

A DNA Fingerprinting Approach for Distinguishing Native and Non-native Milfoils

Ryan A. Thum¹ and Jay T. Lennon²

Department of Biological Sciences
Dartmouth College
Hanover, NH 03755

Jody Connor and Amy P. Smagula

New Hampshire Department of Environmental Services
29 Hazen Drive
Concord, NH 03301

Abstract

Thum, R.A., J.T. Lennon, J. Connor and A.P. Smagula. 2006. A DNA fingerprinting approach for distinguishing among native and non-native milfoils. *Lake and Reserv. Manage.* Vol. 22(1):1-6.

Variable-leaf water milfoil, *Myriophyllum heterophyllum*, is a non-native aquatic plant that has become a major management concern in New England. One key obstacle for effective management is accurate identification of native and non-native milfoil species. We used DNA sequences from the nuclear ribosomal DNA internal transcribed spacers (ITS) to identify non-native populations of *M. heterophyllum*. We found a number of discrepancies among morphological and genetic identifications, including individuals that were morphologically identified as natives but genetically identified as non-natives, and vice versa. We attribute these discrepancies to inaccurate identifications arising from morphological similarities among milfoil species. To help remedy this problem, we developed a restriction enzyme assay that distinguishes non-native *M. heterophyllum* from native milfoils. The assay provides a reliable method for identifying *M. heterophyllum* and therefore should facilitate lake management decisions concerning native and non-native milfoil populations.

Key Words: invasive species, macrophyte, ribosomal DNA, species identification

Invasive species pose potential threats to biodiversity and functioning of freshwater ecosystems (Wilcove *et al.* 1998, Ruiz *et al.* 1999, Mack *et al.* 2000). In particular, the colonization and aggressive growth of non-native aquatic plants are major concerns for lake managers because of high costs associated with management (Pimentel *et al.* 2000). Recently, non-native variable-leaf water milfoil (*Myriophyllum heterophyllum*) has received considerable attention in New England waterbodies and is now recognized as the most commonly occurring non-native aquatic plant in New Hampshire (NH – Department of Environmental Services). It has been estimated that lake-front property values in New Hampshire may decline by as much as 20% after *M. heterophyllum* infestation (Halstead *et al.* 2003). As such, there are obvious

economic and recreational interests in controlling the spread and aggressive growth of this species.

Aside from a lack of understanding of the factors that facilitate invasions, management efforts for *M. heterophyllum* are complicated by an inability to accurately distinguish it from native milfoil species. Milfoil species are notoriously difficult to distinguish because of morphological similarities. For example, *M. heterophyllum* has commonly been mistaken for the native New Hampshire water-milfoil species, *M. verticillatum*, in field and herbarium specimens (Les and Mehrhoff 1999). Milfoils are often identified based on floral characters, but these are not always available for inspection due to short flowering duration and the propensity for vegetative propagation. Morphologically based identification of milfoil species can be further complicated by hybridization (Aiken 1981, Ceska and Ceska 1985), which has been reported for field-collected samples from some New England lakes (Moody and Les 2002). Accurate, reliable and year-round diagnostic

¹Current address: Cornell University, Dept. of Ecology and Evolutionary Biology, E303 Corson Hall, Ithaca, NY 14853

²Current address: Dept. of Ecology and Evolutionary Biology, Box G, Brown University, Providence, RI, 02912

characters are therefore needed to distinguish native milfoils from non-native *M. heterophyllum*.

The capacity to accurately distinguish among native and non-native milfoils has at least two direct implications for management efforts. First, accurate identification is often required for the disbursement of funds for non-native plant management (*e.g.*, NH Law RSA 487:17 III). Second, misidentification of native and non-native milfoils could lead to increased biological and economic impacts, especially if time and money is spent managing the wrong species. For example, the removal of native milfoils that are mistakenly identified as non-native *M. heterophyllum* might increase the susceptibility of a lake to invasion by non-native species, such as *M. heterophyllum*, by creating an open niche. Similarly, the misidentification of non-native *M. heterophyllum* as a native milfoil allows for the persistence of a source population of *M. heterophyllum* that could potentially colonize nearby lakes. Thus, correctly identifying *M. heterophyllum* populations, even where they do not grow aggressively, may allow for early treatment and effective lake management (Moody and Mack 1988).

In this paper, we evaluate the utility of nuclear ribosomal internal transcribed spacers (ITS) to reliably identify native and non-native milfoils in New Hampshire lakes. Using a phylogenetic approach, we demonstrate that morphological and genetic identifications of different milfoil species are often incongruent; these discrepancies most likely result from misidentification of specimens based on morphology alone. We therefore developed a restriction enzyme assay to distinguish non-native *M. heterophyllum* from native New Hampshire milfoils without having to actually sequence their DNA.

Methods

Sample Collection and Morphological Identification of Milfoil Specimens

We collected 53 milfoil samples from 44 New Hampshire lakes. For each sample we cut an apical meristem from a single plant to use for DNA extraction. We identified species by examining morphological characters (leaflet count, number of leaves per whorl, spacing between leaf whorls, and the presence of nutlets) according to Crow and Hellquist (2000). The identified samples were then briefly submersed in clean tap water and gently agitated to remove attached algae and macro-invertebrates. We then freeze-dried specimens for approximately 48 h in a Labconco 77500 bench-top freeze dryer (-45°C).

DNA Extraction, Polymerase Chain Reaction (PCR) and Sequencing

We used standard laboratory techniques for obtaining, amplifying and sequencing DNA; we refer readers to Palumbi *et al.* (1996) for a simple explanation of many of these techniques. Total genomic DNA was extracted from milfoils using DNEasy Plant Mini Kits (Qiagen). We amplified the internal transcribed spacers 1 and 2 and the 5.8S ribosomal DNA subunit between them (referred to as ITS from this point forward) with primers ITS1 and ITS4 from Soltis and Kuzoff (1995). This gene is commonly employed in molecular systematics because it generally shows variation among species but little to no variation within species. Polymerase Chain Reaction recipes contained: 2.5 µL buffer (GibCo), 1 µL MgCl₂ (2 mM), 2.5 µL of each primer, 2.5 µL dNTPs, 1 unit of Taq (GibCo), 2 µL template DNA filled to a final volume of 25 µL with sterile, distilled and deionized water. Thermal cycling for DNA amplification was carried out as the following: one cycle at 94°C for 2 min followed by 25 cycles of 94°C for 1 min, 56°C for 30 seconds, 72°C for 1 min and a final extension following those 25 cycles at 72°C for 8 min. We ran PCR products on an agarose gel (1.5%) to check for correct size and purity. To ensure effective sequencing, we purified PCR products (*i.e.*, removed unincorporated PCR reagents) using the Qiagen PCR Purification Kit.

We sequenced the purified PCR products using the BigDye® Terminator chemistry (version 3.1; Applied Biosystems (ABI)) on an ABI-3100 automated DNA sequencer at Dartmouth College's Molecular Biology Core Facility. Two *M. verticillatum* samples yielded poor sequences with multiple peaks when they were directly sequenced. Therefore we cloned these PCR products using the T-Easy cloning kit (Promega) and sequenced in the manner above.

Prior to initiating any phylogenetic analyses, the DNA sequences from different samples must be properly aligned with one another so that homologous nucleotide sites are compared. To accomplish this, we first visually inspected the electropherograms of our milfoil sequences with Sequencher (version 4.0.5) to double check base calling at nucleotide positions where sequences differed. Next, we aligned the DNA sequences from our milfoil samples with the *Myriophyllum* sequences available on GenBank (accessions AF513822-AF513850) from Moody and Les (2002) using ClustalX (version 1.81, Thompson *et al.* 1997) to identify species genetically based on their phylogenetic relationships with respect to those on GenBank.

Phylogenetic Analysis

We employed standard phylogenetic techniques to compare our DNA sequence data to the *Myriophyllum* sequences on GenBank (for a general text see Hall 2004). We performed

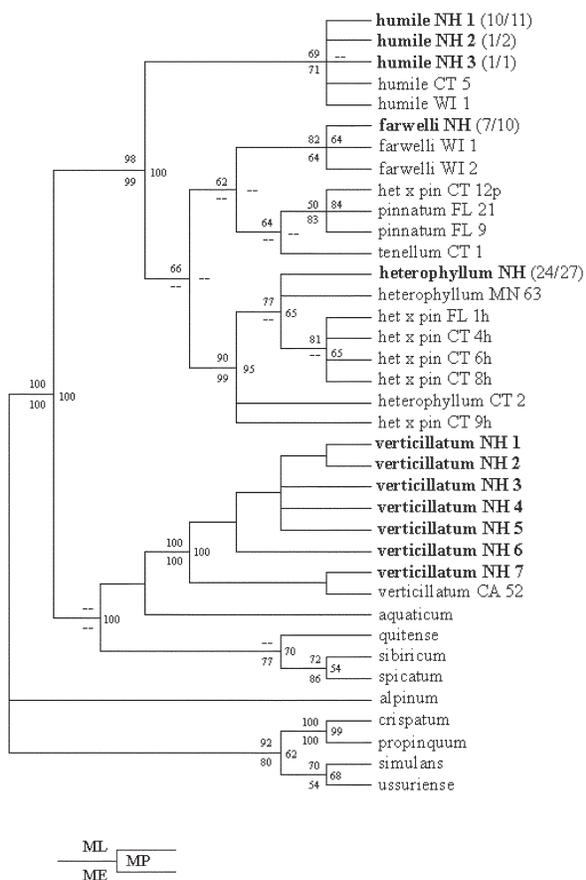


Figure 1.—Phylogenetic relationships among ITS alleles found in our samples (bold) and GenBank accessions (non-bold). “het x pin” samples refer to heterophyllum or pinnatum alleles found in hybrid plants in Moody and Les (2002). Numbers in parentheses refer to the number of lakes each allele was found in and the total number of individuals across all samples with each allele, respectively. All “verticillatum” alleles were found only once and came from a single lake (Balch Lake). Tree shown is the strict consensus of the two most likely trees. ML, ME and MP refer to statistical support of nodes from maximum likelihood, minimum evolution, and maximum parsimony analyses, respectively.

maximum parsimony, minimum evolution, and maximum likelihood heuristic searches in PAUP version 4.0 (Swoford 1998). We used the Tamura-Nei model of nucleotide substitution (A-G=1.6751, C-T=4.1141) in maximum likelihood and minimum evolution analyses with parameters for nucleotide substitution rates, base frequencies, proportion of invariable sites, and gamma shape as estimated by ModelTest (Posada and Crandall 1998). We evaluated statistical support for nodes in the phylogenetic analyses by bootstrapping (maximum parsimony and minimum evolution analyses; 1,000 replicates) and quartet puzzling (maximum likelihood analysis; 1,000 puzzling steps). Finally, we assessed the de-

gree to which species identifications based on morphology and genetics were concordant.

Restriction Enzyme Identification and Digestion

Based on our DNA sequence data, we looked for restriction enzyme sites that could distinguish *M. heterophyllum* from native milfoils. Restriction enzymes cut DNA at specifically recognized sequences, and different enzymes recognize different sequences for cutting. Thus, restriction enzymes that cut *M. heterophyllum* but not natives, and vice versa, can be used to visually determine the genetic identity of plants through gel electrophoresis. We used MapDraw, version 5.08 (DNASTAR, Inc.) to construct restriction enzyme maps for all unique ITS alleles found in our study. We used the restriction maps to identify two different restriction sites: one that was present in the ITS regions of *M. heterophyllum* but not in native milfoils, and one that was present in the ITS region of all of native milfoils but was absent from *M. heterophyllum*. We then digested each ITS PCR product (*i.e.*, the same product used for DNA sequencing) separately with the two enzymes identified to cut at the above restriction sites. Restriction digests were visualized on 1% agarose gels to confirm the ability of these enzymes to effectively distinguish among *M. heterophyllum* and native milfoils.

Results

Genetic Identification of Milfoils

We found 12 different ITS alleles among our milfoil samples. One allele, “heterophyllum NH”, grouped phylogenetically with *M. heterophyllum* GenBank accessions (Fig. 1), from which it differed by only one to three nucleotide substitutions. One allele, “farwellii NH”, grouped phylogenetically with *M. farwellii* GenBank accessions (Fig. 1) and was identical to *M. farwellii* GenBank accession “farwellii WI 2”. Three alleles, “humile NH 1-3”, grouped phylogenetically with *M. humile* GenBank accessions (Fig. 1). The most common of these (humile NH 1) was identical to *M. humile* GenBank accession “humile WI 1”. The other two *M. humile* alleles differed from “humile WI 1” by 1 and 3 nucleotide substitutions (Table 1). Finally, seven unique alleles, “verticillatum NH 1-7” grouped phylogenetically with *M. verticillatum* GenBank accessions (Fig. 1). The two *M. verticillatum* samples contained multiple non-identical copies of ITS alleles within each individual. However, these different *M. verticillatum* alleles formed a monophyletic group with the *M. verticillatum* GenBank accession demonstrating that they represent species-specific alleles, and not alleles from different species as would be expected if they were hybrids (Fig. 1).

Table 1.—Species-specific nucleotide substitutions that distinguish among *M. humile*, *M. heterophyllum*, and *M. farwellii* and the number of lakes and samples where each allele was found. Numbers in the top row refer to nucleotide positions (from a 690 base pair alignment of the five alleles) where at least one allele differed from all others. Although phylogenetic resolution among these three species was low, species-specific differences in their ITS DNA sequences are evident and can be used for DNA fingerprinting. *M. verticillatum* alleles are not included because this species is different from all others at more nucleotide positions than can be shown here. A dash refers to a nucleotide deletion with respect to other sequences.

Allele	No. lakes (samples)	Nucleotide positions																
		40	77	78	85	119	120	169	213	382	437	438	439	449	453	581	595	610
<i>humile</i> NH 1	10 (11)	A	C	C	A	A	C	T	A	C	A	C	C	G	T	T	C	G
<i>humile</i> NH 2	1 (2)	A	C	C	A	A	C	T	A	T	A	C	C	A	T	T	C	G
<i>humile</i> NH 3	1 (1)	A	C	C	A	A	T	T	A	C	A	C	C	G	T	T	C	G
<i>farwellii</i> NH	7 (10)	T	-	-	G	T	C	T	C	C	-	-	C	G	T	T	C	A
<i>heterophyllum</i> NH	24 (27)	T	-	-	A	T	C	C	C	C	-	-	A	G	G	C	G	G

The phylogenetic resolution among *M. humile*, *M. farwellii*, and *M. heterophyllum* was low, but statistical support for the monophyly of each species was relatively high (>60%, data not shown). This is because species-specific nucleotide substitutions, insertions and deletions distinguish the alleles found in each of these species (Table 1). In other words, although ITS DNA sequences are not sufficient to resolve the phylogenetic relationships among these three species, they can be used for species identification.

A number of discrepancies arose between the genetically and morphologically identified milfoils (Table 2), usually involving pairs of native species. For example, nine samples (from seven different lakes) were morphologically identified as native *M. humile* but were genetically more similar to *M. farwellii*. Similarly, one sample morphologically identified as native *M. verticillatum* was genetically related to *M. farwellii*. However, we also encountered discrepancies involving native/non-native species pairs. For example, one sample identified as native *M. verticillatum* was genetically identical to non-native *M. heterophyllum*. Similarly, two samples identified as non-native *M. heterophyllum* were genetically identical to native *M. humile*.

Restriction Enzyme Identification of *M. heterophyllum*

We identified two restriction sites that distinguished ITS DNA of non-native *M. heterophyllum* from native milfoils (Fig. 2A). The first restriction site, *Fsp*I, cut the non-native ITS DNA from *M. heterophyllum* but not native milfoils; in contrast, the second restriction site, *Sac*I, cut native milfoil ITS DNA, but not *M. heterophyllum* (Fig. 2B).

Table 2.—Morphological and genetic identifications of water milfoils from the 10 New Hampshire lakes where species misidentifications occurred, as inferred by discrepancies between genetic and morphological identification. An asterisk indicates native/non-native misidentifications.

Population	Morphological Identification	Genetic Identification
*Captains Pond	<i>M. heterophyllum</i>	<i>M. humile</i>
*Long Pond	<i>M. heterophyllum</i>	<i>M. humile</i>
*Balch Lake, Molson	<i>M. verticillatum</i>	<i>M. heterophyllum</i>
Mountain Stream Reservoir	<i>M. verticillatum</i>	<i>M. farwellii</i>
Newfound Lake	<i>M. humile</i>	<i>M. farwellii</i>
Bearcamp Pond	<i>M. humile</i>	<i>M. farwellii</i>
Messer Pond	<i>M. humile</i>	<i>M. farwellii</i>
Mascoma River 5 Acre Pond	<i>M. humile</i>	<i>M. farwellii</i>
Otter Pond	<i>M. humile</i>	<i>M. farwellii</i>
Redhill Pond	<i>M. humile</i>	<i>M. farwellii</i>

Discussion

Our genetic data suggest that milfoil species are commonly misidentified when identifications are based on morphological characters alone. We found discrepancies between the morphological and genetic identifications for 31% of the samples in our survey. The majority of these discrepancies (70%) involved misidentifications between native species. Although less common, native milfoils and non-native *M. heterophyllum* were also misidentified. These misidentifications underscore the importance of having reliable methods for distinguishing among native and non-native milfoils.

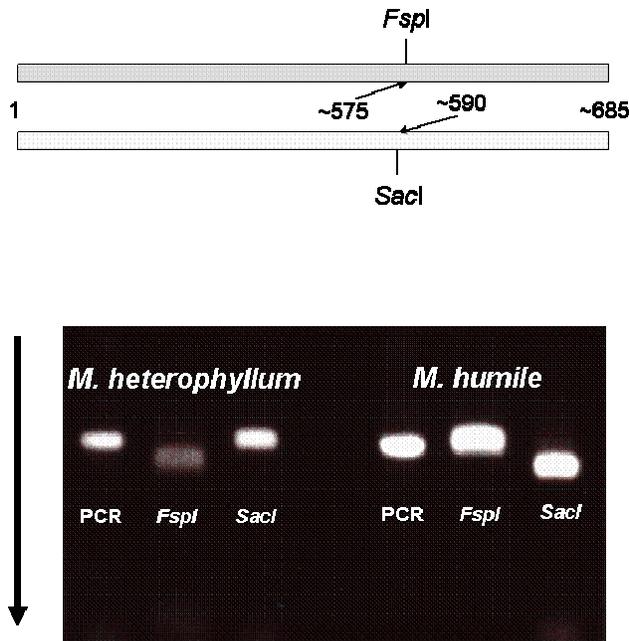


Figure 2.—Restriction enzyme map (top panel) and gel image of ITS PCR products and their restriction digests (lower panel). Top panel: upper bar represents *M. heterophyllum* ITS DNA; lower bar represents ITS DNA from native *M. farwellii*, *M. humile*, and *M. verticillatum*; numbers indicate approximate nucleotide positions. Lower panel: The three left-most samples are *M. heterophyllum* ITS PCR products, ITS PCR products digested with *FspI*, and ITS PCR products digested with *SacI*, respectively. The three right-most samples are *M. humile* ITS PCR products, ITS PCR products digested with *FspI*, and ITS PCR products digested with *SacI*, respectively. The black arrow indicates the direction of movement of the DNA through the agarose gel; smaller fragments (cut) run faster than uncut PCR products. It is clear that *M. heterophyllum* is cut by *FspI* but not *SacI*; likewise, *M. humile* is cut by *SacI* and not by *FspI*. *M. verticillatum* and *M. farwellii* restriction digests are exactly the same as *M. humile* (data not shown).

It is possible that the inconsistencies between morphologically and phylogenetically identified samples represent F_2 and backcrossed progeny with interspecific hybrid origins. In other words, mating between hybrids or backcrossing of hybrids with parental species could also produce a pattern whereby an individual plant showed the genetic signature of one parental species but the morphological characteristics of the other. However, we feel that discrepancies between morphological and genetic identifications as a result of backcrossing are less probable than simple misidentification given the morphological similarity of many milfoil species (Ceska and Ceska 1985) and the known history of misidentification of some species in herbarium records (Les and Mehrhoff 1999).

The ITS alleles that distinguish milfoil species can be used to accurately identify milfoil populations to species through DNA sequencing, but our restriction enzyme method provides

a reliable and simpler method for distinguishing non-native *M. heterophyllum* from native milfoil species. This restriction enzyme method is less costly than DNA sequencing. This method can only distinguish *M. heterophyllum* from native milfoils, however. It cannot distinguish among all native milfoil species. Restriction enzymes could be developed that distinguish among native species, but the number of enzyme combinations required would be considerably larger than the two that were necessary to accomplish the primary objective of this study. Thus, we recommend DNA sequencing of ITS when the identification of native milfoils is of primary concern and morphological characters are ambiguous or insufficient.

Although we did not find any hybrids in our study, hybrid lineages of milfoils have been found in New England (Moody and Les 2002), and some managers may be interested in whether hybrid lineages occur in particular lakes. Our restriction enzyme fingerprinting method should, in theory, identify hybrid lineages of *M. heterophyllum*. Specifically, both restriction enzymes should cut only half of the ITS PCR product because half of the PCR product came from a parental DNA sequence that is not cut by the enzyme. Thus, each restriction digest would have one larger band for the uncut ITS DNA sequence from one parental species and a smaller, cut band from the other parental species.

The use of two restriction enzymes together, one that cuts natives and one that cuts non-native *M. heterophyllum*, buffers against false positive or negative identifications of native milfoils or *M. heterophyllum*. False positives or false negatives in distinguishing native milfoils from *M. heterophyllum* could arise if an individual plant had a unique mutation resulting in the loss or formation of a restriction site cut by one of the two enzymes. Such mutations would lead to the incorrect assignment of native versus non-native milfoils for a single restriction enzyme digestion. However, the probability is small that an individual would have two mutations such that it both formed a restriction site not usually present and lost one usually present within that species.

The ability to distinguish native milfoils from non-native *M. heterophyllum* in New Hampshire lakes is critical for lake management decisions and cannot always be achieved through morphological identification of specimens alone. However, we have demonstrated that species-specific DNA sequences can be used to distinguish among milfoils and that restriction enzyme digests of ITS PCR products with *FspI* and *SacI* can accurately distinguish non-native *M. heterophyllum* from native New Hampshire milfoils. These techniques are relatively inexpensive and easy to employ and may be useful for monitoring the spread of invasive species and post-treatment monitoring.

Acknowledgments

We thank Brandy Penna for extensive help with field collections, and NH-DES for financial support.

References

- Aiken, S. 1981. A conspectus of *Myriophyllum* (Haloragaceae) in North America. *Brittonia* 33:57-69.
- Ceska, O. and A. Ceska. 1985. *Myriophyllum* (Haloragaceae) in British Columbia: problems with species identifications. P. 39-50. *In*: L. Anderson (ed.). 1st International Symposium on Water Milfoil (*Myriophyllum spicatum*) and Related Haloragaceae Species. APMS, Vicksburg, Mississippi.
- Crow, G. and C. Hellquist. 2000. Aquatic and Wetland Plants of Northeastern North America. The University of Wisconsin Press, Madison. 464 p.
- Hall, B.G. 2004. Phylogenetic trees made easy: a how-to manual, 2 edition. Sinauer, Sunderland. 221 p.
- Halstead, J.M., J. Michaud, S. Hallas-Burt and J.P. Gibbs. 2003. Hedonic analysis of effects of a nonnative invader (*Myriophyllum heterophyllum*) on New Hampshire (USA) lakefront properties. *Environ. Manage.* 32:391-398.
- Les, D.H. and L.J. Mehrhoff. 1999. Introduction of nonindigenous aquatic vascular plants in southern New England: a historical perspective. *Biol. Invasions* 1:281-300.
- Mack, R.N., D. Simberloff, W.M. Lonsdale, H. Evans, M. Clout and F.A. Bazzaz. 2000. Biotic invasions: Causes, epidemiology, global consequences, and control. *Ecol. Appl.* 10:689-710.
- Moody, M.E. and R.N. Mack. 1988. The spread of plant invasions: the importance of nascent foci. *J. Appl. Ecol.* 25:1009-1021.
- Moody, M.L. and D.H. Les. 2002. Evidence of hybridity in invasive watermilfoil (*Myriophyllum*) populations. *Proc. Natl. Acad. Sci. USA.* 99:14867-14871.
- Palumbi, S.R., A. Martin, S. Romano, W.O. McMillan, L. Stice and G. Grabowski. 1996. The Simple Fool's Guide to PCR. Kewalo Marine Laboratory and Univ. of Hawaii.
- Pimentel, D., L. Lach, R. Zuniga and D. Morrison. 2000. Environmental and economic costs of nonindigenous species in the United States. *Bioscience* 50:53-65.
- Posada, D. and K.A. Crandall. 1998. MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14:817-818.
- Ruiz, G.M., P. Fofonoff, A.H. Hines and E.D. Grosholz. 1999. Non-indigenous species as stressors in estuarine and marine communities: Assessing invasion impacts and interactions. *Limnol. Oceanogr.* 44:950-972.
- Soltis, D. and R. Kuzoff. 1995. Discordance between nuclear and chloroplast phylogenies in the *Heuchera* group (Saxifragaceae). *Evolution* 49:727-742.
- Swofford, D.L. 1998. PAUP*. Phylogenetic analysis using parsimony. Sinauer, Sunderland, MA.
- Thompson, J.D., T.J. Gibson, F. Plewniak, F. Jeanmougin and D.G. Higgins. 1997. The CLUSTAL X-windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25:4876-4882.
- Wilcove, D., D. Rothstein, J. Dubow, A. Phillips and E. Losos. 1998. Quantifying threats to imperiled species in the United States. *Bioscience* 48:607-615.