# An investigation of genetic divergence among invasive crayfish (*Orconectes virilis*) populations using microsatellites

A Major Qualifying Project Report

submitted to the Faculty

of the

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree of Bachelor of Science

by

Nicholas Kohlstrom

Date: April 24, 2008

Approved:

Professor Lauren Mathews, Major Advisor

Professor Jill Rulfs, Major Advisor

Professor Michael Buckholt, Co- Advisor

## Abstract

Species are usually geographically distributed further than an individual can disperse, thus populations are often genetically differentiated through isolation by distance and barriers. Crayfish have been shown to have low dispersal patterns, possibly due to their biology and environment. Using four microsatellite loci as molecular markers, I studied the degree of genetic differentiation of three populations of *Orconectes virilis* in Massachusetts. I hypothesized the populations would show some differentiation. The data imply that the populations maybe genetically differentiated from one another, but additional data are needed for a stronger inference.

## Acknowledgements

I would like to thank my advisors Prof. Lauren Mathews and Prof. Michael Buckholt for their time, patience and guidance, in addition to Prof. Jill Rulfs for overseeing this endeavor. I would also like to thank Tahiyyah Muhammad for all the work we did together.

# Table of Contents

# 1 Introduction

In trying to understand the diversity of life, evolutionary biologists have long been studying the phenotypic traits of organisms. With the recent advances in molecular biology, it is now possible to study the ecology of natural populations through genetic data. This field of research is known as molecular ecology and is a rapidly expanding branch of biology. Molecular ecology uses a wide variety of techniques and molecular tools such as microsatellites, DNA sequencing, cDNA libraries and PCR (polymerase chain reactions) among others to study ecological questions about gene flow, population dynamics, behavior, reproduction, evolution and biodiversity among other questions.

In this study I probe the population dynamics of an invasive species of crayfish, *Orconectes virilis*, using a molecular marker known as a microsatellite, to determine the amount of differentiation between populations that are isolated by physical barriers and distance. In this section, I will first present some general background information on crayfish biology and ecology, then I will discuss some background topics in population genetics, including the theoretical and empirical literature relevant to the phylogeography of crayfish and ecologically similar organisms, finally I will review the use of microsatellites as a molecular marker.

### *1.1 Crayfish: Orconectes virilis*

There are >540 species of freshwater crayfish, with 70 to 77% of the freshwater crayfish diversity occurring in North America (Halliburton 2004; Holdich, 2004). Crayfish are decapod crustaceans that are divided into three families: Astacidae and Cambaridae in the Northern Hemispshere, and Parastacidae in the Southern Hemisphere (Holdich, 2004; Harm, 2002). As seen Figure 1, the main distributions of Astacidae are located in Europe and Western U.S., Cambaridae in most of North America, and Parastacidae in Australia, Madagascar and parts of South America (Holdich, 2004). In the family Cambaridae there are three prominent genera: *Procamburus, Cambarus*, and *Orconectes*. The genus *Orconectes* is comprised of 11 subgenera, 81 species, and 11 subspecies, of which *Orconectes virilis* is one of the species (Halliburton 2004; Holdich, 2004; Harm, 2002).

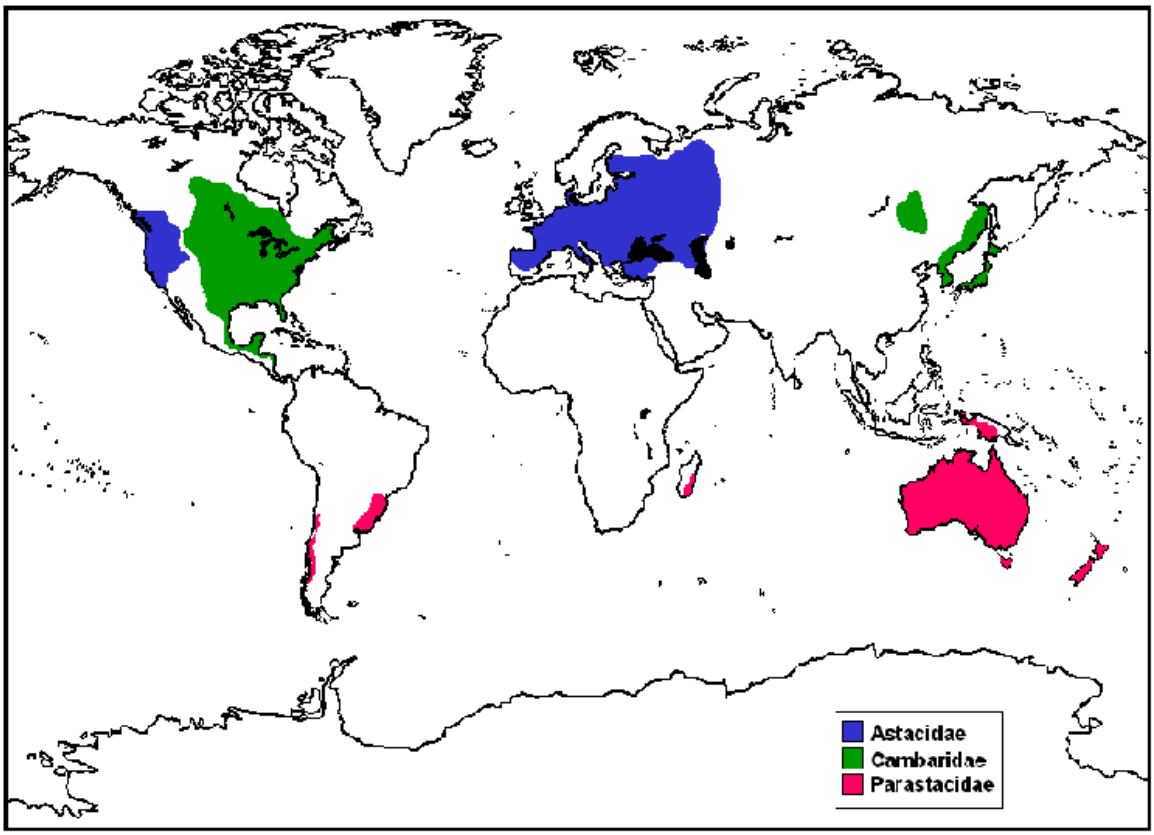


**Figure 1: Family distribution of freshwater crayfish (Carnegie Museum 2006).**

*O. virilis*, also commonly called the northern crayfish or the virile crayfish, grows on average to 10-12 cm in length excluding its antennae and large chelipeds (Figure 2). The color of the body and abdomen are brownish-red dappled with dark brown spots. The chelae, or the major component of the large chelipeds, are wide, flattened and posses a straight dactyl margin. The chelae and legs have a bluish tint with yellow tubercles (Harm, 2002).

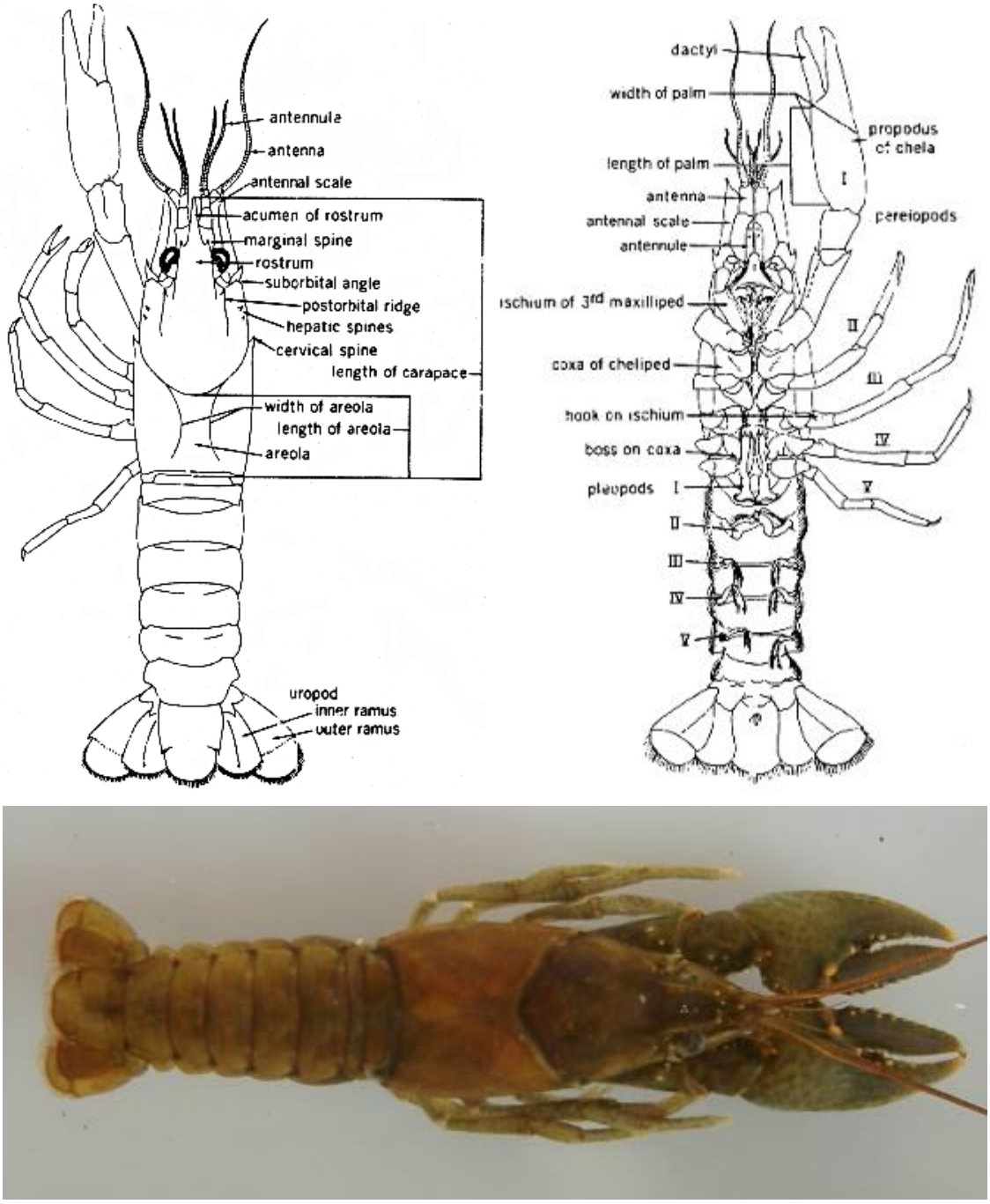


**Figure 2: Dorsal and ventral views of general male crayfish (Hobbs, 1974) and a dorsal view of *O. virilis.***

*O. virilis* has a wide natural range from Alberta to Quebec, Canada, throughout more than half of the United States from Texas to Maine, and Chihuahua, Mexico (Harm, 2002). But in Massachusetts, *O. virilis* is believed to be an invasive species (Hobbs, 1989), and is listed as such by the Global Invasive Species Database

(http://www.issg.org/database/). Its habitats include rivers, streams, lakes, ponds, and marshes. Most often *O. virilis* can be found in bodies of water that contain rocky bottoms, but they can also be found in mud, silt, sand or weedy bottomed rivers and lakes (Harm, 2002). *O. virilis* populations generally live under excavated rocks, logs, and thick vegetation but some populations are known to construct burrows for shelter and to hide from predators in river banks; one population was even observed constructing chimneys at the entrance of the burrows (Harm, 2002). Individuals are able to survive the harsh winters in their northern range by migrating to deeper water that will not freeze (Aiken, 1968). *O. virilis* even holds the record for being the northern most crayfish species found in North America (Harm, 2002). They are most active from May to September due to the warmer water temperatures and their trapability is influenced by bait types, water temperatures and conditions and the presence of predatory fish (Somers *et al*, 1986; Richards *et al*, 1996; Harm, 2002). Most collection methods are sex biased due to sex related differences in behavior and physiology which correlates to reproductive state and molting. Males may be as much as four times more likely to be caught than females (Momot *et al*, 1972; Somers *et al*, 1986). Population densities have been reported as ranging from <0.01 to 8.75 crayfish $m^{-2}$. *O. virilis* are omnivorous and feed on both living and dead vegetation and animals, including fish, snails, tadpoles, leaf litter and insects. They also have been shown to have a significant impact on biomass (Harm, 2002; Nystrom, 2002).

Northern populations of *O. virilis* are reported to mate from July to September, but may also mate during spring depending on seasonal conditions, while more southern populations mate both in the spring and fall. Sperm is stored in the annulus ventralis (the sperm receptacle in the female crayfish), and it can be held over the winter. There may be geographic variation in fecundity, with females in northern populations laying about 20 to 320 eggs from late May to early June, and females in southern populations laying up to 490 eggs starting as early as the beginning of March. Eggs are attached to the underside of the female's abdomen until they hatch. The hatched eggs are called juveniles; these resemble miniature crayfish, and will stay attached to the female until they mature enough to leave. This type of development, from egg to juvenile attached to the mother, is known as direct development. Most other crustaceans undergo indirect development

where hatched eggs enter a free-living nauplius or larval stage. Since they stay attached to their mothers until they become mature enough to leave, crayfish have some interesting patterns of dispersal (Scholtz *et al*, 2002; Reynolds, 2002). Juveniles molt about 5 times their first summer and around 3 or 4 times the next. Crayfish reach sexual maturity by their first or second year. Mature males are cyclically dimorphic and molt in mid-June from Form II (a nonbreeding form) to Form I (a breeding form) and back again to Form II in August. *O. virilis* are also sexually dimorphic, with Form I males usually possessing larger claws than females. The average life span for *O. virilis* is 3 to 4 years (Holdich, 2004; Harm, 2002; Reynolds, 2002).

### *1.2 Population Genetics*

Population genetics is the study of factors that affect the evolution and genetic composition within and among populations. A population is a localized group of members of the same species that can potentially mate with one another (Halliburton, 2004). The four major evolutionary forces that influence evolution and genetic composition in populations are natural selection, mutation, genetic drift, and gene flow (Halliburton, 2004). The majority, if not all populations of species show some degree of genetic structuring due to evolutionary forces acting on populations with different environmental barriers, historical processes and life histories (Balloux *et al*, 2002).

Understanding gene flow and its effects is useful in many fields of research, such as population genetics, population ecology and conservation biology. Gene flow, which is also known as gene migration, is the transfer of alleles from one population to another. When genes are exchanged between populations the allele frequencies between populations become homogenous. A high rate of gene flow will prevent the fixation of alleles that may have been favorable in a specific population due to adaptation to the local environment. This will inhibit the process the speciation. But gene flow can also introduce new/novel alleles into a population or generate new combination of alleles and prevent random genetic drift. Genetic drift is a random variation of allele frequencies due to random sampling of gametes and other spontaneous events. The larger the population size, the smaller amount of genetic drift is displayed, and vice versa because the magnitude of genetic drift is inversely proportional to population size. The long-term

effect of genetic drift decreases the amount of variation within a population and causes populations to diverge or separate. It eventually causes a decrease in heterozygosity and loss of alleles. Two examples of genetic drift are displayed in population bottleneck and founders effect. A population bottleneck is a suddenly reduction in population size, possibly due to a natural disasters like fires or floods, which deletes or reduce an allele or alleles in that population. Founder's effect is the emergence of a new population established by a few individuals from an original population, which carries the same alleles or a small variety due to the number of individuals. Both examples create a narrowed sample of allele frequencies, which should exhibit less variation (Halliburton, 2004).

Phylogeography is the study of the relationship between genetic variation and the geographic distribution of a species or closely related species. One interest of this field is determining the effects that distance and geographic barriers have on gene flow, which can be associated with historical biogeographic events, migration patterns, life cycles, or other factors (Avise *et al*, 1987; Turner *et al*, 1996). Organisms that reside only in rivers and other freshwater bodies have interesting patterns of gene flow since they are separated linearly by distance and watershed. Given that species are usually geographically distributed further than an individual can disperse, populations are often genetically differentiated through isolation by distance. This would mean populations that are closer together would be genetically more similar than populations that are more distant (Balloux *et al*, 2002). This is the same for organisms that are in the same watershed (excluding other factors such as currents, water conditions, dams, waterfalls, etc.). But organisms, such as fish and many types of crayfish, which are restricted to aquatic environments, are inhibited from traveling between watersheds. Even if two populations in separate rivers are close to each other geographically does not mean they are close in linear river distance, since the rivers may connect a great distance downstream (Fetzner *et al*, 2003). All these factors result in the creation of subpopulations. Population subdivision occurs when a large population divides into smaller subpopulations due to partial or full geographical isolation. This results in individuals being more likely to mate with another within the same subpopulation versus another subpopulation. This lack of gene flow in conjunction with mutations and

evolutionary pressures will result in the genetic divergence between the populations (Halliburton, 2004). Differing rates of genetic divergence among taxa are usually attributed to differences in life-history patterns, dispersal mechanisms and migration (Turner *et al*, 1996).

Phylogeography is commonly studied on a larger scale, such as the distribution of species across a continent, particularly in conjunction with major historical events such as the Pleistocene ice ages which reshaped the water way in both northern North America and Europe. For example, in a study of freshwater fish from Scandinavia, the fish were found to be genetically differentiated from one another in patterns that correlate with historic glacial activity and the barriers they created (Mäkinen *et al*, 2007).

On a smaller geographic scale, a study on the phylogeography of the shovel-nosed salamander in the southern Appalachians was performed to observe the differentiation between subpopulations within a river basin and from populations in an adjacent river basin separated by the Eastern Continental Divide (where rivers separated into different basins), which was believed to last be connected about 23.8-1.8 million years ago. This species, like crayfish, lives an entirely aquatic existence. It was found that there was a high degree of differentiation across the Eastern Continental Divide, but a lower degree of differentiation between subpopulations on either side of the divide (Jones *et al*, 2006).

Population genetic studies, in Australia, on the Australian freshwater crayfish, *Cherax destructor*, have found that there is population differentiation between watersheds, particularly between northern and southern watersheds (Hughes *et al*, 2003; Nguyen *et al*, 2005). But there is still some debate over differentiation within watersheds. Hughes and Hiller (2003) found that in the northern watershed there was almost no differentiation between populations in connected rivers, which they attributed to a high amount of connectivity and dispersal. Nguyen *et al*, (2005) found that gene flow was constricted in multiple contiguous watersheds, which resulted in the differentiation of populations. They attributed this to behavioral or life history features restricting dispersal.

Fetzner and Crandall (2003) analyzed the genetic variation of golden crayfish (*Orconectes luteus*) populations from the Ozarks region of Missouri. They found high levels of divergence among populations corresponding to watershed fragmentation created by Pleistocene glacial events. They also found general patterns of within

watershed haplotype uniformity, where each watershed appeared to contain a unique haplotype or set of haplotypes. This suggests that the transfer of individuals between watersheds is a rare event. It was also found that one population was more closely related to another population in another watershed, which was attributed to river capture. River capture is where a body of water will connect or change its flow from one watershed or river to another neighboring one. This connects or transfers populations or individuals that were once separate together (Burridge *et al*, 2007).

Buhay and Crandall (2005) studied the phylogeography of subterranean (cave) freshwater crayfish from the southeastern United States, and compared the data to common local surface dwelling crayfish. Cave crayfish were generally hypothesized to be even more isolated than surface species and should thus show even lower rate of gene flow, lower genetic diversity and more differentiation among populations. Contrary to the common hypothesis, cave crayfish were found to have high levels of genetic diversity and gene flow. When compared to the surface crayfish, it was found that the surface crayfish studied were showing a decline in genetic variability.

### *1.3 Microsatellite Loci as population genetic markers*

In early population genetic investigations, individuals and populations of species were assessed genetically by observing morphologies of individuals and making inferences about their genotypes. This method can be highly unreliable, because some morphologies may be conserved between species, or convergent evolution may result in species that appear the same. In addition, phenotypes maybe modified due to behavior or environmental factors. With the advent of molecular markers that use DNA, RNA and proteins from both nuclear and cytoplasmic genomes, which are fixed in individuals and cannot be effected by the environment, all the problem researchers had using morphologies to study populations can be bypassed. Most eukaryotic species are diploid and contain about equal amounts of nuclear DNA from both parents (except where sex chromosomes apply), as opposed to mitochondrial DNA, which is inherited only from the female parent and is therefore haploid. The rate of evolution of these DNA sequences can vary from locus to locus throughout the genome. As a result, genomes offer a wide range of tools to address many questions in ecology and evolution.

There are a wide array of molecular techniques and markers that can be used to answer specific questions with different advantages/limitations and expenses. In general there are two classes of molecular markers available to analyze nuclear DNA; these are dominant and codominant markers. Dominant markers, for example Random Amplified Polymorphic DNAs (RAPDs) and Amplified Fragment Length Polymorphisms (AFLPs), allow analysis of multiple loci and allow direct differentiation between homozygous and heterozygous individuals, hence, investigators can estimate expected heterozygosity assuming Hardy-Weinberg equilibrium (HWE), but, with no data for observed heterozygosity, cannot determine if the population conforms to HWE or not. Codominant markers, for example allozymes and microsatellites, analyze one locus at a time and allow the detection of both the heterozygous and homozygous condition of an individual (Nguyen *et al*., 2006).

Microsatellites, also known as simple sequence repeats (SSR), short tandem repeats (STR), or variable number tandem repeats (VNTR), are specific sequences of nuclear DNA, which usually contain a variable number of nucleotide tandem repeats of 1-6 bases. An example of different numbers of repeats can be seen in Table 1. Microsatellites are found in the genomes of both eukaryotes and prokaryotes (Goldstein *et a*., 1995; Guillermo *et al*, 1994; Jin *et al*, 1993). Microsatellites have a high mutation rate with mutations happening on the order of $10^{-5}$ - $10^{-2}$ mutations per generation (Schlötterer, 2001; Weber *et a*., 1993). These mutations alter microsatellite lengths through the addition or deletion of repeat units, which makes microsatellites highly polymorphic in natural populations, with an average expected heterozygosity above 50% (Queller *et al*, 1993; DiRienzo *et al*, 1994).

**Table 1: Example of microsatellites with different types/numbers of tandem repeats.**

| Repeat type | 5' flanking region microsatellite locus 3'flanking region | Notation |
|---|---|---|
| Mono | GCCTTGCATCCTTAAAAAAAAAAAATCGGTACTAC | $(A)_{11}$ |
| Di | GCCTTGCATCCTTCACACACACACACATCGGTACTAC | $(CA)_6$ |
| Tri | GCCTTGCATCCTTCTGCTGCTGCTGTCGGTACTAC | $(CTG)_4$ |
| Tetra | GCCTTGCATCCTTACTCACTCACTCACTCTCGGTACTAC | $(ACTC)_4$ |

Microsatellites are normally found dispersed throughout the non-coding regions of the genome. About 3% of the human genome is made up of microsatellites; of that about 90% is found in non-coding regions (Subramanian, 2003). Most microsatellite DNA is considered "junk" DNA. Variations are mostly neutral and produce no measureable effect on phenotype because they are located in non-coding regions away from exons. But when they are present in or near coding regions, they sometimes have an effect on phenotype. Microsatellites have been shown to provide a source of genetic variation and help regulate gene expression and protein function. When they are found in coding regions, microsatellites mutations are known to cause diseases like Fragile X and Huntington's disease. However, most microsatellites in coding regions are trinucleotide repeats which helps prevent frameshifts and diseases (Zhong, 1995; Rubinsztein *et al*, 1999).

Microsatellites are useful as genetic markers because of some key advantages. They are highly polymorphic compared to other sequences of DNA, locus specific (as opposed to randomly amplified markers like AFLPs), codominant, polymerase chain reaction (PCR) based and can be used to gather data between individuals and groups. Since they are codominant, by studying the differences in microsatellite lengths, researchers can make inferences about when two populations diverged and about the gene flow within and among populations. These data can then be used for phylogeographic studies (Selkoe *et al*, 2006; Yazdani, 2003).

To obtain data from a microsatellite, it must first be targeted and amplified from the organism's genome. This is done by using primers that flank the repeat sequence, which can then be used in a PCR (Erlich et. al, 1991) to multiply the microsatellite locus for easy analysis. There are few data on microsatellites for many organisms and there are very few primers available, if any, for some species, which makes identifying and characterizing specific loci the first step for many researchers. Primers that were developed to amplify microsatellite loci for one species can often be used with another closely related species for the same locus, but the success of annealing and amplifying decreases with increased genetic distance due to point mutations in flanking regions. Failure of the primers to anneal and amplify due to these point mutations will result in null alleles (Zane *et al*, 2002).

Microsatellite fragment sizes after PCR amplification can be determined in multiple ways. One way is to run the PCR product in a gel using electrophoresis. The resulting bands on the gel correspond to different sized microsatellite fragments. Fragment sizes can also be determined by analysis in an automated DNA sequencer. This device requires one of the primers, either forward or reverse, to be labeled with a fluorescent tag. The PCR products and a sized standard labeled with a different fluorescence are then added to a capillary tube which separates the fragment sizes just like the gel electrophoresis. As the fragments run through the capillary tubes, a laser records its fluorescence and, by comparing it to the size standard, the fragment sizes can be determined. Another benefit of using the automated DNA sequencer is the ability to multiplex the samples. By fluorescently labeling primers for each different microsatellite locus with a different fluorescence, multiple loci can be run at a single time (Toh *et al*, 1996).

Microsatellites are thought to mutate through one of two proposed mechanisms. The first one is called polymerase slippage (also known as slipped-strand mispairing). In polymerase slippage, shown in Figure 3, when the microsatellite locus is being replicated, the polymerase loses track of its place and temporarily dissociates from the DNA strand. The terminal end of the newly synthesized DNA separates and anneals with another section of DNA, forming a small loop. After DNA synthesis is completed, the mismatch repair system corrects the loops by adding or deleting repeat units, but it may also correct the loop. If this loop forms on the nascent strand the microsatellite will increase in length, but if the loop forms on the template strand the microsatellite will decrease in length. If polymerase slippage occurs during PCR amplification, particularly during the early cycles, small peaks known as stutter bands will form due to the amplification of different sized fragments. These stutter bands can sometimes be mistaken for actual alleles (Voet *et al*, 2006).

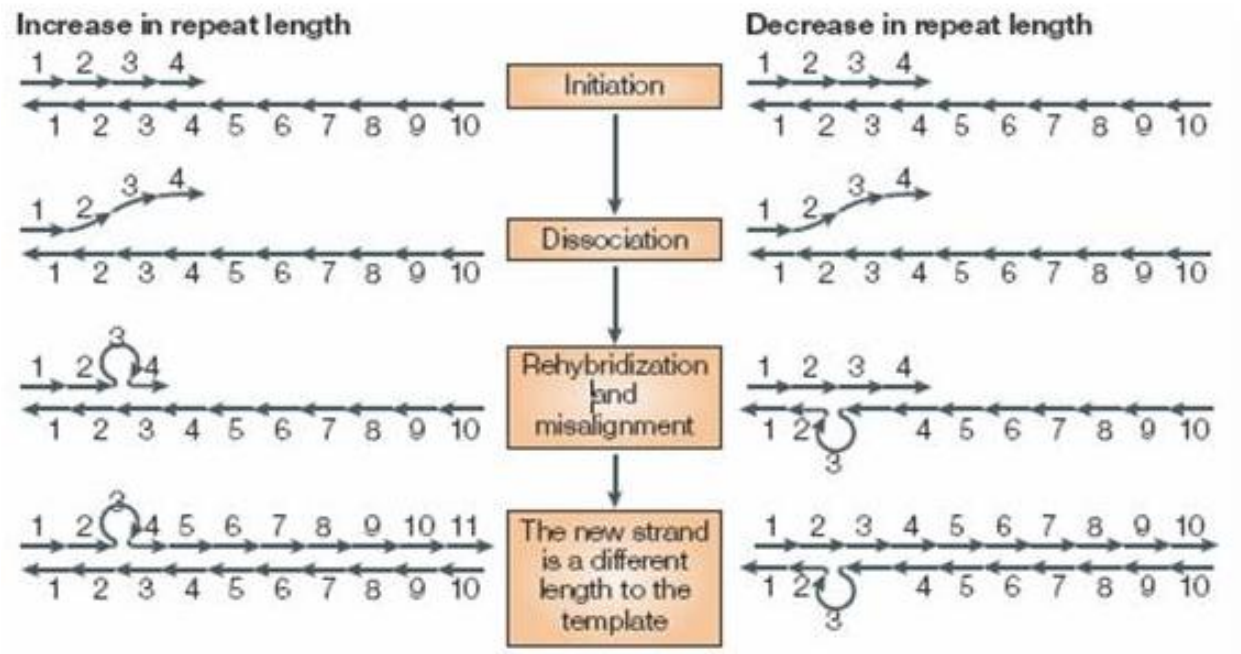


**Figure 3: The process of mutations formed in microsatellites by polymerase slippage from Al-Jawabreh (2006).**

The second mechanism to explain microsatellite mutations is unequal crossing during meiosis. During meiosis, chromosomal crossover occurs, where both of the chromosomes of a diploid organism line up and matching regions break and reconnect to the other chromosome resulting in genetic recombination. If the crossover occurs in a microsatellite locus and the fragment sizes are not equal on both chromosomes, then both chromosomes will result in different sized fragments (Figure 4). This mode of mutation is best for explaining extreme changes in microsatellite locus size (Voet *et al*, 2006).

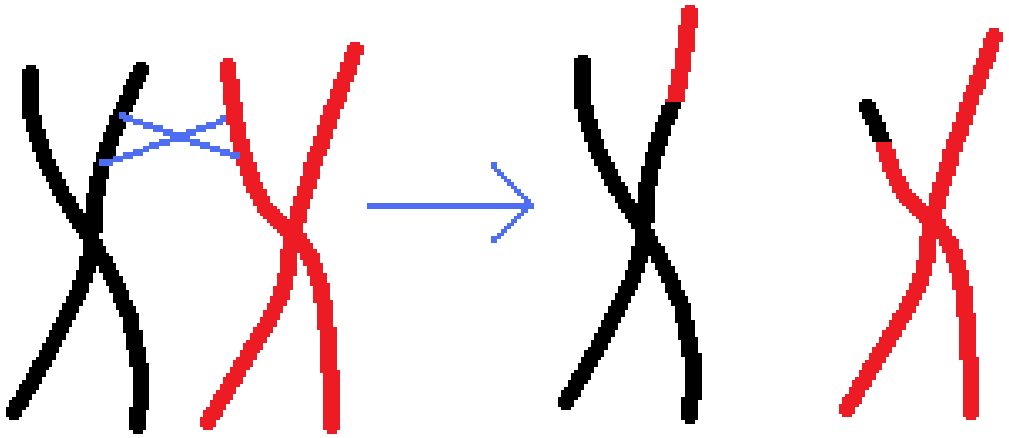

**Figure 4: Example of crossover resulting in uneven recombination sizes.**

Three evolutionary models of mutation have been created that can be used to study microsatellites with their high amount of polymorphism. The first model is the infinite alleles model (IAM), which says that each mutation creates a new allele at a certain rate, such that alleles that are identical in state are also identical by decent, because that allele could only have evolved once (Kimura & Crow, 1964). The second model is the K-allele model in which K allelic states are possible. The chance of a mutation happening at a given allele is equal to the rate of mutation and the K probability (Crow & Kimura, 1970). The third model is the stepwise mutation model (SMM) where each mutation adds or subtracts a single repeat from the allele. In the first two models all alleles can equally differ from one another but under the stepwise mutation model, allele sizes that are similar are more closely related that those that differ more (Ohta & Kimura, 1973).

If microsatellites do evolve according to a SMM, one of the potential problems a researcher might encounter is size homoplasy. Homoplasy is when alleles are identical in size but are not identical in descent (Liepelt *et al*, 2001; Gaggiotti *et al*, 1999). This can best be explained by the example shown in Figure 5. If someone was to only look at microsatellites 4, 5, 6 and 7 it would appear that there were two allele types, allele 4, 5 and 6 which are all one allele size and allele 7 which is another allele size. Using the stepwise mutation model, this would indicate that microsatellites 4, 5 and 6 are more

closely related to each other than to microsatellite 7. This is not the case, as you can see microsatellite 6 is actually more closely related to microsatellite 7.

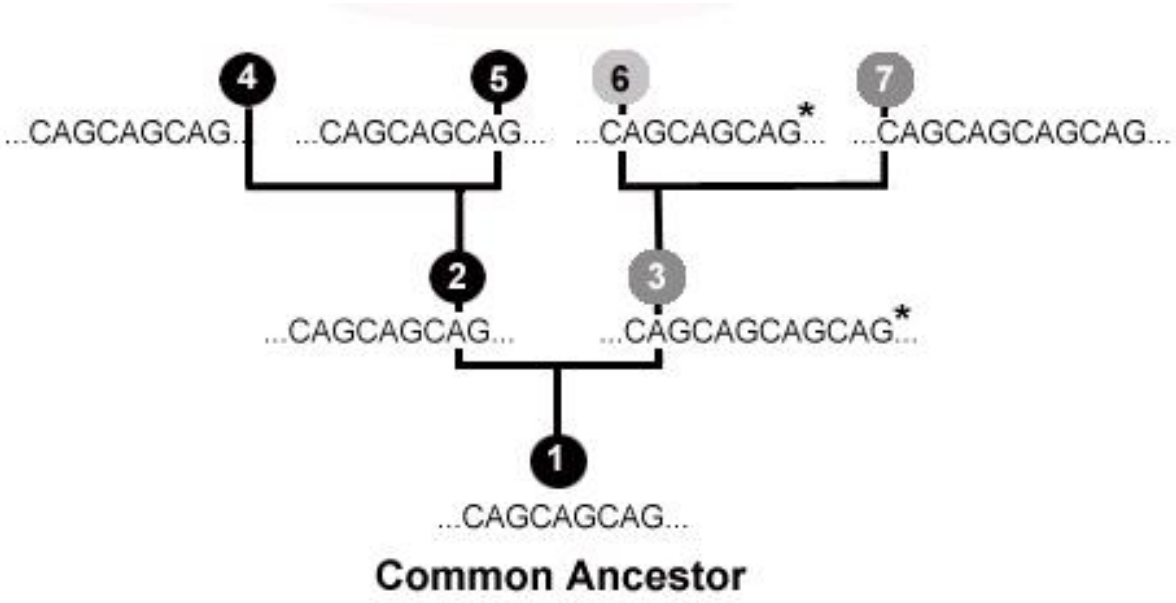


**Figure 5: An example of homoplasy from Melinek (2008).**

### *1.4 Statement of Hypothesis*

Crayfish, unlike many crustaceans, have direct development of young. In addition, their freshwater ecosystems are likely separated by both linear distance and frequent barriers, particularly the decreased capability to migrate to other watersheds. Consequently, crayfish populations probably experience low rates of dispersal and gene flow among populations in separate drainage basins, and even among sites within a drainage basin. Therefore, I hypothesize that three populations of *O. virilis* in Massachusetts, separated in different rivers and watersheds, will be differentiated from one another. I tested this hypothesis by comparing four microsatellite loci to determine the degree of genetic differentiation of these three populations.

# 2 Materials and Methods

### *2.1 Field collection*

Live crayfish of the species *Orconectes virilis* were collected from three sites in Massachusetts (Table 2; Figure 6) between March and November of 2005-2007. These three sites were Institute Pond (Site code: M4; Figure 7) and the Blackstone River (Site code: M9; Figure 8) in Worcester, MA and the French River (Site code: M8; Figure 9) in Oxford, MA. Crayfish were caught by hand, dip net, seine or crayfish/shiner traps baited with sardines or broccoli. Most crayfish were collected for earlier investigations, but additional specimens were collected for this study by trapping from the M4 site to supplement sample sizes.

**Table 2: Location of collection sites for *Orconectes virilis* and sample sizes.**

| Sample location | Location code | Latitude (N) | Longitude (W) | Sample size |
|---|---|---|---|---|
| Institute Pond- Worcester, MA | M4 | +42°16'38.47" | -71°48'19.60" | 29 |
| Blackstone River- Worcester, MA | M9 | +42°12'58.23" | -71° 47' 6.78" | 30 |
| French River- Oxford, MA | M8 | +42° 6'54.62" | -71°54'31.35" | 30 |

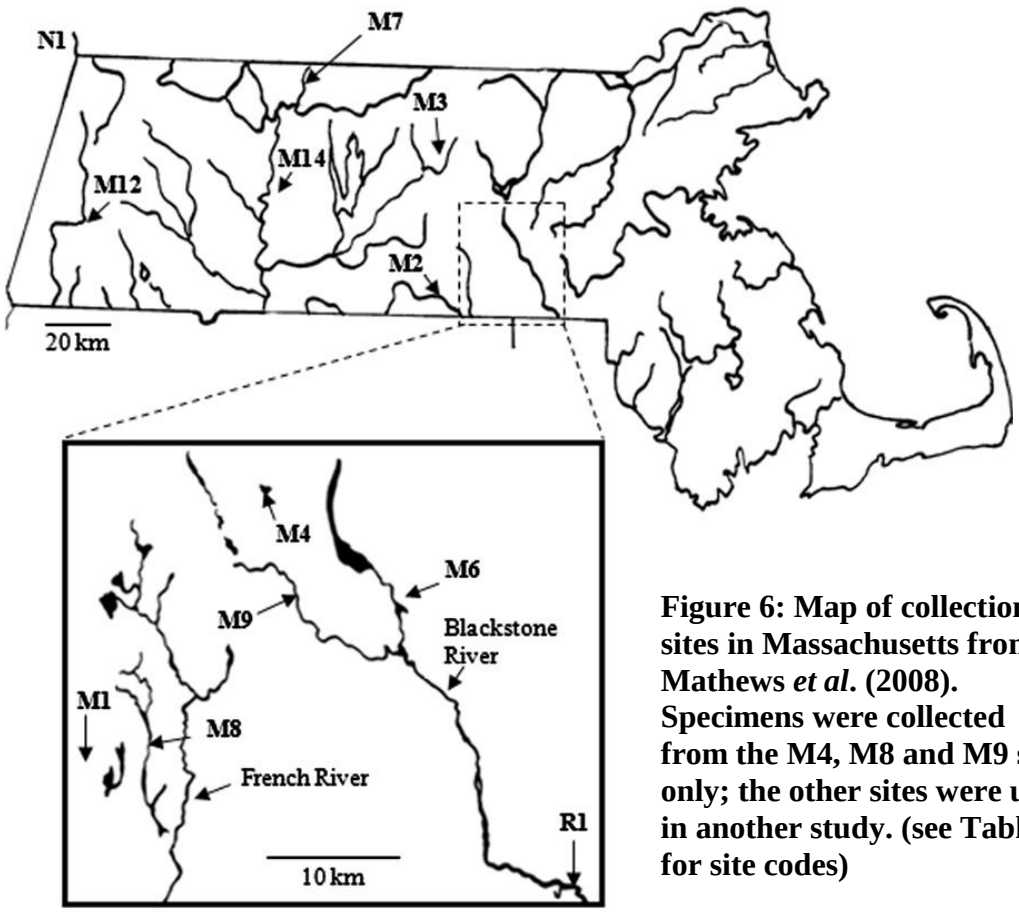


**Figure 6: Map of collection sites in Massachusetts from Mathews *et al*. (2008). Specimens were collected from the M4, M8 and M9 site only; the other sites were used in another study. (see Table 2 for site codes)**

**Figure 7: M4 Site: Institute Pond, Worcester, MA from Adams *et al* (2007).**

**Figure 8: M9 Site: Blackstone River, Worcester, MA.**

**Figure 9: M8 Site: Hodges Village Dam –French River, Oxford, MA from Adams *et al* (2007).**

The specimens collected were keyed as *O. virilis* or another species according to Hobbs (1972). Due to the difficulties in keying females and immature males, specimens that could not be keyed to a particular species were classified to a certain species if they matched the general morphology of that species (same chelae, coloration, pereiopods and personal observations) and came from a population where no other species was found (i.e., where previous investigations had indicated a "pure" population of *O. virilis*: Mathews *et al*, 2008). All crayfish were returned alive to Worcester Polytechnic Institute where they were euthanized and kept frozen at -80°C until needed.

### *2.2 DNA extraction*

I obtained tissue from 4 specimens of the M4 site for DNA extraction, while all the other specimens had already been extracted by others using a similar protocol. About 5-10mg of tissue was acquired from a pereiopod or chela and DNA was extracted according to the Solid Tissue Protocol in Gentra's Puregene: Genomic DNA Purification Kit. The tissue was place in 300µL of Cell Lysis Solution and 1.5µL of Puregene Proteinase K, mixed by inverting and incubated over night at 55°C for maximum yield. The Cell Lysis Solution causes the cells in the tissue to rupture, releasing their cellular components. The Puregene Proteinase K solution digests proteins and removes contaminants such as nucleases, which would degrade the DNA. After incubating overnight, 100µL of Protein Precipitation Solution was added to the tissue and vortexed at high speed for ~20 seconds. The sample was then cooled on ice for ~1 minute and centrifuged at 13,000 x g for 3 minutes to pellet the protein precipitate. If a tight pellet was not formed the tissue was placed back on ice and then re-centrifuged. The supernatant was then carefully added to 300µL of isopropanol and mixed by inverting gently ~50 times, which causes the DNA to precipitate. The DNA was then pelleted by centrifuging at 13,000 x g for 1 minute. The supernatant was then discarded carefully to avoid removing the pellet. The pellet was gently washed with 300µL of 70% ethanol and repelleted by centrifuging for 1 minute at 13,000 x g. The supernatant was discarded and the pellet was allowed to dry for ~15 minutes. After drying, the DNA pellet was resuspended in 50µL of DNA Hydration Solution with vortexing. The solution was then

incubated at 65°C for ~1 hour and overnight at room temperature to dissolve the DNA. After incubation the solution was stored at -20°C until needed.

The DNA extractions were run on a 1% agarose gel to confirm DNA extraction success and to quantify the DNA. If the DNA concentration was very low or not present at all a second extraction was carried out on that specimen, this included specimen from the previous extractions which were stored at -80°C. The amount of DNA was estimated by running 100ng/µL of lambda bacterial DNA and comparing it to the specimen DNA. The DNA was then diluted to ~10ng/µL accordingly.

### *2.3 Polymerase Chain Reaction (PCR)*

Polymerase chain reaction (PCR) is a technique widely used by biologist to quickly multiply fragments of DNA. PCR reactions were carried out for each crayfish locus separately in 20µL volumes containing 1X PicoMaxx Reaction Buffer (Stratagene), 0.15mM dNTPs (New England BioLabs), 0.2mM of both the forward and reverse primers, 0.25U PicoMaxx Enzyme (Stratagene), 0.25U Taq Extender (Stratagene) and ~10ng of genomic DNA.

Stratagene's PicoMaxx is a blend of cloned Taq and Pfu DNA polymerases and polymerase enhancing factors which results in a higher success rate than other PCR enzymes at amplifying DNA fragments (Stratagene, 2007). Taq is a thermally stable enzyme that was originally isolated from the thermophilic bacterium *Thermus aquaticus* (Voet *et al*, 2006). Its ability to withstand denaturing at high temperature makes it a useful tool in PCR. Before Taq polymerase, other DNA polymerases were used which needed to be added to samples at each cycle because of denaturing from heating, but Taq only needs to be added once. One of the drawbacks to Taq polymerase is its low replication fidelity, because it lacks 3' to 5' exonuclease proofreading activity. However, PicoMaxx also contains Pfu polymerase, which is a thermally stable proofreading DNA polymerase isolated from *Pyrococcus furiosus*; Pfu does posses 3' to 5' exonuclease proofreading activity. But one of the drawbacks to Pfu polymerase is it increases the length of time of PCR during elongation (Stratagene, 2007).

Four separate microsatellites loci were chosen to be amplified from each specimen. The primers that were used, the expected DNA fragment size, the fluorescent

label, and the microsatellite sequence for each locus can be found in Table 3. The forward primers were fluorescently labeled. Each pair of primers was run in a separate PCR.

**Table 3: List of loci and corresponding primers used.**

| Primer names | Sequence (5' to 3') | Primer names | Sequence (5' to 3') | Expected product size (bp) | Fluorescent label (forward) | Msat seq. |
|---|---|---|---|---|---|---|
| Ov50 F | GGAACTGACAGTAG AAACAA | Ov50 R | GTTGAGTGGCGGG ACCAAAG | 217 | VIC | $GT_{10}$ |
| Ov62 F | GTGAGTGTTAGACA CCTTTAC | Ov62 R | GAAATCTTAGGAG GACAGC | 203 | 6-FAM | $CT_{12}$ |
| Ov82 F | GTCTCCCGCTATTCA ATTAC | Ov82 R | CGGATATTGATCA ATGATGC | 211 | NED | $GA_{21}$ |
| Ov86 F | GACGATAAGGTCTC CATA | Ov86 R | ACGTCTTGAGCTCC ACTAC | 150 | PET | $GA_{14}$ |

The PCR amplifications were performed in a MJ Research PTC-225 Peltier Thermal Cycler with the cycling conditions as follows: 95°C for 2mins, then 40 cycles of 95°C for 30s, 56°C for 30s and 72°C for 1min, followed by 72°C for 10mins. Reactions were then held at 4°C.

The 95°C for two minutes is called the initialization step (Voet *et al*, 2006). The high temperature is required to activate some types of DNA polymerase, but is not necessary for all types. The next step, called the denaturation step, heats the PCR solution to 95°C for 30 seconds. This causes the double stranded DNA and primers to "melt", which separates the double stranded DNA into single stranded and causes the primers to dissociate (Figure 10). The next step, called the annealing step, cools the solution to 56°C for 30 seconds. The cooling allows the primers to bind to their complementary sequence on the DNA and for the polymerase to bind to the primer-DNA template, which then begins DNA synthesis. The solution is then heated to 72°C for one minute in a step called elongation. This temperature is the optimum activity temperature for the polymerase. The polymerase synthesizes a new strand of DNA complementary to the DNA template by adding dNTPs. The polymerase will continue to synthesis DNA until it reaches the end of the template or is denatured again (Voet *et al*, 2006). The denaturing step, annealing step and elongation step are then repeated (also known as cycling) 39 more times. Since both

forward and reverse primers were used, the DNA sequence between them was copied each cycle. The last step, called the final elongation, holds the temperature at 72°C for 10 minutes to make sure any remaining single stranded DNA is fully copied. The solution is then cooled to 4°C for storage. Figure 10 shows four cycles during PCR. In a preliminary experiment, I found that the annealing temperatures of each of the primers were insensitive to a ten degree gradient change during PCR amplification and all provided an adequate amplification of the microsatellite of interest at 56°C. This allowed all four primers sets to be used with the same cycling conditions, even though each of the primers predicted optimal annealing temperatures may have been different. The PCR products were run on 2% agarose gels along with negative controls to confirm pure amplification and concentration.

From the PCR amplification, 1.0µL from each of the loci were multiplexed by adding together with 0.5µL of Gene Scan- 600LIZ size standard, (Applied Biosystems) and 9.0µL of formamide to one well of a 96-well plate. This was sent to Cornell University Life Sciences Core Laboratories Center for automated DNA fragment size determination on their Applied BioSystems 3730xl DNA Analyzer, which uses a fluorescence based detection system.

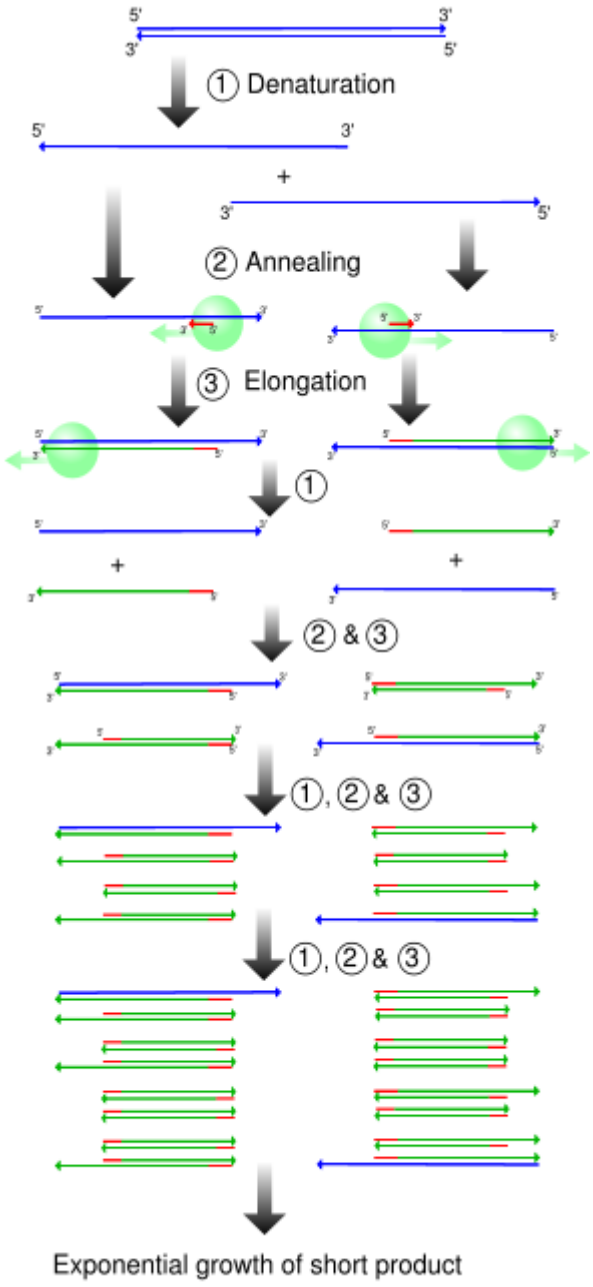


**Figure 10: An example of four cycles during PCR from Ball (2007).**

### *2.4 Fragment Analysis*

The data received from the automated fragment analysis were analyzed on the Applied Biosystems GeneMapper Software v4.0 to determine the number of alleles and fragment size(s). The number of alleles, which correlated to the number of fragments, indicated whether the specimen was homozygous or heterozygous for the microsatellite.

Each locus was scored differently depending on "background noise", contamination fluorescence and expected fragment sizes. Only fluorescence greater than the background noise, contamination fluorescence and within the expected fragment sizes were scored. For the Ov50 locus, only fragment sizes which resulted in fluorescence greater than 200 on the y-axis and had fragment sizes between 250 and 300bp were considered. For the Ov62 locus, only fragment sizes which resulted in fluorescence greater than 800 on the y-axis and had fragment sizes between 200 and 250 were considered. For the Ov82 locus, only fragment sizes which resulted in fluorescents greater than 300 on the y-axis and had fragment sizes between 215 and 300 were considered. Finally for the Ov86 locus, only fragment sizes which resulted in fluorescents greater than 120 on the y-axis and had fragment sizes between 135 and 280 were considered. In addition to these general criteria for scoring, there were other factors I had to consider. For example, while scoring I found evidence of stutter bands and contaminants in the negative controls. Stutter bands are easily distinguished from the actual fragment sizes by their consistent increase in fragment size from the actual fragment size and their almost identical peak and noise patterns. When a contaminant was found in a negative control and resulted in peaks that reached the loci inclusion fluorescence intensity and fell within the expected fragment size range then the locus was rejected. This happened once in the Ov62 locus for samples from the M4 site, so this locus was excluded for this site. Also, to increase the reliability of my data only specimens which were scored successfully for three loci were kept while all others were excluded.

### *2.5 Data analysis*

The program Cervus 2.0 (Kasumovic *et al*, 2003) was used to calculate the observed and expected heterozygosities. GENEPOP 3.3 (Raymond *et al*, 1995) was used to estimate the departures from Hardy-Weinberg equilibrium by the Markov chain method and the number of migrants between populations was estimated using the private allele method ($N_m$). An AMOVA in Arlequin 3.01 (Excoffier *et al*, 2005) was performed to estimate $F_{ST}$ and $R_{ST}$ values and to estimate the population pairwise $F_{ST}$ and $R_{ST}$ values based on 10,000 permutations.

# 3 Results

Genetic variation of three populations of *O. virilis* was investigated using four polymorphic microsatellite loci, except in the M4 population, where only three microsatellite loci were used due to apparent contamination (strong amplifications in the negative controls). The contamination of the Ov62 locus resulted in the higher percent failure (Table 4) of attempted locus amplifications that failed to produce usable results. The number of different alleles per locus in each population ranged from 4 to 12. The relative distribution of alleles for each locus in the populations can be seen in Figure 11. The observed heterozygosity ranged from 0.214 to 1.000 while the expected heterozygosity ranged from 0.611 to 0.915 (Table 4). The Ov82 and Ov86 loci for all three populations showed significant heterozygote deficiencies ($p < 0.05$).

$F_{ST}$ and $R_{ST}$ were interpreted using the following guidelines: 0.00-0.05 indicated little genetic differentiation, 0.05-0.15 indicated moderate genetic differentiation, 0.15-0.25 indicated very large genetic differentiation and >0.25 indicates extensive genetic differentiation (Balloux *et al*, 2002; Nguyen *et al*, 2006). The AMOVA results using $F_{ST}$ analysis showed that all of the allele variation was partitioned within the populations, and the $F_{ST}$ was not significantly different from zero (Table 5). On the other hand, $R_{ST}$ analysis showed that, while the majority of the allele variation was partitioned within the populations, 16.79% was partitioned among the populations and, using this method, the $R_{ST}$ was significantly different from zero (Table 5). The pairwise population comparisons for $F_{ST}$ showed no significant genetic divergence between populations, while $R_{ST}$ showed extensive significant genetic differentiation between the M4:M8 and M4:M9 populations (Tables 6 and 7). Using the private allele method (alleles found in only one population), the estimated number of immigrants between each population ranged from 1.22 to 2.37 immigrants per generation (Nm) (Table 8).

**Table 4: Summary of information for the microsatellite loci from three sites.**

| Locus | % Failed | | Collection site | | |
|---|---|---|---|---|---|
| | | | M4 | M8 | M9 |
| Ov50 | 44.9 | $H_o$ | 0.857 | 1.000 | 0.815 |
| | | $H_e$ | 0.611 | 0.767 | 0.664 |
| | | $N_a$ | 4 | 5 | 6 |
| Ov62 | 57.3 | $H_o$ | - | 0.846 | 0.800 |
| | | $H_e$ | - | 0.760 | 0.773 |
| | | $N_a$ | - | 5 | 5 |
| Ov82 | 44.9 | $H_o$ | 0.214* | 0.571* | 0.238* |
| | | $H_e$ | 0.685 | 0.839 | 0.722 |
| | | $N_a$ | 6 | 9 | 9 |
| Ov86 | 39.3 | $H_o$ | 0.286* | 0.333* | 0.360* |
| | | $H_e$ | 0.833 | 0.915 | 0.829 |
| | | $N_a$ | 8 | 12 | 12 |

$H_o$ = observed heterozygosity; $H_e$ = expected heterozygosity; $N_a$ = number of alleles; Asterisks indicate significant heterozygote deficiencies (p-values < 0.05).
The percent failed is the proportion of attempted locus amplifications that fail to produce usable results.

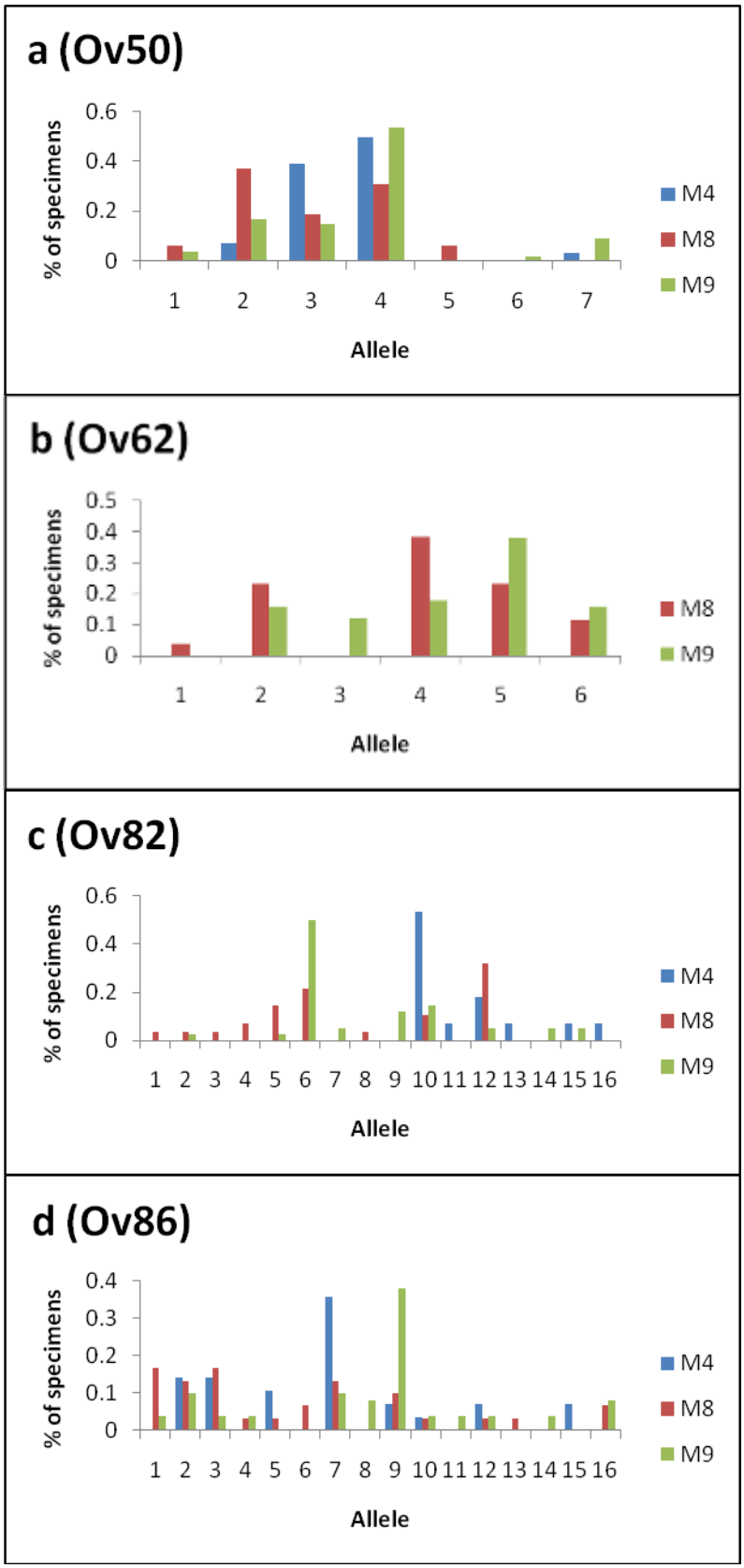


**Figure 11: A comparison of the number alleles between populations by locus. Alleles are ordered smallest (Allele 1) to largest within a locus. In b (Ov62) the M4 site was not scored for this population.**

**Table 5: AMOVA results using $F_{ST}$ and $R_{ST}$ analysis.**

| Analysis | Source of variation | Degrees of freedom | Sum of squares | Percentage of Variation | Fixation Index (P-value) |
|---|---|---|---|---|---|
| $F_{ST}$ | Among populations | 2 | 0.586 | -2.09 | -0.021 (0.994) |
| | Within populations | 111 | 127.1 | 102.09 | |
| $R_{ST}$ | Among populations | 2 | 1148.2 | 16.79 | 0.168 (0.000) |
| | Within populations | 111 | 7656.7 | 83.21 | |

**Table 6: Estimates of pairwise $F_{ST}$ for all population comparisons.**

| | M4 | M8 | M9 |
|---|---|---|---|
| M4 | | 0.997±0.000 | 0.999±0.000 |
| M8 | -0.071 | | 0.138±0.003 |
| M9 | -0.039 | 0.016 | |

Below diagonal are pairwise $F_{ST}$ values and above are P values

**Table 7: Estimates of pairwise $R_{ST}$ for all population comparisons.**

| | M4 | M8 | M9 |
|---|---|---|---|
| M4 | | 0.000±0.000 | 0.000±0.000 |
| M8 | 0.307 | | 0.598±0.004 |
| M9 | 0.261 | -0.007 | |

Below diagonal are pairwise $R_{ST}$ values and above are P values

**Table 8: Estimated number of migrants ($N_m$) per generation between populations.**

| Populations | Mean sample size | Mean frequency of private alleles | Nm (after correction for size) |
|---|---|---|---|
| M4 and M8 | 13.167 | 0.073 | 1.831 |
| M4 and M9 | 19.167 | 0.080 | 1.223 |
| M8 and M9 | 18.500 | 0.056 | 2.375 |

# 4 Discussion

I hypothesized that the three populations of *O. virilis* studied would be differentiated from one another. This was believed to be due to the direct development of young, low dispersal/migration rate and the separation of populations between rivers and distance. Using variable microsatellite loci, I conclude that there is some differentiation between the populations.

The microsatellite data indicated a significant heterozygote deficiency in the Ov82 and Ov86 loci relative to the expectations of Hardy-Weinberg equilibrium, but these loci also showed an increased number of alleles. Heterozygote deficiency can be caused by a number of factors, including inbreeding, nonrandom sampling, Walhund effect, founder effect, null alleles, selection against heterozygotes, fishing pressures or any combination of these, but null alleles is the most common cause when using microsatellites (Wang *et al*, 2007; Foltz, 1986). It should be noted that the populations may not be heterozygote deficient, since the other two loci used, Ov50 and Ov62, were not deficient. These differences in results between loci maybe due to the microsatellite locus itself or my high failure rate after excluding unusable microsatellite data. The high failure rate was mainly due to a contaminant found in my negative controls, which resulted in the removal of many data sets/alleles. This contaminant may have also resulted in the differences in heterozygote deficiencies between loci, due to its unknown nature. An example of this would be the scoring/use of false alleles in the data analysis. More samples and microsatellite loci data are needed before the heterozygosity of these populations can be determined.

To estimate the connections and pattern of gene flow among populations, the majority of studies use Wright's $F_{ST}$ and/or Slatkin's $R_{ST}$ statistics (Balloux *et al*, 2002). The AMOVA results for the $F_{ST}$ analysis indicates that there was no differentiation ($F_{ST}$ = -0.021, P = 0.994) among the populations, but statistically these results were not significant. Seeing absolutely no differentiation among the populations seems counterintuitive considering gene flow should be hindered by physical barriers and distance between the populations and should result in some differentiation (higher $F_{ST}$). The pairwise $F_{ST}$ values between all the population combinations did not show any significant pairwise differentiation, with the highest $F_{ST}$ of 0.016 occurring between the

M8:M9 populations. These results imply that there is sufficient gene flow among these populations to maintain genetic homogeneity, which is further supported by the Nm data that indicated there are 1.22 to 2.37 migrants per generation between the populations.

The AMOVA results for the $R_{ST}$ analysis support a much larger genetic differentiation ($R_{ST}$ = 0.168, P = 0.000) among the populations, which is in support of my hypothesis. The pairwise $R_{ST}$ values between the M4:M8 and M4:M9 populations did show very significant genetic differentiation between them, while the M8:M9 populations did not show any significant differentiation. It should be pointed out that there was more differentiation between the M4:M8 populations than the M4:M9 or the M8:M9 populations. It was expected that the M4:M8 populations would be the most differentiated since they were the furthest apart of the populations and located in separate watersheds, but the differentiation between the M8:M9 populations, which are in different watersheds, was expected to be greater than that between the M4:M9 populations, which are in the same watershed. The actual results may be different from what was expected due to unexpected migration patterns and/or human transfer of individuals between populations, or possibly because of skewed results due to the contaminant.

There is no clear consensus about the accuracy of $F_{ST}$ and $R_{ST}$ statistical estimations of population differentiation but each of these statistics has its drawbacks. $F_{ST}$ assumes the infinite allele model (IAM), while $R_{ST}$ assumes the stepwise mutation model (SMM). The SMM is believed to reflect the actual way microsatellites mutate the majority of the time. When the microsatellites mutate by adding or subtracting a single nucleotide repeat, the SMM is highly accurate. Also, as long as the mutations follow the SMM, $R_{ST}$ is independent of the mutation rate since it is based only on the addition and subtraction of a single nucleotide repeat, and thus has a "memory" of the mutations. But microsatellite can deviate from the SMM, such as when large changes in fragment sizes occur during chromosomal crossover. Eventually the IAM will become more accurate because it allows all mutations to result in novel alleles at a certain mutation rate, which may be closer to the larger fragment changes than the single repeat changes in the SMM. In general, according to the literature, $R_{ST}$ provides a less biased estimate of population dynamics than $F_{ST}$ in most instances (Balloux *et al*, 2002; Slatkin, 1995).

If my $R_{ST}$ estimates for the differentiation among the populations are correct then there is some differentiation between the populations. This could be due to direct development of young, low dispersal/migration rate and the separation of populations between rivers and distance, which is also seen in other studies of similar species (Mäkinen *et al*, 2007; Jones *et al*, 2006; Hughes *et al*, 2003; Nguyen *et al*, 2005; Fetzner and Crandall, 2003). More specimen and microsatellite loci are needed to make any reliable conclusion on the differentiation among these populations. It may also be useful to collect more specimens from other watersheds and within the same watersheds to determine if differentiation is cause by isolation between bodies of water and/or distance.

# References


Adams L, Basile M (2007) A study of the local invasive crayfish species (*Orconectes virilis*) using population genetics. Worcester, MA: W.P.I. Unpublished MQP report.

Aiken DE (1968) The crayfish *Orconectes virilis*: survival in a region with severe winter conditions. *Canadian Journal of Zoology* **46**, 207-211.

Al-Jawabreh, A. (2006) *Molecular Epidemiology, Clinical Molecular Diagnosis and Genetic Diversity of Cutaneous Leishmaniasis*. Retrieved on Jan 29 2008. Available at: http://edoc.hu-berlin.de/dissertationen/al-jawabreh-amer-2005-11-25/HTML/front.html

Applied Biosystems. (2000). GeneScan® Reference Guide Chemistry Reference for the ABI PRISM® 310 Genetic Analyzer. Applied Biosystems, Foster City, CA.

Avise JC, Arnold J, Ball RM, Bermingham Jr E, Lamb T, Neigel JE, Reed CA, Saunders NC (1987) Intraspecific phylogeography: the mitochondrial DNA bridge between population genetics and systematics. *Annual Review of Ecology and Systematics,* **18**, 489-522.

Ball M (2007). *Madeleine Price Ball*. Retrieved on Jan 2 2008. Available at: http://mad.printf.net/

Balloux F, Lugon-Moulin N (2002) The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* **11**, 155–165.

Bourgain C, Abney M, Schneider D, Ober C, McPeek MS (2004) Testing for Hardy-Weinberg Equilibrium in samples with related individuals. *Genetics,* **168**, 2349-2361.

Brookfield JFY (1996) A simple new method for estimating null allele frequency from heterozygote deficiency. *Molecular Ecology* **5**, 453–455.

Buhay JE, Crandall KA (2005) Subterranean phylogeography of freshwater crayfishes shows extensive gene flow and surprisingly large population sizes. *Molecular Ecology*, **14**, 4259–4273.

Burridge CP, Craw D, Waters JM (2007) An empirical test of freshwater vicariance via river capture. *Molecular Ecology,* **16**, 1883-1895.

Carnegie Museum of Natural History (2006) *The crayfish and lobster taxonomy browser*. Retrieved on Feb 8 2008. Available at: http://iz.carnegiemnh.org/crayfish/NewAstacidea/infraorder.asp?io=Astacidea

Chafee JH (2006) *Blackstone River Valley*. Retrieved on Feb 8 2008. Available at: http://www.nps.gov/blac/planyourvisit/index.htm

Chunyan W, Xiaomu Y, Jingou T (2007) Microsatellite diversity and population genetic structure of redfin culter (*Culter erythropterus*) in fragmented lakes of the Yangtze River. *Hydrobiologia*, **586**, 321-329.

Crandall KA (2006) Applications of phylogenetics to issues in freshwater crayfish biology. *Bull. Fr. Pêche Piscic.*, **380-381**, 953-964.

Crow JF, Kimura M (1970) An Introduction to Population Genetics Theory. Harper and Row, New York.

Di Rienzo A, Peterson AC, Garza JC, Valdes AM, Slatkin M, Freimer NB (1994) Mutational processes of simple-sequence repeat loci in human populations. *Proc. Natl. Acad. Sci.* **91**, 3166-3170.

Diethard T (1989) Hypervariability of simple sequences as a general source for polymorphic DNA markers. *Nucl. Acids Res.*, **17**, 6463-6471.

Erlich HA, Gelfand D, Sninsky JJ (1991) Recent advances in the polymerase chain reaction. *Science*, **252**, 1643–1651.

Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, **11**, 1591–1604.

Excoffier LGL, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online,* **1**, 47-50.

Fetzner Jr JW, Crandall KA (2003) Linear habitats and the Nested Clade Analysis: An empirical evaluation of geographic vs. river distances using an Ozark crayfish (Decapoda: Cambaridae). *Evolution*, **57**, 2101-2118.

Foltz D (1986) Null alleles as a possible cause of heterozygote deficiencies in the oyster *Crassostrea virginica* and other bivalves. *Evolution,* **40**, 869-870.

Gaggiotti OE, Lange O, Rassmann K, Gliddon C (1999) A comparison of two indirect methods for estimating average levels of gene flow using microsatellite data. *Molecular Ecology,* **8**, 1513–1520.

Gherardi F (2002) Chapter 7: Behaviour. In Biology of Freshwater Crayfish. (Edited by Holdich DM). Oxford, England. Blackwell Science Ltd.

Goldstein DB, Linares AR, Cavalli-Sforza LL, Feldman MW (1995) An evaluation of genetic distances for use with microsatellite loci. *Genetics,* **139**, 463-471.

Halliburton R (2004) Introduction to Population Genetics. Pearson Prentice Hall: Upper Saddle River, NJ.

Harm P (2002) Chapter 15: Orconectes. In *Biology of Freshwater Crayfish*. (Edited by Holdich DM). Oxford, England. Blackwell Science Ltd.

Hobbs HH (1972). Crayfishes (Astacidae) of North and Middle America. Biota of Freshwater Ecosystems Identification Manual 9, Environmental Protection Agency.

Hobbs Jr H (1972) *Biota of Freshwater Ecosystems: Identification Manual No. 9. Crayfishes (Astacidea) of North and Middle America*. Retrieved on Jan 4 2008. Available at: http://iz.carnegiemnh.org/crayfish/Keys/index2.htm

Hobbs Jr. HH (1989) An illustrated checklist of the American crayfishes. *Smithsonian Contributions to Zoology* **480**, 236.

Holdich DM (2004) Chapter 1: Background and Functional Morphology. In Biology of Freshwater Crayfish. (Edited by Holdich DM). Oxford, England. Blackwell Science Ltd.

Hughes JM, Hillyer MJ (2003) Patterns of connectivity among populations of Cherax destructor (Decapoda: Parastacidae) in western Queensland, Australia. *Mar. Freshwater Res.*, **54**, 587-596.

Jarne P, Lagoda PJL (1996) Microsatellites, from molecules to populations and back. *Trends Ecol. Evol.*, **11**, 424-429.

Jin L, Macaubas C, Hallmayer J, Kimura A, Mignot E (1993) Mutation rate varies among alleles at a microsatellite locus: Phylogenetic evidence. *Proceedings of the National Academy of Sciences*, **93**, 15285-15288.

Jones TM, Voss RS, Ptacek MB, Weisrock D, Tonkyn D (2006) River drainages and phylogeography: An evolutionary significant lineage of shovel-nosed salamander (*Desmognathus marmoratus*) in the southern Appalachians, *Molecular Phylogenetics and Evolution*, **38**, 280-287.

Kalinowski ST, Taper ML, Marshall TC (2007) Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology*, **16**, 1099-1006.

Kasumovic MM, Ratcliffe LM, Boag PT (2003) A method to improve confidence in paternity assignment in an open mating system. *Canadian Journal of Zoology*, **81**, 2073-2076.

Kimura M, Crow JF (1964) The number of alleles that can be maintained in a finite. population. *Genetics* **49**, 725-738.

Liepelt S, Kuhlenkamp V, Anzidei M, Vendramin GG, Ziegenhagen B (2001) Pitfalls in determining size homoplasy of microsatellite loci. *Molecular Ecology Notes*, **1**, 332–335.

Mäkinen HS, Merilä J (2007) Mitochondrial DNA phylogeography of the three-spined stickleback (*Gasterosteus aculeatus*) in Europe – evidence for multiple glacial refugia. *Molecular Phylogenetics and Evolution*, in press.

Mathews LM, Adams L, Anderson E, Basile M, Gottardi E, Buckholt MA (2008) Genetic and morphological evidence for substantial hidden biodiversity in a freshwater crayfish species complex. *Mol. Phylogenet. Evol*, doi:10.1016/j.ympev.2008.02.006

Melinek R. (2008) *Fishing Out Microsatellites*. Retrieved on Jan 29 2008. Available at: http://www.woodrow.org/teachers/esi/2002/Biology/Projects/p3/index.htm

Momot WT, Gowing H (1972) Differential seasonal migration of the crayfish, *Orconectes virilis* (Hagen) in Marl Lakes. *Ecology*, **53**, 479–483.

National Biological Information Infrastructure (NBII) & IUCN/SSC Invasive Species Specialist Group (ISSG). (2005). Orconectes virilis (crustacean). The Global Invasive Species Database. Retrieved on Jan 29 2008. Available at: http://www.issg.org/database/species/ecology.asp?si=218&fr=1&sts

Nguyen T, Hurwood D, Mather P, Na-Nakorn U, Kamonrat W, Bartley D (2006) Manual on application of molecular tools in aquaculture and inland fisheries management. Part 2. Laboratory protocols and data analysis. NACA Monographs, 2. Network of Aquaculture Centres in Asia-Pacific, Bangkok, Thailand.

Nguyen TTT, Burridge CP, Austin CM (2005) Population genetic studies on the Australian freshwater crayfish, *Cherax destructor* (Parastacidae: Decapoda) using allozyme and RAPD markers. *Aquatic Living Resources*, **18**, 55–64.

Nystrom P (2002) Chapter 5: Ecology. In Biology of Freshwater Crayfish. (Edited by Holdich DM). Oxford, England. Blackwell Science Ltd.

Ohta T, Kimura M (1973) A model of mutation appropriate to estimate the number of electrophoretically detectable alleles in a genetic population. *Genet. Res*., **22**, 201-204.

Ortí G, Pearse DE, Avise JC (1994) Phylogenetic assessment of length variation at a microsatellite locus. *Proceedings of the National Academy of Sciences*, **94**, 10745-10749.

Queller DC, Strassmann LE, Hughes CR (1993) Microsatellites and kinship. *Trends Ecol. Evol*., **8**, 285-288.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J. Heredity*, **86**, 248-249.

Reynolds JD (2002) Chapter 4: Growth and reproduction. In Biology of Freshwater Crayfish. (Edited by Holdich DM). Oxford, England. Blackwell Science Ltd.

Richards C, Kutka FJ, McDonald ME, Merrick GW, Devore PW (1996) Life history and temperature effects on catch of northern orconectid crayfish. *Hydrobiologia* **319**, 111-123

Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics*, **145**, 1219-1228.

Rubinsztein D, Amos B, Cooper G (1999) Microsatellite and trinucleotide-repeat evolution: Evidence for mutational bias and different rates of evolution in different lineages. *Philosophical Transactions: Biological Sciences*, **354**, 1095-1099.

Schlötterer C (2001) Evolutionary dynamics of microsatellite DNA. *Chromosoma*, **109**, 565-571.

Scholtz G, Kawai T (2002) Aspects of embryonic and postembryonic development of the Japanese freshwater crayfish *Cambaroides japonicus* (Crustacea, Decapoda) including a hypothesis on the evolution of maternal care in the Astacida. *Acta Zoologica*, **83**, 203–212.

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, **9**, 615–629.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

Somers KM, Stechey DPM (1986) Variable trappability of crayfish associated with bait type, water temperature and lunar phase. *American Midland Naturalist*, **116**, 36-44.

Stratagene. (2007). Pfu DNA Polymerase; Instruction Manual. Stratagene, La Jolla, CA.

Stratagene. (2007). PicoMaxx® High Fidelity PCR System; Instruction Manual. Stratagene, La Jolla, CA.

Subramanian S, Mishra R, Singh L (2003) Genome-wide analysis of microsatellite repeats in humans: their abundance and density in specific genomic regions. *Genome Biol*., **4**, R13.

Toh Y, Oki E, Oda S (1996) An integrated microsatellite length analysis using an automated fluorescent DNA sequencer. *Cancer Res.*, **56**, 2688-91.

Turner TF, Trexler JC, Kuhn D, Robison H (1996) Life history variation and comparative phylogeography of darters (Pisces:Percidae) from the North American central highlands. *Evolution*, **50**, 2023-2036.

Voet D, Voet J, Pratt C (2006) Fundamentals of Biochemistry; Life at the Molecular Level 2nd. Hoboken: Wiley.

Weber JL, Wong C (1993) Mutation of human short tandem repeats. *Hum. Mol. Genet.*, **2**, 1123-l128

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution* **38,** 1358-70.

Wright S (1978) Evolution and the Genetics of Populations Vol. 4. Variability Within and Among Natural Populations. Univ. of Chicago Press, Chicago.

Yazdani R, Scotti I, Jansson G, Plomion C, Mathur G (2003) Inheritance and diversity of simple sequence repeat (SSR) microsatellite markers in various families of *Picea. abies*. *Hereditas* **138**, 219-227.

Zane L, Bargelloni L, Patarnello T (2002) Strategies for microsatellite isolation: a review. *Molecular Ecology,* **11**, 1–16.

Zhong N, Yang W, Dobkin C, Brown WT (1995) Fragile X gene instability: anchoring AGGs and linked microsatellites. *Am. J. Hum. Genet*., **57**, 351–61.