Introduction & Background

The invasive water-milfoil *Myriophyllum heterophyllum* is a major economic and ecological pest throughout lakes of New England. The majority of management studies have focused on the removal and prevention of milfoil. However, little is known about the population genetics of the invasion. Population genetic studies have the potential to identify the number of lineages that have invaded, as well as the geographic locations from which the invasive lineages are derived. The identification of such lineages can then be used in comparative ecological and genetic studies.

The necessary first step in any population genetic study is the development of appropriate genetic markers for the questions being asked because different classes of molecular markers vary in their rates of molecular evolution. For example, nuclear DNA sequences from the internal transcribed spacers (ITS) of the ribosomal DNA complex have been used to distinguish among native versus invasive water-milfoils. However, ITS does not show any significant pattern of variation within *M. heterophyllum*, so this marker is not appropriate for questions about patterns of genetic variation within and among *M. heterophyllum* populations.

We are developing microsatellite markers for use in population genetic studies of *M. heterophyllum*. Microsatellites are well-suited to population level studies because they are highly variable. Microsatellites are simple sequence repeats of the four nitrogenous bases that compose DNA (A, G, C, T). For example, a dimeric repeat (repeat of a two base pair sequence) might contain the sequence CA repeated ten times, yielding the microsatellite CACACACACACACACACACA [denoted $(CA)_{10}$; Figure 1]. Microsatellites are highly variable in size because repetitive regions of DNA are extremely error prone during DNA replication because of a process known as strand slippage. If the original microsatellite was $(CA)_{10}$, a mutation during DNA replication caused by strand slippage might produce the microsatellite $(CA)_9$, or $(CA)_{11}$. For a given locus then, populations may vary in their distributions of repeats for a given microsatellite. Because they differ in size, different microsatellite alleles at a given locus can be easily distinguished using gel electrophoresis. For example, a $(CA)_{10}$ allele will run faster than a $(CA)_{11}$ allele on a gel because the $(CA)_{10}$ allele is smaller. By genotyping multiple microsatellite loci for multiple individuals within and among populations, we can identify unique lineages and reconstruct patterns of ancestry among populations.

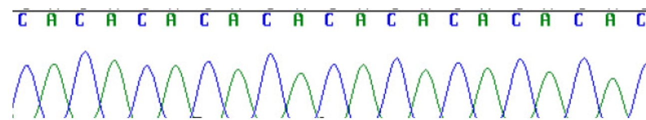

Figure 1. Electropherogram of microsatellite DNA sequence. The sequence shows a dimeric sequence (CA) repeated 10 times.

Here, I briefly describe the steps for the library construction and sequencing of the *M. heterophyllum* microsatellite library. Then, I briefly describe the primer design and genotyping steps that we are currently working on. I end the report with a summary of my educational experiences that resulted from my NEAPMS scholarship.

Microsatellite Library Development

The microsatellite library was constructed using a protocol developed by Steve Bogdanowicz at Cornell University. Steve Bogdanowicz and Ryan Thum assisted greatly in the library construction. To construct the microsatellite library, DNA was first extracted from a single milfoil sample (kindly provided by Robert Johnson). The extracted DNA was then digested (cut into small pieces) using restriction enzymes. Microsatellite DNA probes were then hybridized to the restricted DNA to enrich for fragments that contained a microsatellite. DNA fragments that did not hybridize with the probes were simply washed away and discarded. Next, the microsatellite-enriched population of DNA molecules was inserted into plasmids, which were then transformed (injected) into individual *E. coli* cells. The *E. coli* cells were then plated onto luria agar plates and allowed to incubate overnight. Each colony that grew on the plate represented a clonal population of an *E.coli* cell that took up a single piece of DNA which, in theory, contained a microsatellite (Figure 2). By cloning, we were therefore able to separate out hundreds to thousands of unique DNA fragments that could contain a microsatellite. We further enrich for microsatellite fragments by hybridizing radioactively-labeled microsatellite probes to the bacterial colonies. In this way, we could determine which *E. coli* colonies contained milfoil microsatellite DNA by exposing the radioactively-labeled plates to autoradiographic film. Clones that 'lit up' during the exposure were then picked for DNA sequencing.

DNA Sequencing of Microsatellite Clones

I picked several hundred positive colonies for DNA sequencing. For each colony, I used the polymerase chain reaction (PCR) to amplify the cloned DNA fragment. PCR products were then sequenced to determine the presence of a microsatellite. DNA sequencing was performed using BidDye cycle sequencing chemistry (Applied Biosystems) and sequenced on an ABI3100 at Cornell University's Evolutionary Genetics Core Facility (EGCF). In total, I sequenced nearly 300 clones.

Each clone was carefully inspected to determine if it contained a microsatellite. Greater than 60% of the sequenced clones contained microsatellites. This percentage represents the minimum percentage of clones with microsatellites because I sequenced in only one direction. Long fragments have not been fully sequenced yet (i.e., they need to be sequenced in the reverse direction), and complete sequences should reveal microsatellites in at least some of these long sequences.

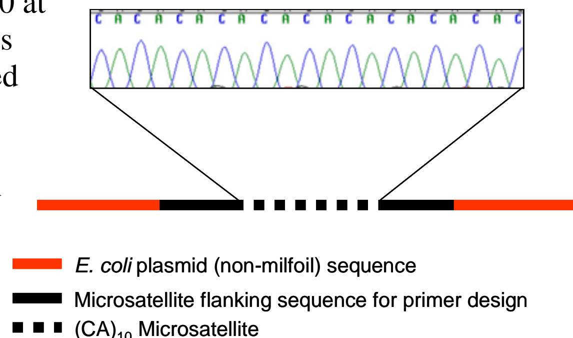

Figure 2. Shematic of a microsatellite clone. A piece of milfoil DNA enriched for microsatellites (black line) is inserted into a bacterial plasmid. After sequencing, the plasmid sequence is trimmed off and primers are made in the milfoil DNA flanking the microsatellite.

Future Work: Primer Design & Genotyping

I carefully inspected each sequenced clone containing a microsatellite to determine whether there was enough flanking sequence to design primers (Figure 2). For example, in some clones, the microsatellite DNA sequence is located too close to the plasmid boundary. Primers cannot be developed for these sequences.

I am currently designing PCR primers in the flanking regions of the microsatellites for each clone that has sufficient flanking region to design primers. Generally, at least 50 base pairs of non-microsatellite DNA are required on either side of the microsatellite for successful primer design. After designing primers for a given clone, I will begin optimizing those primers for PCR. Once the primers are optimized, they will be tested on several *M. heterophyllum* individuals to determine whether they consistently amplify the microsatellite region. For example, individuals may differ not only in the microsatellite alleles, but in the DNA sequences flanking those alleles. If the flanking regions differ enough among individuals, then the PCR primers may not consistently work. After confirming that the PCR primers consistently work across individuals, I will begin genotyping individuals to determine whether any informative genetic variation exists for each microsatellite locus. We can then begin to answer the questions outlined in the introduction.

Educational Experience

Working on this project has exposed me to many issues related to invasion biology in general, and to the problem of milfoil invasions in particular. Prior to this research experience, I had no exposure to these issues. As a result, I have become aware of the concerns and challenges associated with the management of invasive species. I also learned about the unique contributions that genetic investigations can make to studies of invasive species. Intellectually it was very exciting to see how evolutionary tools, such as population genetic studies, could be brought to bear on the management and study of invasive water-milfoils.

My participation in this project has contributed greatly to my growth and education as a developing molecular biologist. The basic techniques that I employed on a daily basis, such as PCR and DNA sequencing, will be invaluable in my future molecular biological work. In additioin, this experience has provided me with a microsatellite-specific skill-set that will likely prove useful in my future research. Microsatellites are commonly employed to answer many of the evolutionary questions that are of general interest to me. My experience working with microsatellites on this project will prove advantageous in the likely event that I must develop microsatellites for another organism.

Although I will be graduating from Cornell at the end of this summer, I look forward to participating further in this research. In particular, we plan to finish optimizing microsatellite loci and publish a paper in the journal *Molecular Ecology Notes* before the end of this summer.