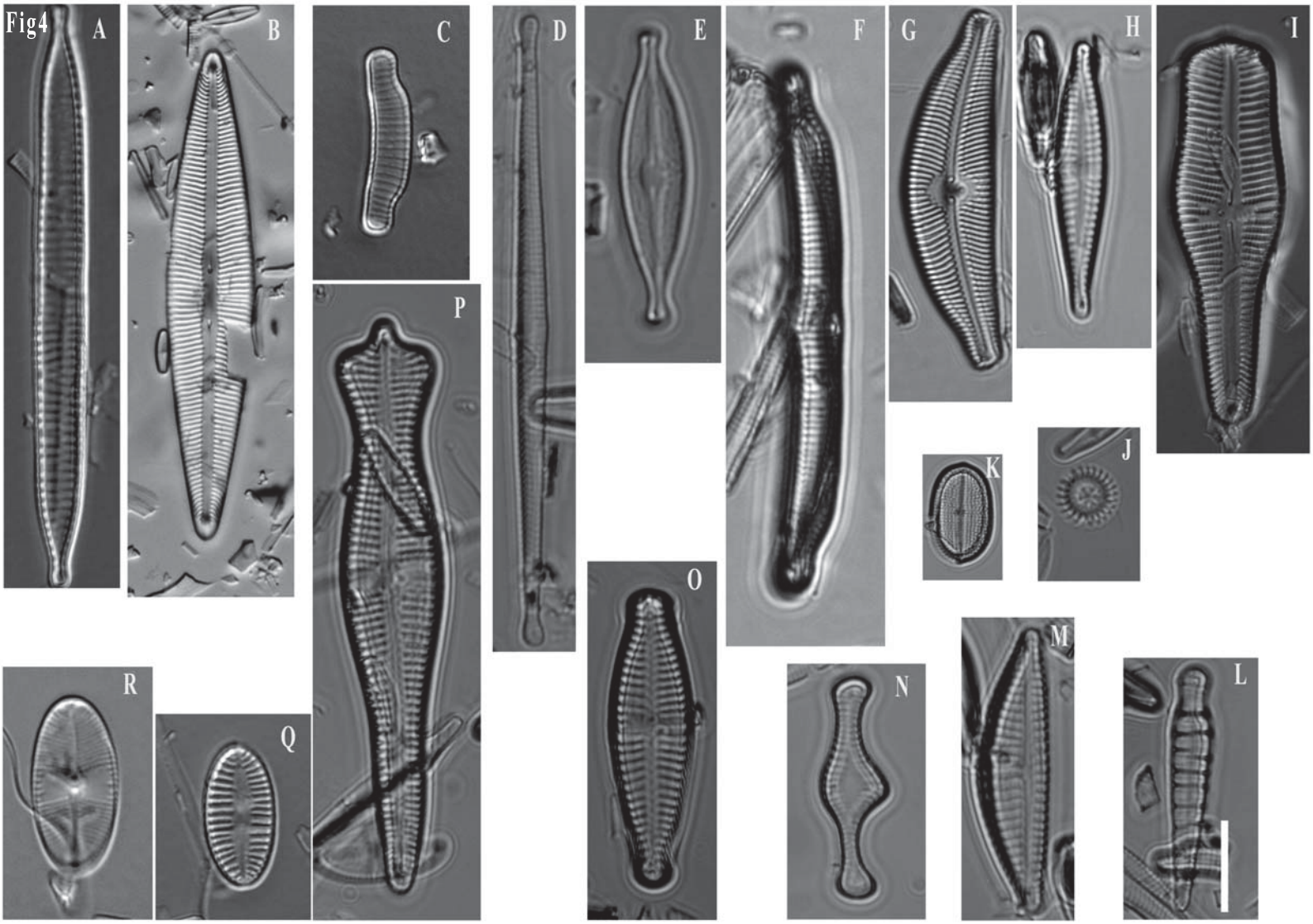


Fig4
A
B
C
D
E
F
G
H
I
P
K
J
O
R
Q
N
M
L

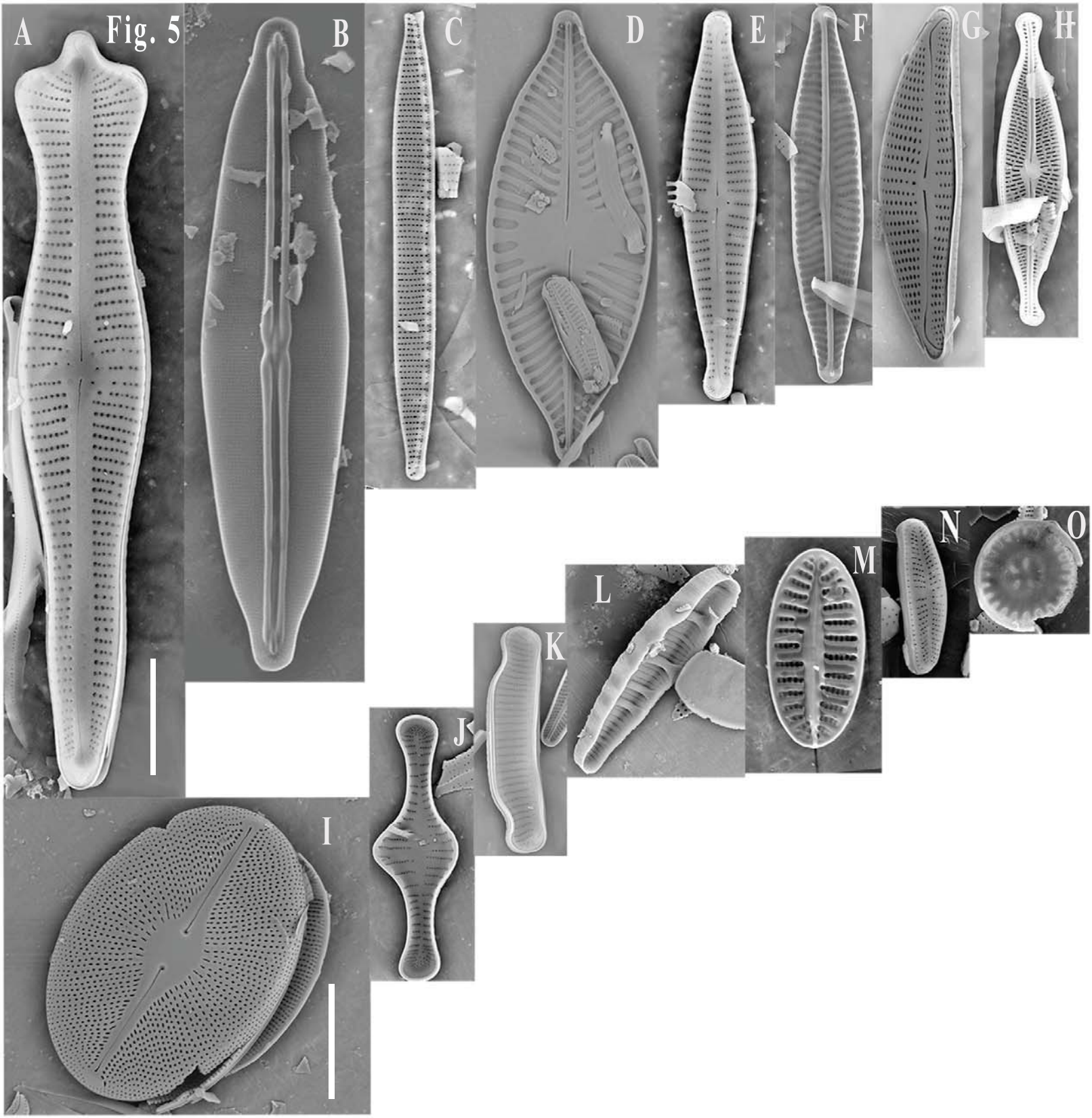
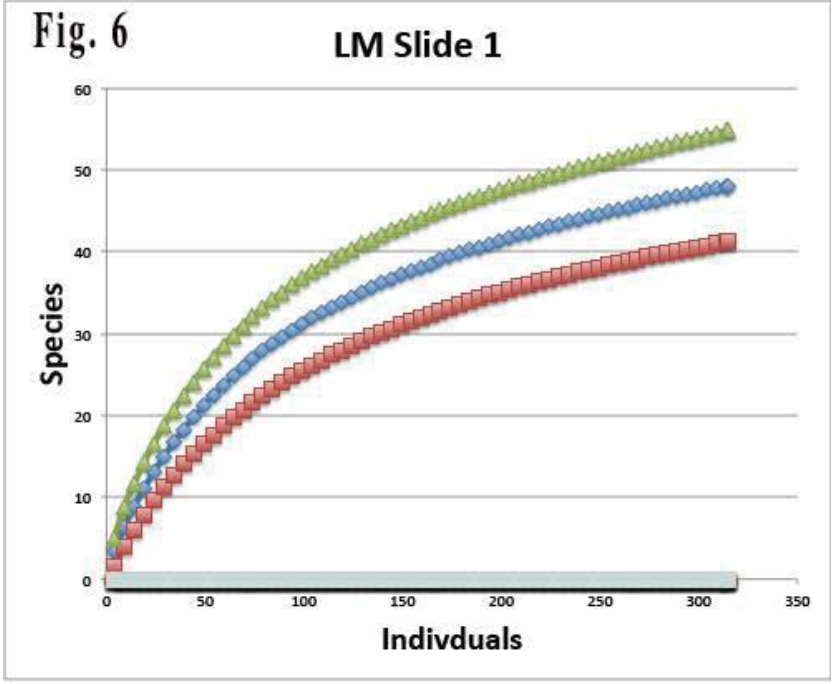
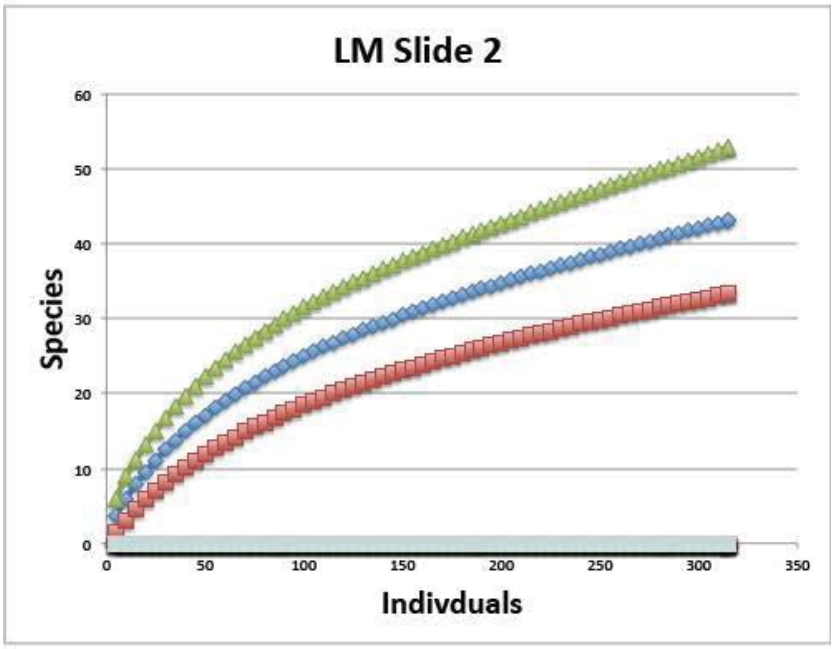

A
Fig. 5
B
C
D
E
F
G
H
I
J
K
L
M
N
O

Fig. 6


LM Slide 1
Species
Indivduals
0
10
20
30
40
50
60
0
50
100
150
200
250
300
350


LM Slide 2
Species
Indivduals
0
10
20
30
40
50
60
0
50
100
150
200
250
300
350

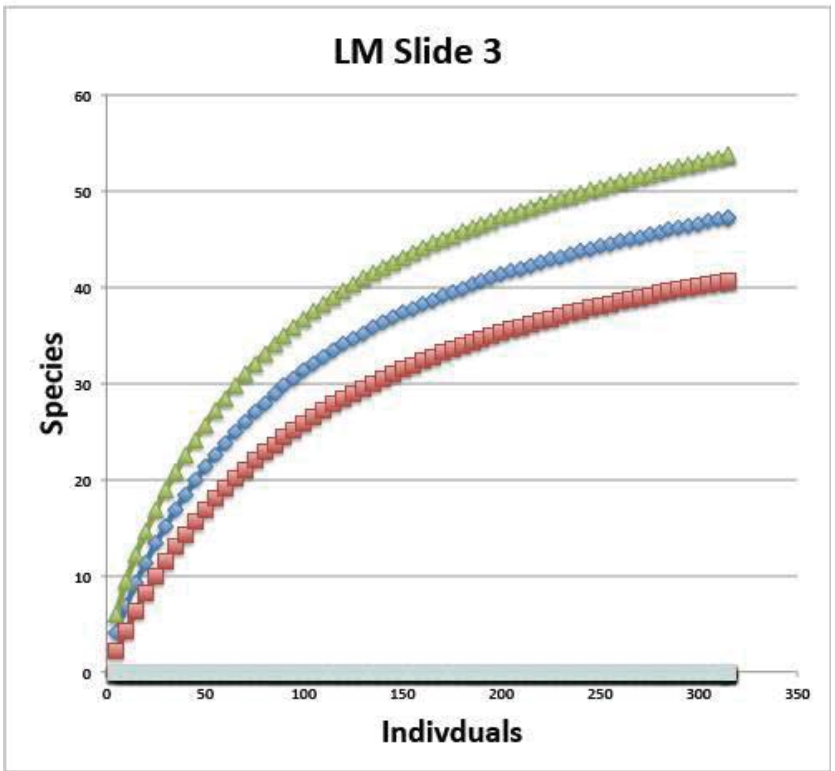

LM Slide 3
Species
Indivduals
0
10
20
30
40
50
60
0
50
100
150
200
250
300
350

Fig. 7

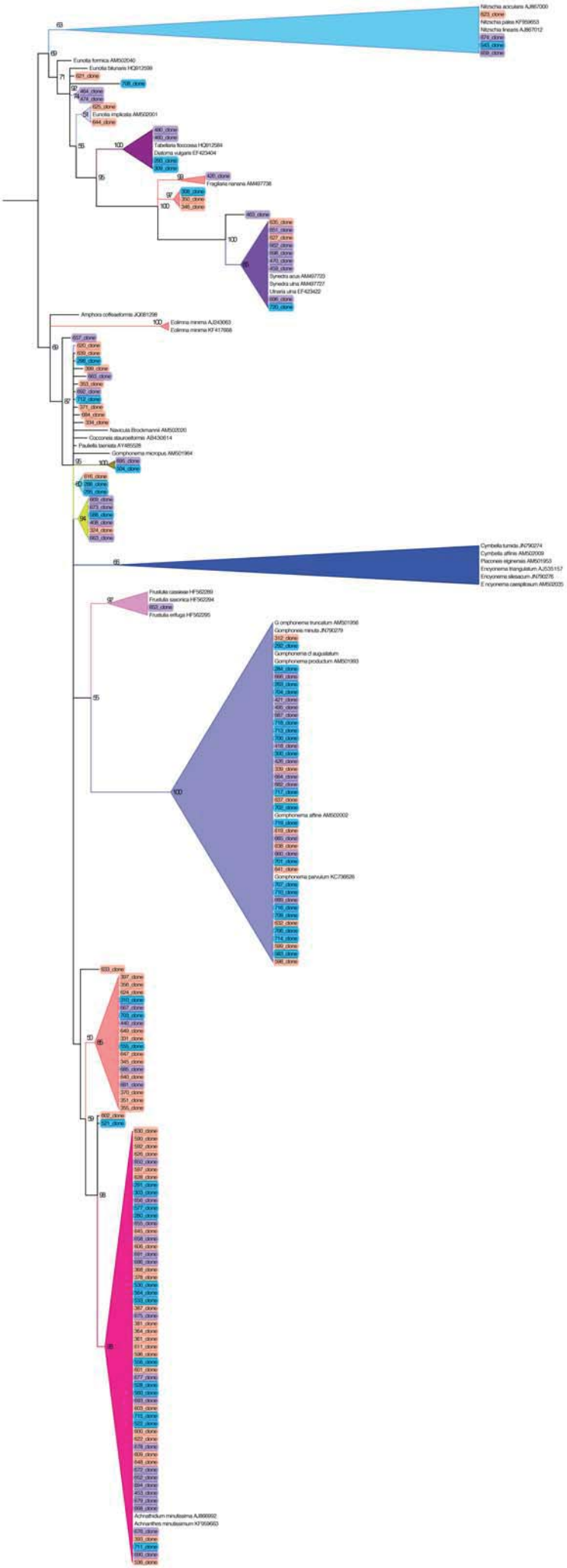

Nitzschia accicularis AJ867000
623_clone
Nitzschia palea KF959653
Nitzschia linearis AJ867012
674_clone
545_clone
659_clone
Eunotia formica AM920040
Eunotia faluvaris HQ912598
621_clone
708_clone
664_clone
674_clone
625_clone
Eunotia implicata AM920001
644_clone
680_clone
Tabellaria flocculosa HQ912584
Diatoma vulgaris EF423404
693_clone
639_clone
625_clone
Fragilaria nanana AM497708
698_clone
350_clone
346_clone
663_clone
635_clone
657_clone
627_clone
652_clone
696_clone
470_clone
458_clone
Synedra acus AM497723
Synedra ulna AM497727
Ulnaria ulna EF423422
696_clone
720_clone
Amphora coffeaeformis JQ081298
Eolimna minima AJ243063
Eolimna minima KF417668
637_clone
635_clone
639_clone
696_clone
398_clone
663_clone
393_clone
692_clone
712_clone
371_clone
684_clone
334_clone
Navicula Brockmannii AM920200
Cocconeis stauroeformis AB430614
Paulella bariseta AY485528
Gomphonema micropus AM921964
695_clone
504_clone
618_clone
709_clone
705_clone
668_clone
673_clone
646_clone
468_clone
324_clone
663_clone
Cymbella tumida JN790274
Cymbella affinis AM920069
Placoneis elginensis AM501953
Encyonema triangulatum AJ535157
Encyonema silesiacum JN790276
Encyonema caespitosum AM920035
Frustulia cassiniae HF562289
Frustulia saxonica HF562294
652_clone
Frustulia erifuga HF562295
G omphonema truncatum AM501956
Gomphonea minuta JN790279
312_clone
292_clone
Gomphonema cf augustatum
Gomphonema productum AM501993
284_clone
668_clone
293_clone
704_clone
421_clone
495_clone
697_clone
718_clone
712_clone
709_clone
418_clone
300_clone
625_clone
330_clone
664_clone
692_clone
717_clone
637_clone
702_clone
Gomphonema affine AM920002
716_clone
619_clone
665_clone
638_clone
690_clone
701_clone
641_clone
Gomphonema parvulum KC736626
707_clone
710_clone
669_clone
716_clone
709_clone
632_clone
706_clone
714_clone
589_clone
683_clone
596_clone
633_clone
397_clone
356_clone
624_clone
310_clone
367_clone
703_clone
440_clone
649_clone
371_clone
535_clone
647_clone
340_clone
665_clone
640_clone
681_clone
370_clone
351_clone
705_clone
602_clone
671_clone
630_clone
590_clone
592_clone
626_clone
602_clone
597_clone
628_clone
391_clone
303_clone
656_clone
627_clone
280_clone
655_clone
640_clone
658_clone
606_clone
691_clone
686_clone
368_clone
278_clone
630_clone
564_clone
703_clone
397_clone
675_clone
381_clone
364_clone
361_clone
611_clone
596_clone
636_clone
621_clone
677_clone
538_clone
682_clone
693_clone
603_clone
701_clone
622_clone
600_clone
622_clone
678_clone
639_clone
648_clone
672_clone
652_clone
694_clone
453_clone
679_clone
668_clone
Achnanthidium minutissima AJ866992
Achnanthes minutissimum KF959663
676_clone
393_clone
711_clone
690_clone
596_clone

Fig. 7A

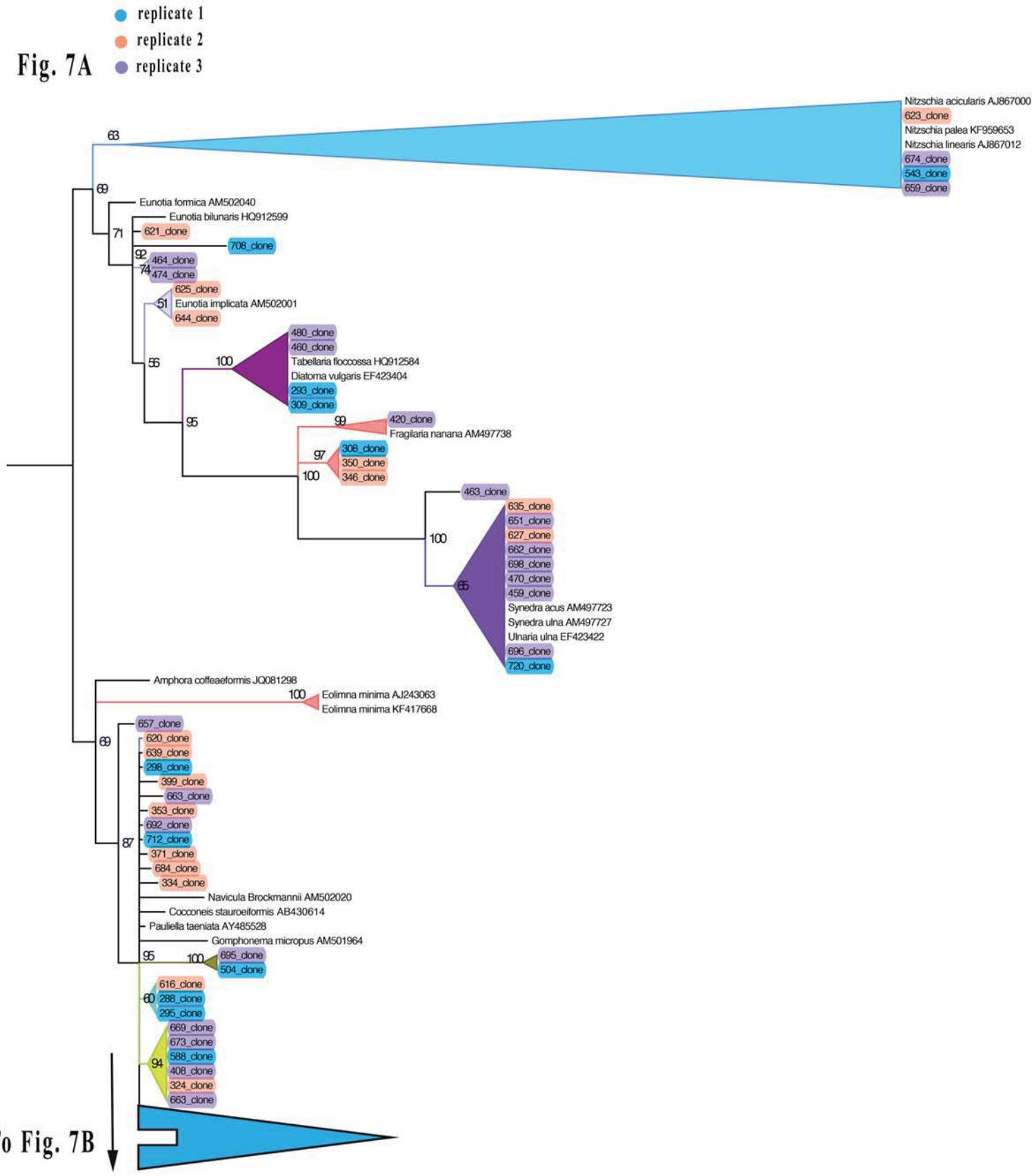

replicate 1
replicate 2
replicate 3
Nitzschia acicularis AJ867000
623_clone
Nitzschia palea KF959653
Nitzschia linearis AJ867012
674_clone
543_clone
659_clone
Eunotia formica AM502040
Eunotia bilunaris HQ912599
621_clone
708_clone
464_clone
474_clone
625_clone
Eunotia implicata AM502001
644_clone
480_clone
460_clone
Tabellaria floccossa HQ912584
Diatoma vulgaris EF423404
293_clone
309_clone
420_clone
Fragilaria nanana AM497738
308_clone
350_clone
346_clone
463_clone
635_clone
651_clone
627_clone
662_clone
698_clone
470_clone
459_clone
Synedra acus AM497723
Synedra ulna AM497727
Ulnaria ulna EF423422
696_clone
720_clone
Amphora coffeaeformis JQ081298
Eolimna minima AJ243063
Eolimna minima KF417668
657_clone
620_clone
639_clone
298_clone
399_clone
663_clone
353_clone
692_clone
712_clone
371_clone
684_clone
334_clone
Navicula Brockmannii AM502020
Cocconeis stauroeiformis AB430614
Pauliella taeniata AY485528
Gomphonema micropus AM501964
695_clone
504_clone
616_clone
288_clone
295_clone
669_clone
673_clone
588_clone
408_clone
324_clone
663_clone
To Fig. 7B

Fig. 7B

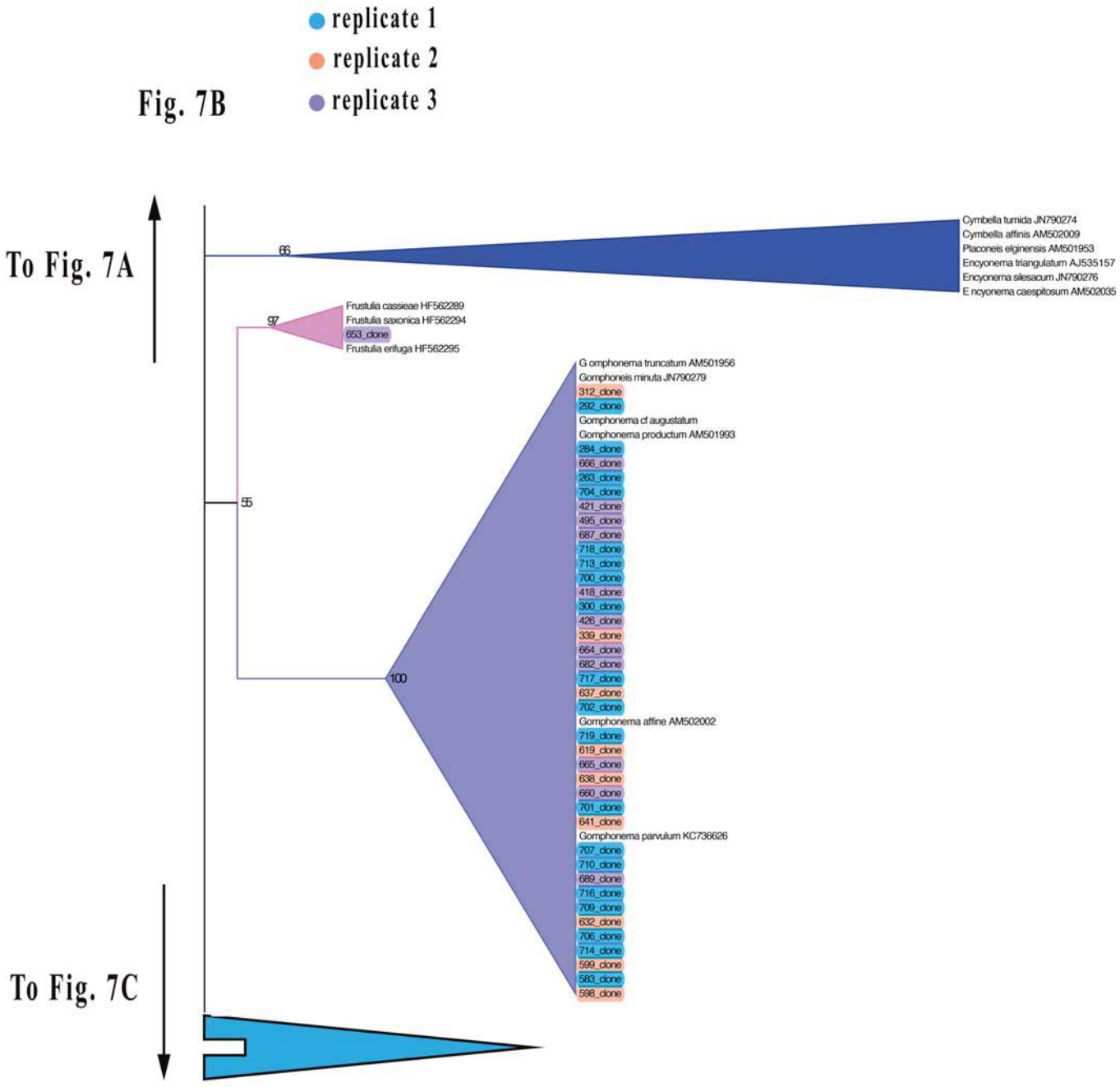

replicate 1
replicate 2
replicate 3
To Fig. 7A
To Fig. 7C
66
97
55
100
Cymbella tumida JN790274
Cymbella affinis AM502009
Placoneis elginensis AM501953
Encyonema triangulatum AJ535157
Encyonema silesacum JN790276
E ncyonema caespitosum AM502035
Frustulia cassieae HF562289
Frustulia saxonica HF562294
653_clone
Frustulia erifuga HF562295
G omphonema truncatum AM501956
Gomphoneis minuta JN790279
312_clone
292_clone
Gomphonema cf augustatum
Gomphonema productum AM501993
284_clone
666_clone
263_clone
704_clone
421_clone
495_clone
687_clone
718_clone
713_clone
700_clone
418_clone
300_clone
426_clone
339_clone
664_clone
682_clone
717_clone
637_clone
702_clone
Gomphonema affine AM502002
719_clone
619_clone
665_clone
638_clone
660_clone
701_clone
641_clone
Gomphonema parvulum KC736626
707_clone
710_clone
689_clone
716_clone
709_clone
632_clone
706_clone
714_clone
599_clone
583_clone
598_clone

Fig. 7C

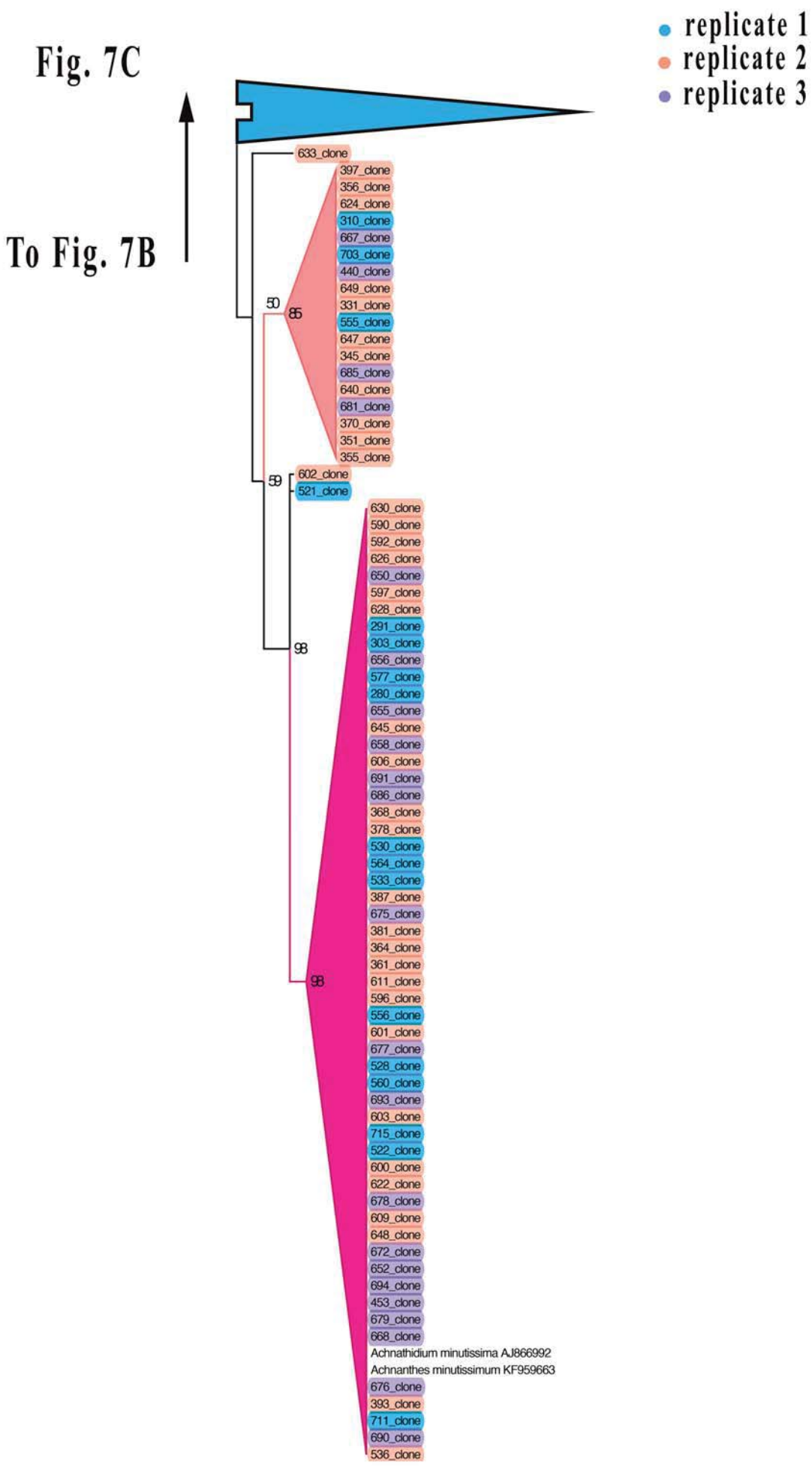
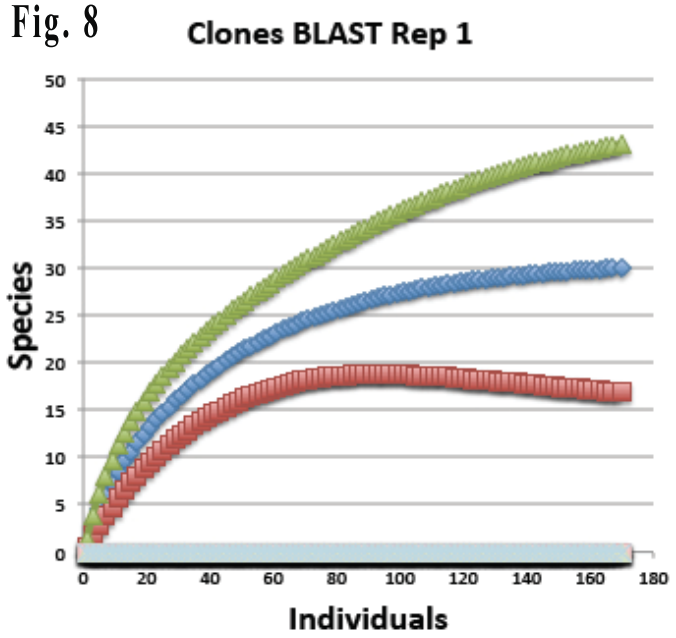
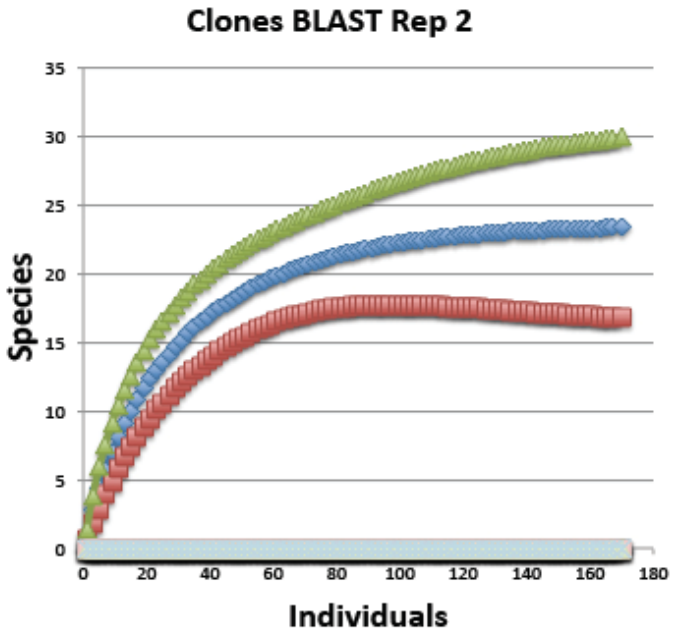

replicate 1
replicate 2
replicate 3
To Fig. 7B
633_clone
397_clone
356_clone
624_clone
310_clone
667_clone
703_clone
440_clone
649_clone
331_clone
555_clone
647_clone
345_clone
685_clone
640_clone
681_clone
370_clone
351_clone
355_clone
50
85
59
602_clone
521_clone
98
630_clone
590_clone
592_clone
626_clone
650_clone
597_clone
628_clone
291_clone
303_clone
656_clone
577_clone
280_clone
655_clone
645_clone
658_clone
606_clone
691_clone
686_clone
368_clone
378_clone
530_clone
564_clone
533_clone
387_clone
675_clone
381_clone
364_clone
361_clone
98
611_clone
596_clone
556_clone
601_clone
677_clone
528_clone
560_clone
693_clone
603_clone
715_clone
522_clone
600_clone
622_clone
678_clone
609_clone
648_clone
672_clone
652_clone
694_clone
453_clone
679_clone
668_clone
Achnathidium minutissima AJ866992
Achnanthes minutissimum KF959663
676_clone
393_clone
711_clone
690_clone
536_clone

Fig. 8


Clones BLAST Rep 1
Species
Individuals
0
5
10
15
20
25
30
35
40
45
50
0
20
40
60
80
100
120
140
160
180


Clones BLAST Rep 2
Species
Individuals
0
5
10
15
20
25
30
35
0
20
40
60
80
100
120
140
160
180

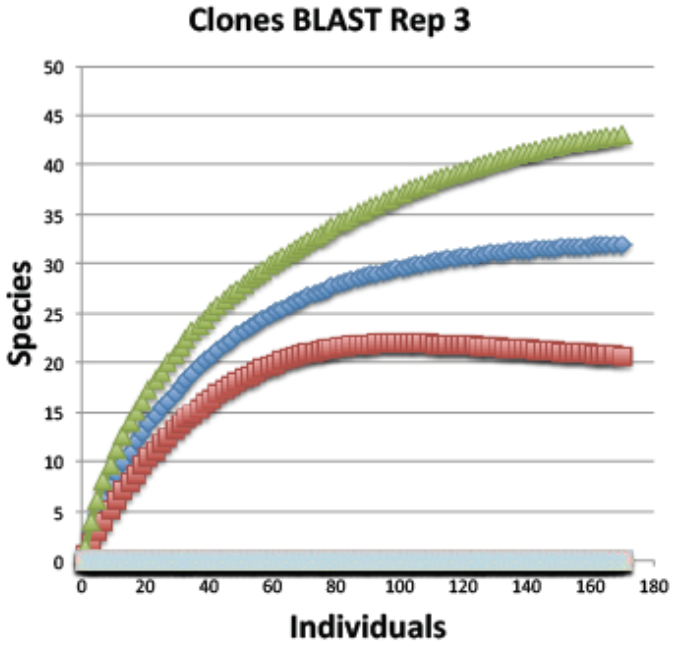
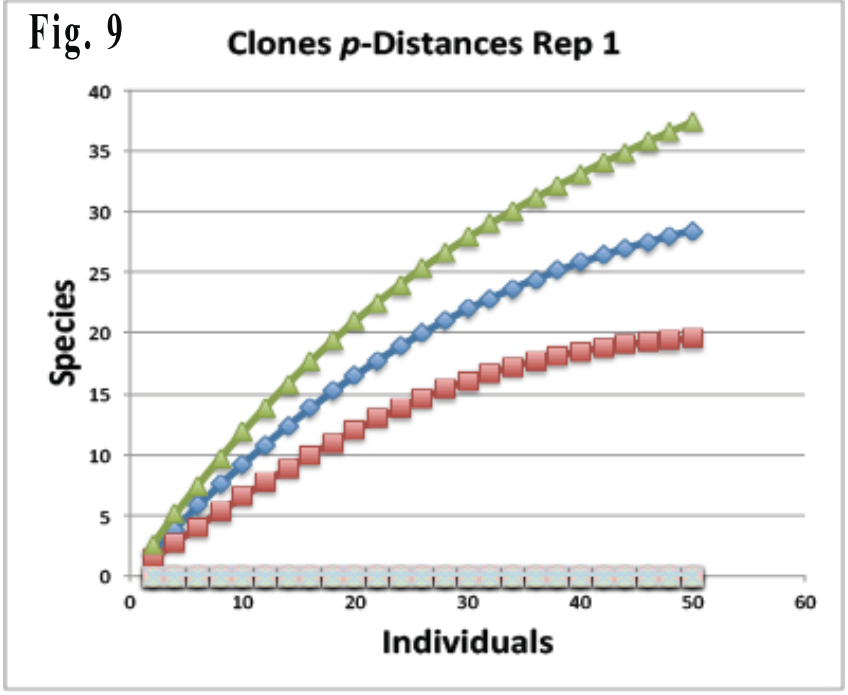
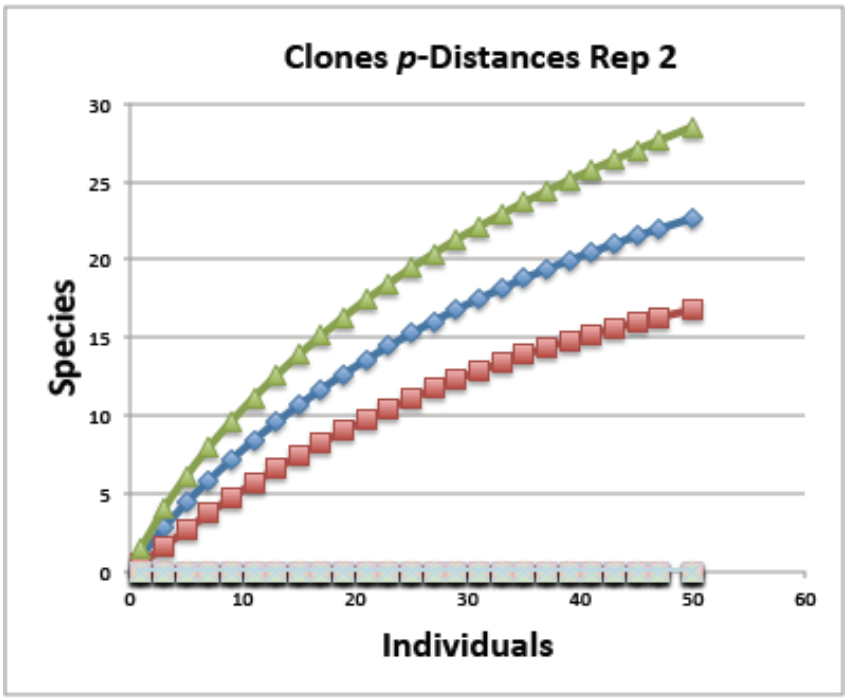

Clones BLAST Rep 3
Species
Individuals
0
5
10
15
20
25
30
35
40
45
50
0
20
40
60
80
100
120
140
160
180

Fig. 9


Clones p-Distances Rep 1
Species
Individuals
0
5
10
15
20
25
30
35
40
0
10
20
30
40
50
60


Clones p-Distances Rep 2
Species
Individuals
0
5
10
15
20
25
30
0
10
20
30
40
50
60

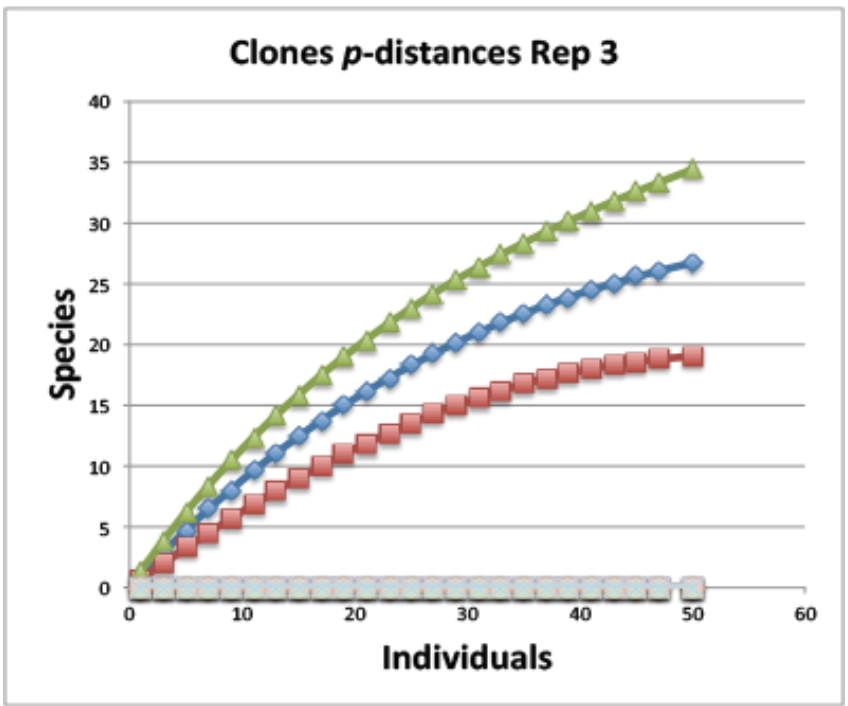

Clones p-distances Rep 3
Species
Individuals
0
5
10
15
20
25
30
35
40
0
10
20
30
40
50
60

Table 1. Diatom taxa detected by LM and SEM are shown and demarcated with +. Diatom clones detected using the NCBI BLASTn tool are demarcated with +. GenBank accession numbers are shown for those taxa that have a sequence on the NCBI website. Taxa that are not represented on the NCBI website are demarcated with -. * Denotes [1]homotypic synonym and [2]basionym taxonomy.

Taxon	GenBank Accession Number	Detected LM	Detected Clones
Adlafia suchlandtii (Hustedt) Lange-Bertalot in Gerd Moser, Lange-Bertalot & D.Metzeltin 1998	-	+	0
Achnanthidium minutissimum (Kützing) Czarnecki 1994	AJ866992	+	+
Achnanthidium gracillimum (Meister) Lange-Bertalot in Krammer & Lange-Bertalot 2004	-	+	0
Achnanthes lanceolata (Brébisson ex Kützing) Grunoin Van Heurck 1880	-	+	0
Achnanthes chlidanos M. H. Hohn & Hellerman 1963	-	0	0
Achnanthes pusilla Grunow in Cleve & Grunow 1880	-	+	0
Achnanthes minutissmum(Kützing) 1833	KF959663	+	+
Amphora coffeaeformis (C.Agardh) Kützing 1844	JQ081298	0	+
Amphora graeffeana Hendey 1973	KJ463430	+	0
Cocconeis placentula Ehrenberg 1838	KM084976	+	0
Cocconeis stauroneiformis (W. Smith) H. Okuno 1957	AB430614	0	+
Cyclotella atomus Hustedt 1937	DQ514858	+	0
Cymbella affinis Kützing 1844	AM502009	+	0
Cymbella tumida (Brébisson) van Heurck 1880	JN790274	+	+
Diatoma vulgaris Bory de Saint-Vincent 1824	EF423404	0	+
Discostella stelligera (Cleve & Grunow) Houk & Klee 2004	DQ514903	+	0
Discostella pseudostelligera (Hustedt) Houk & Klee 2004	DQ514905	+	0

Encyonema caespitosum Kützing 1849	AM502035	+	0
Encyonema silesiacum (Bleisch) D.G.Mann 1990	JN790276	+	0
Encyonema triangulum (Ehrenberg) Kützing 1849	AJ535157	+	0
Eolimna minima (Grunow) Lange-Bertalot & W. Schiller in W. Schiller & Lange-Bertalot 1997	AJ243063	+	0
Eolimna minima (Grunow) Lange-Bertalot & W.Schiller in W.Schiller & Lange-Bertalot 1997	KF417668	+	0
Eunotia bilunaris (Ehrenberg)Schaarschmidt 1880	HQ912599	+	+
Eunotia flexuosa (Brébisson ex Kützing) Kützing 1849	-	+	0
Eunotia formica Ehrenberg 1843	AM502040	+	0
Eunotia implicata Nörpel, Lange-Bertalot & Alles in E.Alles, M.Nörpel-Schempp, & H.Lange-Bertalot 1991	AM502001	+	+
Eunotia incisa W.Smith ex W.Gregory 1854	-	+	0
Eunotia sp Ehrenberg, 1837	AM501963	+	+
Eunotia cf *praerupta* Ehrenberg 1843	-	+	0
Eunoti pectinalis var *undulata* (Ralfs) Rabenhorst 1864	AB085832	+	0
Fragilaria bidens Heiberg 1863	AB430599	0	+
Fragilaria capucina Desmazières 1830	KC736619	+	0
Fragilaria rumpens (Kützing) G. W. F. Carlson 1913	KF959661	+	+
[1]*Fragilaria nanana* Lange-Bertalot 1993* (homotypic synonym)	AM497738	+	+
Frustulia cassieae Lange-Bertalot & Beier	HF562289	+	0
Frustulia crassinervia (Brébisson) Lange-Bertalot and Krammer 1996	HF562259	+	0
Frustulia erifuga Lange-Bertalot & Krammer in Lange-Bertalot & Metzeltin 1996	HF562295	0	+
Frustulia saxonica Rabenhorst 1853	HF562294	0	+
Gomphonema micropus (Kützing) 1844	AM501964	+	0
Gomphoneis minuta (Stone) Kociolek & Stoermer	JN790279	+	0
Gomphonema affine (Kützing) 1844	AM502002	+	0

Gomphonema angustatum (Kützing) Rabenhorst 1864	AM502005	+	+
Gomphonema productum (Grunow) Lange-Bertalot & Reichardt in Lange-Bertalot 1993	AM501993	+	+
Gomphonema parvulum (Kützing) Kützing 1849	KC736626	+	+
Gomphonema truncatum Ehrenberg 1832	AM501956	+	0
Gomphonema sp Ehrenberg 1832	KM507847	+	0
Karayevia oblongella (Østrup) M.Aboal in Aboal et al. 2003	-	+	0
Meridion circulare (Greville) C.Agardh 1831	-	+	0
Navicula cryptocephala Kützing 1844	KM084881	+	0
Navicula notha Wallace 1960	-	+	0
Navicula rhyncocephala (Patrick & Reimer 1966)	-	+	0
Navicula brockmannii Hustedt 1934	AM502020	+	0
Nitzschia palea var. debillis (Kützing) Grunow 1880	KM507859	+	+
Nitzschia acidoclinata Lange-Bertalot 1976	KC736632	+	0
Nitzschia hantzschiana Rabenhorst 1860	-	+	0
Nitzschia linearis W. Smith 1853	AJ867012	+	0
Pauliella toeniata (Grunow) F. E. Round & P. W. Basson 1997	AY485528	0	+
Placoneis elginensis (Gregory) E. J. Cox 1988	AM501953	+	0
Planothidium apiculatum (R.M.Patrick) Lange-Bertalot 1999	-	+	0
[2]*Synedra acus* Kützing 1844*	AM497723	+	+
[1]*Synedra nana* F. Meister 1912*	-	+	0
Synedra goulardii Brébisson ex Cleve & Grunow 1880	-	+	0
Synedra rumpens Kützing 1844	-	+	+
Synedra ulna (Nitzsch) Ehrenberg 1832	AM497727	+	+
Tabellaria flocculosa (Roth) Kützing 1844	HQ912584	+	+
Ulnaria ulna (Nitzsch) P. Compère in Jahn et al. 2001	EF423422	+	+
Ulnaria sp (Kützing) Compère 2001 (Kützing) Compère 2001	-	+	0
[2]*Ulnaria acus* (Kützing) M. Aboal in Aboal, Alvarez Cobelas, Cambra & Ector 2003* (Basionym)	KF959659	0	+

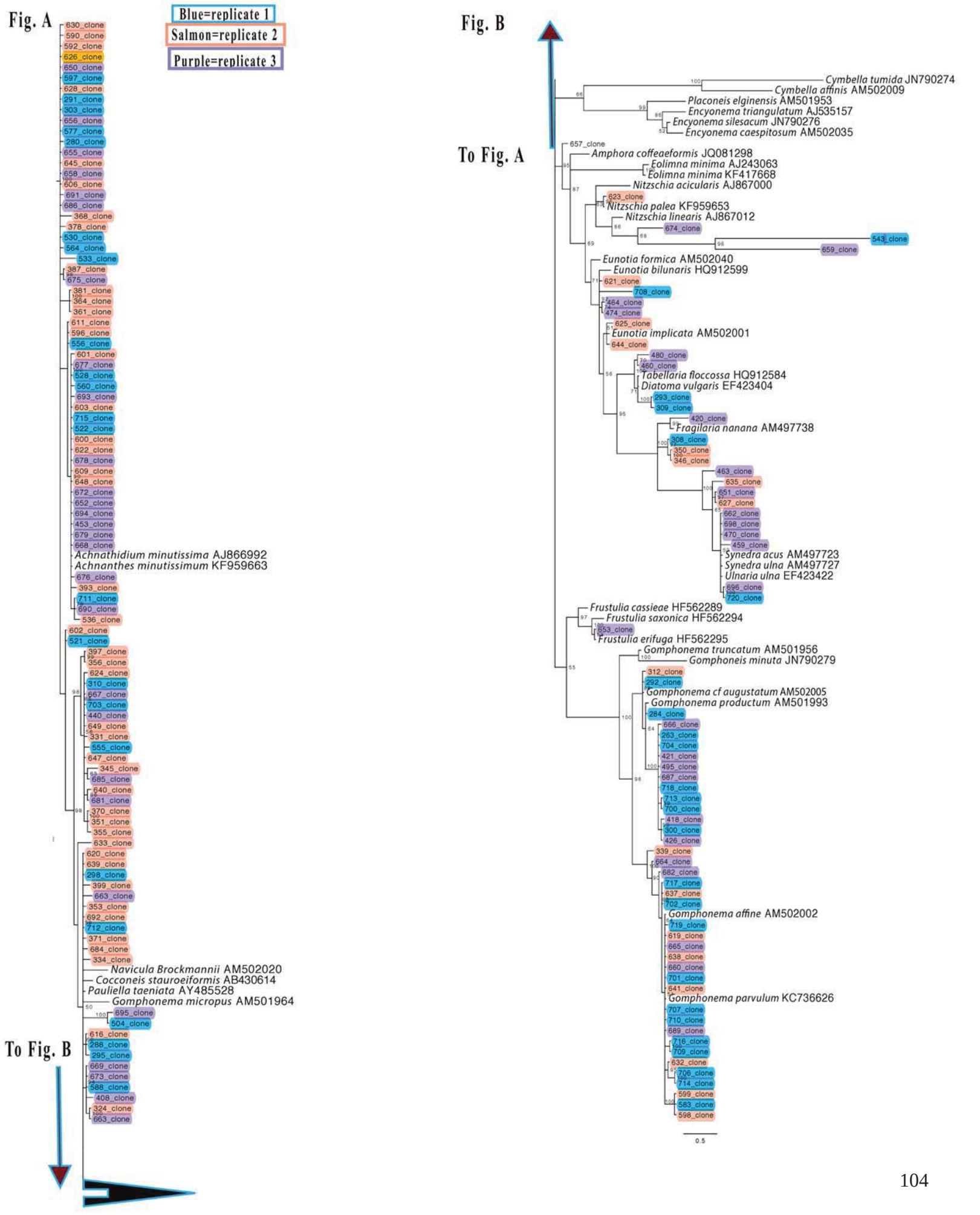
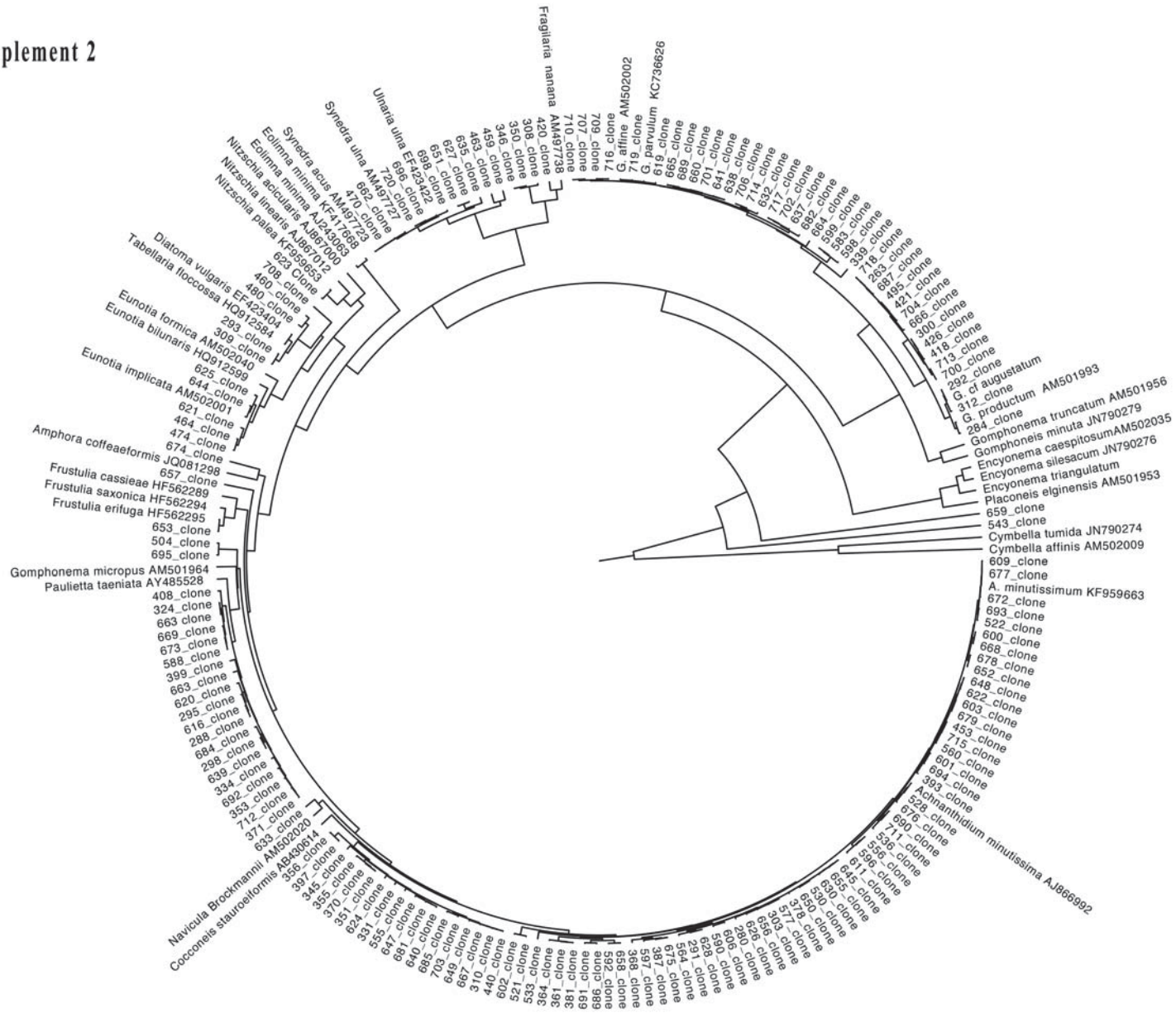


Fig. A
Blue=replicate 1
Salmon=replicate 2
Purple=replicate 3
630_clone
590_clone
592_clone
626_clone
650_clone
597_clone
628_clone
291_clone
303_clone
656_clone
577_clone
280_clone
655_clone
645_clone
658_clone
606_clone
691_clone
686_clone
368_clone
378_clone
530_clone
564_clone
533_clone
387_clone
675_clone
381_clone
364_clone
361_clone
611_clone
596_clone
556_clone
601_clone
677_clone
528_clone
560_clone
693_clone
603_clone
715_clone
522_clone
600_clone
622_clone
678_clone
609_clone
548_clone
672_clone
652_clone
694_clone
453_clone
679_clone
668_clone
Achnathidium minutissima AJ866992
Achnanthes minutissimum KF959663
676_clone
393_clone
711_clone
690_clone
536_clone
602_clone
521_clone
397_clone
356_clone
624_clone
310_clone
667_clone
703_clone
440_clone
649_clone
331_clone
555_clone
647_clone
345_clone
685_clone
640_clone
681_clone
370_clone
351_clone
355_clone
633_clone
620_clone
639_clone
298_clone
399_clone
663_clone
353_clone
692_clone
712_clone
371_clone
684_clone
334_clone
Navicula Brockmannii AM502020
Cocconeis stauroeiformis AB430614
Pauliella taeniata AY485528
Gomphonema micropus AM501964
695_clone
504_clone
616_clone
288_clone
295_clone
669_clone
673_clone
588_clone
408_clone
324_clone
663_clone
To Fig. B
Fig. B
Cymbella tumida JN790274
Cymbella affinis AM502009
Placoneis elginensis AM501953
Encyonema triangulatum AJ535157
Encyonema silesacum JN790276
Encyonema caespitosum AM502035
To Fig. A
657_clone
Amphora coffeaeformis JQ081298
Eolimna minima AJ243063
Eolimna minima KF417668
Nitzschia acicularis AJ867000
623_clone
Nitzschia palea KF959653
Nitzschia linearis AJ867012
674_clone
543_clone
659_clone
Eunotia formica AM502040
Eunotia bilunaris HQ912599
621_clone
708_clone
464_clone
474_clone
625_clone
Eunotia implicata AM502001
644_clone
480_clone
460_clone
Tabellaria floccossa HQ912584
Diatoma vulgaris EF423404
293_clone
309_clone
420_clone
Fragilaria nanana AM497738
308_clone
350_clone
346_clone
463_clone
635_clone
651_clone
627_clone
662_clone
698_clone
470_clone
459_clone
Synedra acus AM497723
Synedra ulna AM497727
Ulnaria ulna EF423422
696_clone
720_clone
Frustulia cassieae HF562289
Frustulia saxonica HF562294
653_clone
Frustulia erifuga HF562295
Gomphonema truncatum AM501956
Gomphoneis minuta JN790279
312_clone
292_clone
Gomphonema cf augustatum AM502005
Gomphonema productum AM501993
284_clone
666_clone
263_clone
704_clone
421_clone
495_clone
687_clone
718_clone
713_clone
700_clone
418_clone
300_clone
426_clone
339_clone
664_clone
682_clone
717_clone
637_clone
702_clone
Gomphonema affine AM502002
719_clone
619_clone
665_clone
638_clone
660_clone
701_clone
641_clone
Gomphonema parvulum KC736626
707_clone
710_clone
689_clone
716_clone
709_clone
632_clone
706_clone
714_clone
599_clone
583_clone
598_clone
0.5

Supplement 2



Supplement 3

Table 2. Estimates of Evolutionary Divergence between Sequences
The number of base differences per site from between sequences are shown.
The analysis involved 131 nucleotide sequences. All positions containing gaps and missing data were eliminated.

There were a total of 12 positions in the final dataset. Evolutionary analyses were conducted in MEGA6

	630_clone	590_clone	592_clone	597_clone	387_clone	675_clone	280_clone	691_clone	686_clone	368_clone	378_clone
630_clone											
590_clone	0.005										
592_clone	0.002	0.002									
597_clone	0.003	0.003	0.000								
387_clone	0.005	0.005	0.002	0.003							
675_clone	0.005	0.005	0.002	0.003	0.000						
280_clone	0.005	0.005	0.003	0.003	0.005	0.005					
691_clone	0.005	0.005	0.002	0.003	0.005	0.005	0.005				
686_clone	0.003	0.002	0.000	0.000	0.003	0.002	0.003	0.002			
368_clone	0.015	0.015	0.012	0.013	0.015	0.015	0.015	0.015	0.012		
378_clone	0.007	0.007	0.005	0.005	0.007	0.007	0.008	0.007	0.005	0.017	
381_clone	0.012	0.012	0.010	0.010	0.012	0.012	0.013	0.012	0.010	0.022	0.015
364_clone	0.012	0.012	0.010	0.010	0.012	0.012	0.013	0.012	0.010	0.022	0.015
564_clone	0.005	0.005	0.002	0.003	0.005	0.005	0.005	0.005	0.002	0.015	0.007