SUMMARY OF COMPLETED RESEARCH

We used a large (720 lake) dataset from the New Hampshire Department of Environmental Services (NH-DES) to evaluate whether the distribution of invasive variable-leaf watermilfoil, *Myriophyllum heteropyllum*, within the state is associated with limnological and geographic attributes. A significant component of our effort went into organizing this existing database so that it was in a form suitable for our multivariate statistical analyses. In addition, we evaluated the genetic diversity and geographic origins of invasive *M. heterophyllum* in New Hampshire with multiple genetic markers and broad geographic sampling.

The establishment and spread of invasive *M. heterophyllum* in New Hampshire is likely controlled simultaneously by restricted natural dispersal potential, stochastic human-mediated dispersal, and local environmental conditions of lakes. **Our multivariate analyses revealed that *M. heterophyllum* occurs in NH lakes that are larger, more alkaline, and more conductive (reflecting higher cation and anion concentrations) than lakes where native milfoils occur.** However, the geographic distribution of *M. heterophyllum* within the state is currently confined primarily to the Merrimack watershed despite the fact that suitable habitat is available in equal relative frequencies in all NH watersheds. Thus, its current confinement to the Merrimack watershed is most likely due to limited natural dispersal potential. While human-mediated dispersal may be the primary cause for colonization of isolated lakes and lakes in other watersheds, *M. heterophyllum* likely spreads via river connections within watersheds. Furthermore, *M. heterophyllum* continues to increase in frequency and abundance in each of the four major watersheds where it has been identified. **Given the abundance of suitable but currently unoccupied habitats, the high potential for human-mediated dispersal via boats, and natural dispersal via rivers, it seems probable that *M. heterophyllum* will continue to spread extensively both within and among watersheds in the region.** In addition, because alkalinity and conductivity are tightly coupled with the loading of inorganic nutrients and total dissolved solids, *M .heterophyllum* distribution may be influenced by human land-use practices. It is therefore also possible that invasive *M. heterophyllum* will displace native plants, including native milfoils, in lakes that are becoming increasingly alkaline and conductive due to anthropogenic activities.

Our genetic analyses revealed that the *M. heterophyllum* invasion in the northeastern US consists of different morphologically cryptic lineages with distinct geographic origins. Contrary to the notion that the NH *M. heterophyllum* invasion originated from lineages historically native to the southeastern US, **our analyses suggest that the geographic origin of the majority of New Hampshire populations instead stems from lineages that were historically native to the Midwestern US**. In contrast, we identified both Southeastern and Midwestern lineages of *M. heterophyllum* in the neighboring New England states of Connecticut and Maine. One New Hampshire population, Hill Top Pond, provided some evidence for introgression between Southeastern and Midwestern lineages. In addition, one *M. heterophyllum* individual obtained from a New Hampshire pet store was genetically similar to the Southeastern *M. heterophyllum*, and provides evidence that the aquarium trade may contribute to the spread of invasive lineages. **New Hampshire may therefore be susceptible to invasion by several genetically distinct lineages of *M. heterophyllum*.** Further geographic sampling within the state and throughout the range of *M. heterophyllum*, and further study of the habitat affinities and ecology of the genetically distinct lineages identified here, are necessary for a comprehensive understanding of this complex biological invasion.

PROJECT DETAILS

Background and Review of Objectives

The main objective, as stated in the original proposal for this research, was to identify lake attributes that influence the distribution of native and non-native (*M. heterophyllum*) milfoils in New Hampshire (NH) using a suite of multivariate statistical approaches with chemical, morphological, biological, and spatial data for NH lakes as provided by NH-DES. Specifically, we set out to address the following questions:

1) Does the spatial distribution of *M. heterophyllum* in NH lakes provide evidence for dispersal limitation?

2) Is *M. heterophyllum* distribution correlated with "local" habitat characteristics, such as chemical, hydrological, morphological, and biological attributes?

In addition to the above, which we refer to as the 'environmental analysis', we extended the genetic work of Thum and Lennon 2006 and Thum et al. 2006, to gain a more detailed perspective of the population genetics of the NH *M. heterophyllum* invasion. Specifically, we collected samples from new regions of the native and non-native range of *M. heterophyllum* and used additional molecular markers – chloroplast DNA (cpDNA) sequences and microsatellite markers – to address the following questions:

1) How genetically similar are NH populations to invasive populations outside of the state (i.e., other portions of New England)?

2) How much genetic variation occurs within and among populations of *M. heterophyllum* in NH?

3) Where have invasive NH lineage(s) originated?

Detailed Methods

Environmental analysis

As stated in the original proposal for this research, we employed a number of complementary statistical techniques to determine the factors associated with the occurrence of invasive *M. heterophyllum* and closely related native milfoil species in New Hampshire lakes. However, the precise methods employed deviated in some cases from the original proposed methods as a result of unforeseen limitations (see below). Nevertheless, the major goals of the proposal were achieved using either the same or modified methods as originally proposed.

The data for our environmental analysis were provided by NH-DES and consisted of a number of limnological and biological variables obtained from 782 lakes sampled between the years of 1976 and 2006. Many of the lakes were sampled two or more times during these years. Because the within-lake samples are not independent, we used only the most recent data from each lake. In addition, because Lake Winnepesaukee has several spatially isolated bays that are hydrologically connected, we randomly selected a single bay for use in our analyses.

The limnological variables that we included in our environmental analyses were: pH, alkalinity, color, conductivity, area, average depth, elevation, watershed area, volume, flushing rate, total Phosphorus, total Nitrogen, Phosphorus retention, chlorophyll a, and macrophyte richness excluding *M. heterophyllum* when present. We were also interested in making explicit habitat comparisons between *M. heterophyllum* and native milfoil species. Although four species are native to NH (*M. humile, M. farwelli, M. tenellum, and M. verticillatum*), the most common native milfoil species are *M. humile and M. farwelli*. In the dataset, 98 occurrences of native milfoils were identified to species using taxonomic keys whereas 12 cases were only classified as native milfoil. Of the 98 identified specimens, 91 were identified as *M. humile* whereas only three were identified as *M. farwelli* and *M. tenellum*, and only one as *M. verticillatum*. However, mistakes in identification of the native milfoils are fairly common on the basis of morphology (Thum et al. 2006). Thus, for our analyses, we simply classified milfoils as either *M. heterophyllum* or native.

Question 1: Does the spatial distribution of M. heterophyllum *imply dispersal limitation?*

A map of the locations where *M. heterophyllum* has been found in NH demonstrates clearly that the species is restricted primarily to the Merrimack River watershed in NH. Such a spatial pattern may result from dispersal limitation. Alternatively, such a pattern may result if appropriate habitat is disproportionately distributed within the Merrimack watershed. To address these alternative hypotheses, we first tested whether lakes occupied by *M. heterophyllum* were disproportionately located within the Merrimack watershed. Then, we compared the observed distribution of suitable habitat (see below) to the predicted distribution of suitable habitat on the basis of watershed size alone.

To generate the expected number of *M. heterophyllum* invasions per watershed, we multiplied the number of observed invasions by the relative proportion of lakes sampled in each of the five major watersheds in NH. We then used a χ^2 analysis to compare the observed number of invasions in each watershed to the expected number.

Similarly, to address whether suitable *M. heterophyllum* lakes were disproportionately distributed among the five major watersheds, we calculated the canonical scores of each lake from a discriminant function analysis (see below) and subsequently performed a χ^2 analysis to test whether lakes with canonical scores within the range of those scores for lakes occupied by *M. heterophyllum* occurred significantly more or less frequently in each of the five major watersheds in NH than expected by chance.

Question 2: Is M. heterophyllum *distribution significantly correlated with local limnological attributes?*

Because only six percent of all sampled NH lakes were invaded, and because 77% of the invaded locations fall within the Merrimack watershed, we could not perform the originally proposed methods for Multiple Logistic Regression and Variance Partialling. The above methods rely on statistical comparisons between 'occupied' versus 'unoccupied' lakes, and can therefore have low statistical power to detect environmental characteristics associated with occupied lakes when many suitable lakes are not currently occupied due to recent and ongoing range expansion. To circumvent this potential lack of power in our analyses with the originally proposed methods, we employed alternative methods that do not compare occupied versus unoccupied lakes, but rather quantify the environmental conditions in occupied lakes only.

First, we used a Canonical Correspondence Analysis (CCA) to determine which lake attributes were associated with the presence of *M. heterophyllum*. For this analysis, we included 11 commonly distributed species, including native milfoils, to determine similarities or differences in lakes occupied by *M. heterophyllum* relative to those occupied by common native species. We then explicitly tested for differences in lakes occupied by *M. heterophyllum* versus native milfoils using Multivariate Analysis of Variance (MANOVA).

Genetic analysis

Because the microsatellite analyses involve complicated population genetic analyses and terminology, we restrict our presentation of the methods and results to a very brief and course overview of the cpDNA sequencing. In addition, we restrict our discussion to the main findings and how they relate specifically to the NH *M. heterophyllum* invasion.

We sequenced 694 base pairs of DNA sequence from the intergenic region between trnL and trnF on the chloroplast for 130 plants using the primers B49317 and A50272 from Taberlet et al. 1991. *M. heterophyllum* plants were collected throughout its range (FL, GA, NC, SC, CT, NY, NH, ME, MA, MI, WI, OK, MN). In addition, we collected plants from *M. laxum*, which is known to hybridize with *M. heterophyllum*, as well as *M. hippuroides* and *M. pinnatum*, two species that are morphologically very similar to *M. heterophyllum*. The sequences were aligned using the ClustalW algorithm as implemented in MEGA 3.1 (Kumar et al. 2004) and then edited by eye. The haplotype network was generated in TCS 1.21 (Clement et al. 2000).

Milfoil samples were identified by other scientists on the basis of either morphology, DNA sequence data from ITS, or both.

Detailed Results

Chronological History and Spatial Configuration of the Invasion

The number of *M. heterophyllum* populations detected in New Hampshire has increased dramatically since 1976. However, these populations are not randomly distributed across the state. Rather, 77% of all *M. heterophyllum* populations are located within the Merrimack watershed. In fact, *M. heterophyllum* is found significantly more frequently in lakes within the Merrimack watershed as compared to the remaining four watersheds when the expected number of invasions within each watershed is corrected for by the percentage of the total sampled lakes that occur within each watershed (χ^2=12.7, p=0.013). However, the number of populations identified within each watershed is increasing in all four watersheds where *M. heterophyllum* has been found (Figure 1), although the proportion of lakes occupied by *M. heterophyllum* is low within each watershed (0-9%; Table 2).

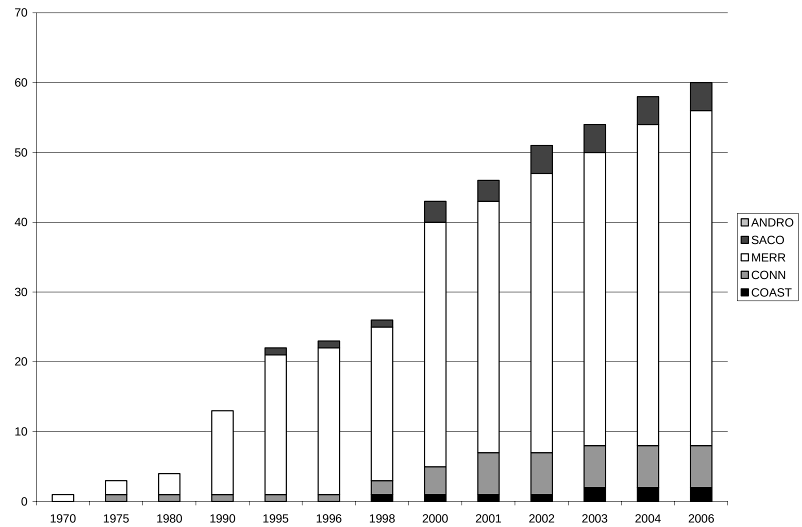

Figure 1. Number of *M. heterophyllum* populations identified through time for each of the five major watersheds in NH. Note that the years are not always consecutive along the x-axis.

Table 2. Number of lakes sampled and invaded by *M. heterophyllum* over the study period. The vast majority of invasions occur in the Merrimack watershed. However, the percentage of lakes invaded in each watershed is fairly low, ranging from 0 to 12%. Note that confirmed invasions in rivers are not included in the table below, since limnological information was not available for them for the environmental analysis. However, the qualitative results do not differ when rivers are included.

Watershed	No. lakes sampled	No. lakes invaded	% lakes invaded	% of invasions
Androscoggin	27	0	0	0
Coastal	72	2	3	5
Connecticut	208	4	2	9
Merrimack	400	34	9	77
Saco	75	4	5	9
Total	782	44	6	

Environmental Analysis

The CCA revealed a significant association between species and limnological variables (F=8.6, P=0.002). Most importantly, it demonstrated clearly that *M. heterophyllum* is associated with lake characteristics that are distinct from all other species in the analysis, including native milfoil (Figure 3). Specifically, *M. heterophyllum* occurs in lakes that are relatively large (high surface area and volume), alkaline, and conductive compared to lakes with other species. In addition to the CCA, we used MANOVA to explicitly test for differences among lakes occupied by *M. heterophyllum* versus native milfoil. The interpretation of the MANOVA is similar to the CCA, and shows that *M. heterophyllum* occupies lakes that are relatively larger, more alkaline, and more conductive that lakes occupied by native milfoils ($F_{9,137}$=15.2, p<0.0001; Table 3). We used the results from the CCA and MANOVA analyses to determine whether ‘suitable’ *M. heterophyllum* lakes occurred more frequently in the Merrimack watershed. For example, we showed above that *M. heterophyllum* lakes were significantly more common in the Merrimack watershed than the remaining four watersheds, and that *M. heterophyllum* occurs in lakes that are relatively larger, more alkaline, and more conductive than native milfoils. Thus, it is possible that the disproportionate frequency of *M. heterophyllum* occurrences in the Merrimack watershed results from relatively higher habitat availability within the Merrimack watershed than within the remaining watersheds. To explicitly test this latter hypothesis, we first calculated the canonical score for each lake in the dataset using the coefficients for the canonical root (i.e., discriminant function) from the MANOVA. We then performed a χ^2 analysis to test whether lakes with canonical scores within the range of those lakes occupied by *M. heterophyllum* occurred significantly more or less frequently in each of the five major watersheds in NH than expected by chance. The distribution of ‘suitable’ M. heterophyllum lakes, as defined by the range of canonical scores of occupied *M. heterophyllum* lakes, is not disproportionately distributed within the Merrimack watershed (χ^2= 5.1, P=.28).

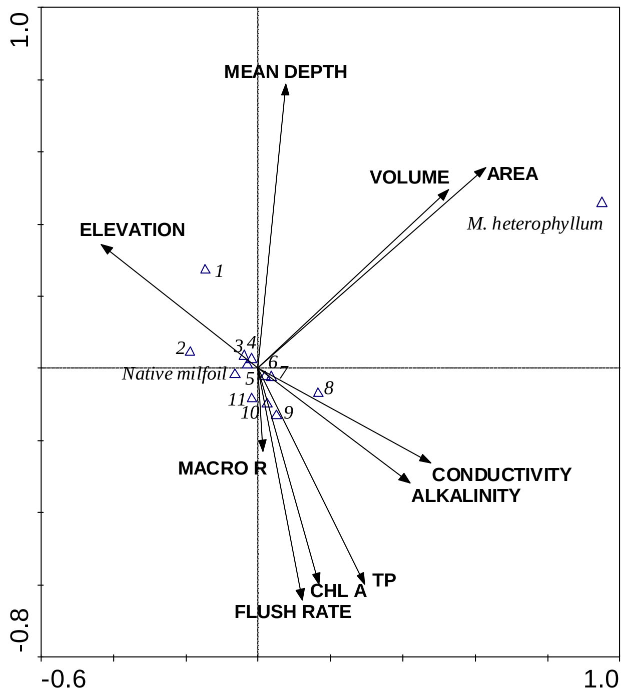

Figure 3. Biplot of species and environmental variables from CCA. Arrows represent gradients of each variable through multivariate space along the first two canonical axes. Triangles represent centroids of species data. 1=*Eriocaulon septangulare*, 2=*Nymphoides cordatum*, 3=*Eleocharis*, 4=*Juncus*, 5=*Nuphar*, 6=*Pontederia cordata*, 7=*Potamogeton*, 8=*Typha*, 9=*Sagittaria*, 10=*Brasenia schreberi*, 11=*Utricularia*. MACRO R=macrophyte species richness, CHL A=Chlorphyll a concentration, TP=Total Phosphorus concentration.

Table 3. Correlations (i.e., Factor Structures) of environmental variables with the canonical root from the MANOVA.

Variable	Correlation
Area	0.63
Conductivity	0.45
Alkalinity	0.35
Mean Depth	0.29
Total P	0.20
Volume	0.20
Chlorophyll a	0.10
Flush Rate	-0.23
Elevation	-0.50

Note: positive values indicate that *M. heterophyllum* lakes are associated with relatively larger values of the variable compared to native milfoil, whereas negative values indicate that *M. heterophyllum* lakes are associated with relatively smaller values of the variable.

Genetic Variation and Structure in the M. heterophyllum *Invasion*

There is relatively little genetic diversity among NH populations of *M. heterophyllum*. Of the 27 NH samples that we obtained that were identified as *M. heterophyllum*, 24 of them carried the same cpDNA haplotype, as indicated in Figure 2. This same haplotype was also found in at least some *M. heterophyllum* samples from CT, MA, ME, MI, NY, OR, RI, WA, and WI, as well as one individual identified as M. hippuroides from WA. Of the three individuals that differed from the above samples, two individuals were collected from Hill Top Pond, and the third individual was obtained from a pet store in NH. These latter samples shared the same DNA sequence as that found in one *M. heterophyllum* sample from GA and one *M. heterophyllum/M. laxum* hybrid found in CT.

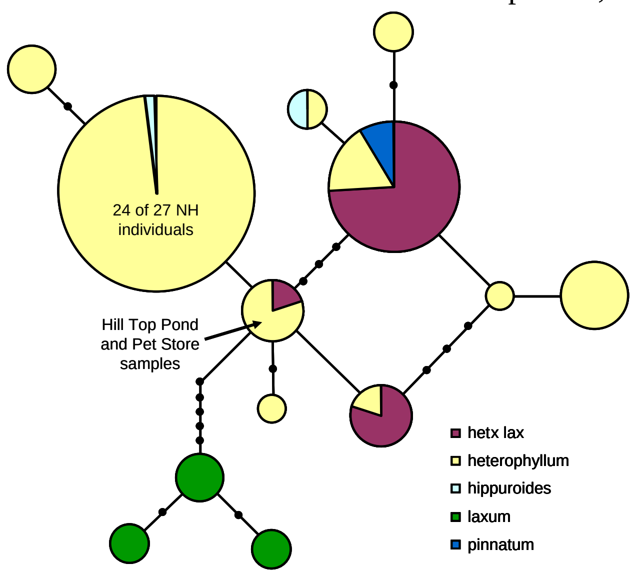

Figure 2. Haplotype network of cpDNA sequences from *M. heterophyllum* and closely related species. Circles represent distinct DNA sequences, or haplotypes. Haplotypes are connected by branches that represent evolutionary 'events' – a change in the DNA sequence through time from a common ancestor in the form of a nucleotide substitution, insertion, or deletion. Small, black circles indicate 'missing' haplotypes – evolutionary events that are inferred to have occurred at some point in history, but were not actually found in any of the samples. For example, two haplotypes separated by a small, closed circle are two evolutionary events different from one another, whereas two haplotypes separated by only a single branch differ by only one evolutionary event. Each species is given a distinct color in the haplotype network, and a distinct color was given to specimens identified as *M. heterophyllum/M. laxum* hybrids. The area of each circle is proportional to the number of individuals that carried that haplotype, and the taxonomic composition of each haplotype is indicated by the relative proportion of each shaded region in each circle.

Conclusions & Recommendations

Our environmental analysis of the *M. heterophyllum* invasion in NH suggests that the current distribution of the species within the state is likely explained by a combination of dispersal and environmental limitation. Specifically, *M. heterophyllum* occurs in lakes that are relatively large, more alkaline, and more conductive relative to unoccupied lakes and lakes occupied by native milfoils. However, *M. heterophyllum* populations are disproportionately distributed within the Merrimack watershed despite the presence of suitable lakes in the other four major watersheds, indicating that limited natural dispersal potential plays an important role in limiting its current distribution. Given the likely spread of the species via human-mediated activities, such as boating, and the availability of suitable but currently unoccupied habitat throughout the state, we predict that *M. heterophyllum* will continue to spread within the state.

Our environmental analysis provides a solid foundation for developing specific and testable hypotheses for further investigation into the causes of its spread, and should be complemented in the future with experimental studies. For example, the high correlation between *M. heterophyllum* presence and relatively higher alkalinities and conductivities suggests that physiological studies of *M. heterophyllum* across gradients in these variables should be conducted. In addition, our statistical analyses should be expanded geographically to determine whether alkalinity and conductivity are strong predictors of *M. heterophyllum* presence in other portions of its native and introduced range. Finally, further analysis of landscape-level changes in lake alkalinity and conductivity, and whether and how these variables are related to anthropogenic activities, should be conducted to determine the extent to which land-use practices may facilitate the spread of *M. heterophyllum*.

Our genetic analyses revealed important genetic variation among invasive *M. heterophyllum* populations in New England. While *M. heterophyllum* is genetically diverse across its range, the *M. heterophyllum* populations in NH are genetically uniform, and the majority of NH populations of *M. heterophyllum* may derive from as little as a single individual. However, contrary to the notion that the NH *M. heterophyllum* invasion originated from lineages historically native to the southeastern US, our analyses suggest that the geographic origin of the majority of New Hampshire populations instead stems from lineages that were historically native to the Midwestern US. In contrast, we identified both Southeastern and Midwestern lineages of *M. heterophyllum* in the neighboring New England states of Connecticut and Maine. One New Hampshire population, Hill Top Pond, provided some evidence for introgression between Southeastern and Midwestern lineages. In addition, one *M. heterophyllum* individual obtained from a New Hampshire pet store was genetically similar to the Southeastern *M. heterophyllum*, and provides evidence that the aquarium trade may contribute to the spread of invasive lineages. New Hampshire may therefore be susceptible to invasion by several genetically distinct lineages of *M. heterophyllum*.

Nothing is currently known about how ecologically different the morphologically cryptic but genetically distinct *M. heterophyllum* lineages are, or whether different lakes are susceptible to invasion by the distinct lineages. Our environmental analysis begins to address this question by quantifying the environmental characteristics associated with *M. heterophyllum* invasion in NH. Further geographic sampling within the state and throughout the range of *M. heterophyllum*, and further study of the habitat affinities and ecology of the genetically distinct lineages identified here, are necessary for a comprehensive understanding of this complex biological invasion.

REFERENCES

Clement, M., D. Posada, and K. A. Crandall. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9.

Kumar, S., K. Tamura, and M. Nei. 2004. *MEGA3*: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Briefings in Bioinformatics 5: 150-163.

Taberlet, P., L. Gielly, G. Pautou, and J. Bouvet. 1991. Universal primers for amplification of 3 noncoding regions of chloroplast DNA. Plant Molecular Biology 17: 1105-1109.

Thum, R. A., and J. T. Lennon. 2006. Is hybridization responsible for invasive growth of non-indigenous water-milfoils? Biological Invasions 8: 1061-1066.

Thum, R. A., J. T. Lennon, J. Connor, and A. P. Smagula. 2006. A DNA fingerprinting approach for distinguishing native and non-native milfoils. Lake and Reservoir Management 22: 1-6.