Assessing the Efficacy of Known *Procambarus clarkii* Primers to Amplify Microsatellite Loci in *Orconectes limosus* and *O. quinebaugensis*

A Major Qualifying Project

Submitted to the Faculty of

WORCESTER POLYTECHNIC INSTITUTE

in partial fulfillment of the requirements for the

Degree in Bachelor of Science

Biology/Biotechnology and Biochemistry

By

Brandon Poole

Meghan Sawicki

Nicole Smolko

Angela Wood

Advisor:

Michael Buckholt

Advisor:

Lauren Mathews

May 1, 2014

Abstract

The goal of this project was to assess the efficacy of microsatellite primers in three crayfish species present in Massachusetts. We began by investigating twenty primer pairs previously developed for *Procambarus clarkii*. After selecting three loci for further study, we explored the applicability of those three primer pairs in two additional species of crayfish, *Orconectes limosus* and *O. quinebaugensis*. Resulting products from *P. clarkii* were sent to Cornell University for fragment analysis, but the analysis proved inconclusive.

Acknowledgements

Thank you to our advisors, Michael Buckholt and Lauren Mathews. The knowledge and skills we have acquired from you are invaluable assets to us as we seek careers in biology and biochemistry. We would also like to thank Abby White, who always ensured that we had the materials we needed to carry out the various phases of our project.

Table of Contents

Figures

Tables

1. Introduction

Scientists have long been interested in understanding and classifying the living organisms that inhabit the Earth. Since Charles Darwin published *The Origin of Species* in 1859, researchers have tried to explain how species relate to one another and how traits are inherited and altered from generation to generation (Beebee & Rowe, 2003). Methods for answering these types of questions have become increasingly sophisticated, and there are now various molecular methods that are utilized to study inheritance. In general, the use of molecular techniques to understand evolution and inheritance is known as molecular ecology (Andrew *et al.*, 2013).

In early molecular ecology research, evolution and genetic variability were studied by observing physical characteristics. This manner of study was exhibited by pioneering scientists such as Darwin and Mendel. These early scientists equated appearance variation with genetic variation. However, it was later demonstrated that closely-related species or members of the same species show few, if any, physical differences resulting from genetic variation alone; environmental pressures also contribute to physical differences (Beebee & Rowe, 2003; Grant & Grant, 2008).

To overcome the limitations of these early studies, scientists turned their attention to other forms of variation. Researchers began studying similarities in chemical products (sex pheromones, defense compounds, chemical cues, etc.) as predictors of genetic inheritance and relatedness (Beebee & Rowe, 2003). This method gave scientists a way to study subtle differences between species. However, studying chemical products presented new limitations, opposite to those of earlier studies. Only a small number of variations can be made to a chemical product without changing its function, so these chemicals rarely (if ever) vary within a species

(Beebee & Rowe, 2003). In other words, their level of polymorphism is often too low to use as the basis for studying inheritance within a single species.

Scientists needed a method for studying molecular ecology that was sensitive enough to study evolution within a species, but not so sensitive that it could be influenced by environmental factors. The solution to this problem came with the discovery of molecular markers. Molecular markers are small sections of the genome chosen to represent the individual organism, and they may or may not be located within protein coding regions (Beebee & Rowe, 2003). There are many different types of molecular markers with varying degrees of polymorphism, and each can have advantages and disadvantages within specific studies (Avise, 1994). For example, parentage analysis benefits greatly from high levels of polymorphism, because this ensures enough diversity that similarities are most likely due to familial relationship rather than mere chance (Avise, 1994). On the other hand, moderate polymorphism is more useful for population studies, where fewer experimental groups with larger numbers are required (Beebee & Rowe, 2003). Molecular markers have become a favored method for studying evolution due to their availability, variability, and adaptability for different studies.

One molecular marker in use today is the microsatellite. Microsatellites are short sequences of nucleotides in tandem repeats (Hearne *et al*., 1992). They have many advantages over other molecular markers, including simple quantification (compared to some phenotypic markers), abundance in the genome (compared to allozymes), and high levels of polymorphism, which make them effective for distinguishing organisms of the same species (compared to many chemical markers) (Andrew, 2013). However, their high level of polymorphism is not suitable for all research endeavors. For instance, microsatellite sequences are not useful for long-term or

interspecies studies because they produce too much variation over generations and between species.

For this project, we have studied microsatellite loci in three local species of crayfish: *Procambarus clarkii*, *Orconectes limosus* and *O. quinebaugensis*. The primary goal of our study was to verify the presence of microsatellite loci in *P. clarkii* using primer pairs identified in previous research (Zue & Yue, 2008; Belfiore & May, 2000). After empirically confirming the presence of three of the loci, we explored the applicability of those three primer pairs in the two other species of crayfish, *O. limosus* and *O. quinebaugensis*. Resulting products from *P. clarkii* were sent to the Cornell University Institute of Biotechnology for fragment analysis. The data received from this analysis proved to be inconclusive.

2. Background

Microsatellites have recently emerged as one of the most widely-used molecular markers for certain questions, due to the fact that they are highly polymorphic, codominant, and replicable (Jones *et al*., 2010). These markers have allowed researchers to not only distinguish relatedness between species, but relatedness between individuals *within* a species.

2.1 Microsatellites

Nearly all known eukaryotic genomes contain microsatellites. Also known as simple sequence repeats (SSRs), microsatellites generally span about 1-6 base pairs (e.g. CACACA or GCGC). They are found in large quantities throughout the genome, and can be contained within both coding and noncoding regions of the DNA (Hearne *et al.*, 1992). Microsatellites in their simplest form have three different variations in the locus: pure (e.g. ATATATATAT),

compounded (e.g. ATATGCGCGC), or interrupted (e.g. ATGCATGCATAT) (Jarne & Lagoda, 1996). These variations can combine with one another to form nucleotide repeats. Dinucleotide repeats are the most common, and are usually found in the noncoding region of the genome (Selkoe & Toonen, 2006; Li *et al.*, 2002). However, when repeats occur in the coding region, they are associated with disease (Selkoe & Toonen, 2006; Li *et al.*, 2002). Microsatellite sequences often repeat 5 to 40 times; in rare cases, they can be longer (Selkoe & Toonen, 2006; Li *et al.*, 2002).

Microsatellites stand apart from other genetic markers due to their high rate of mutation. It has been published that microsatellites have an average mutation rate per locus per generation of 5×10^{-4}, which is high in comparison to the normal mutation rate in humans, 2.5×10^{-8} mutations per nucleotide (Selkoe & Toonen, 2006; Nachman & Crowell, 2000). Due to the high mutation rates of microsatellite regions, large allelic diversity can be created over only a few generations. Scientists can study how a microsatellite locus varies between individuals through the use of DNA fragment analysis and other molecular techniques. With this information, researchers are able to study ecological topics such as parentage, diversification, and environmental effects on population (Monsen-Collar & Dolcemascolo, 2010). Different microsatellite loci can be selected for different types of studies; questions about gene flow are best answered using a locus which has a low mutation rate while questions about demography will need a locus with a high mutation rate (Selkoe & Toonen, 2006).

While high mutation rates can be useful in many fields of study, they simultaneously make higher level systematics questions more challenging to answer. One issue that arises is distinguishing between microsatellites of identical length that have different sequences. These are known as homoplastic sequences (Estoup *et al*., 2002; Selkoe & Toonen, 2006). There are

two types of homoplasy: detectable and undetectable. Detectable homoplasy occurs when a mutation forms a new allele which is the same size as the original. Thus, the alleles appear the same in gel electrophoresis, but sequence analysis reveals a difference. Undetectable homoplasy is when two alleles are identical in both length and sequence, but have different genetic origins. These can occur in two ways. First, two identical alleles can be created through back-mutation (reverting back to a previous form). Second, mutations in two unrelated alleles can cause them to become the same length and sequence through different means (Selkoe & Toonen, 2006). Overall, homoplasy can depress the allelic diversity of populations and inflate estimates of gene flow (Selkoe & Toonen, 2006). However, homoplasy only becomes a significant concern when researching large populations or a highly divergent group. Estoup *et al*. describe how to address homoplasy post-discovery, allowing for a correction in statistical evaluations (2002). Homoplasy provides one example of why microsatellite divergence between species is difficult to determine.

The high mutation rate of microsatellites makes them highly polymorphic. In other words, many distinct sequences can be identified within the same population. Many scientists speculate that polymorphism is a result of slippage during DNA replication, though this is still under some debate (Zane, 2002; Jarne & Lagoda, 1996; Dewoody *et al*., 2006). Slippage is known to occur during DNA replication due to tandem repeats as a new strand is replicated. Replication slippage occurs when the DNA polymerase accidently reads a portion of the template strand that was previously assembled. This process is illustrated in Figure 1 below.

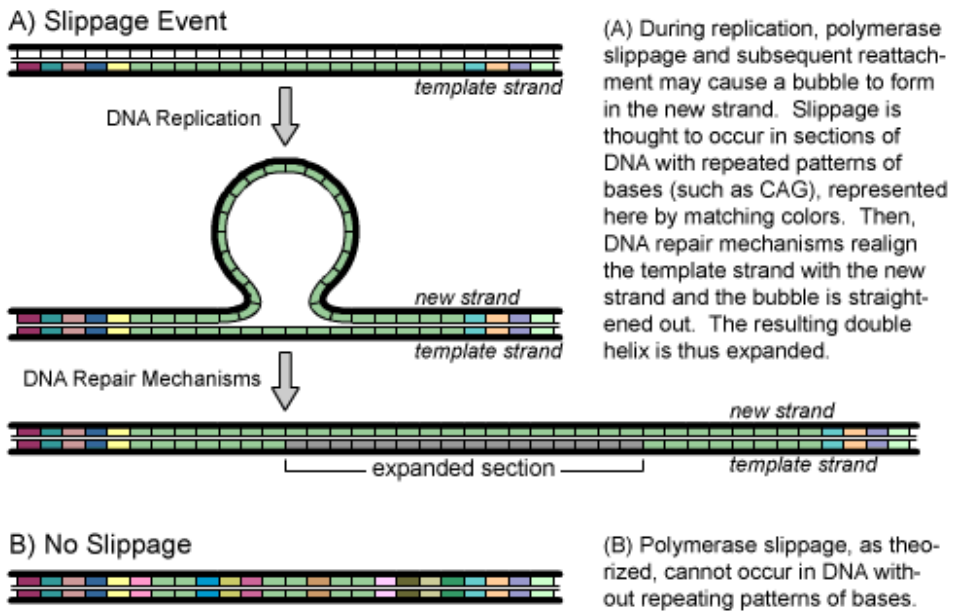

Figure 1: Diagram of slippage event (Stenerson, 2012)

This addition causes the daughter and template strand to be mismatched, so nucleotide excision repair proteins are used to either elongate or shorten the strand so they can match (Fazekas *et al.*, 2010). This slippage leads to polymorphism, which can be studied to better understand how a species' lineage changes over time.

A third noteworthy feature of microsatellites are the conserved flanking regions found on either side of the tandem repeat. These flanking sequence regions allow for primers to be designed for a specific species (Zane, 2002; Jarne & Lagoda, 1996; Dewoody *et al*., 2006). Some primer sequences are even conserved between species, allowing scientists to investigate similar microsatellites in a closely related ancestor or cousin of the species of study (Selkoe & Toonen, 2006). Therefore, the flanking region is important for generating a primer sequence that captures a specific microsatellite locus exclusively.

Scientists can use microsatellites to address a number of questions regarding genome sequence, genetic diseases, forensic analysis, and other topics in biology due to their high specificity and inheritance patterns. The work completed in this project and similar studies have created groundwork for understanding many biological factors such as relatedness and levels of inbreeding, genome differences within a population, population size estimation, gene flow through lineage, and phylogeographic studies (Blounin *et al.*, 1998; Blounin, 2003). For example, Wieczorek and Geber (2002) researched microsatellite loci in *Solidago semperviren* (seaside goldenrod) and were able to detect a highly polymorphic population in Delaware Bay, USA. That study displayed cross-amplification of microsatellites working across 11 species of seaside goldenrod, aiding in phylogeographic studies of this group. In another study, Clementino *et al.* (2010) were able to study chicken ecotypes and identify lineage within Brazilian chickens. From their data, they were able to determine polymorphic microsatellite sites that can be used to study genetic variability in the chickens and provide resources for future animal breeding programs (Clementino *et al.*, 2010).

Similar studies have been performed using various species of crayfish. For instance, after investigating 15 wild *Orconectes placidus* broods, Walker *et al.* (2002) discovered that most females had multiple mates (usually two), and the numbers of offspring in a brood were often skewed toward one father. This was verified by assaying the population using three polymorphic microsatellite loci and analyzing the results using parentage analysis software. Overall, their data resulted in a new understanding of breeding and offspring care in *O. placidus* (Walker *et al.*, 2002). Lastly, Gouin *et al*. (2006) studied the endangered crayfish species *Austropotamobius pallipes* to better understand their limited population in France. After analyzing five microsatellite loci in 44 different populations, they discovered that there are two distinct groups

of *A. pallipes* located in France, and a strong suggestion of an ancient divergence between them. From this data, they proposed a demographic management plan so that *A. pallipes* can be sustained for years to come (Gouin *et al.*, 2006). These studies represent only a small portion of the molecular ecology research being done using microsatellites, and they continue to be a popular molecular tool for new fields of study.

2.1.1 Comparison of Microsatellites to Other Molecular Markers

As scientists continue to study genetics, they also continue to develop different techniques. Of the other genetic markers available, such as allozymes, randomly amplified polymorphic DNA (RAPDs), and minisatellites (also known as DNA fingerprinting), microsatellites are the preferred genetic marker in understanding parentage across multiple generations and answering questions in behavioral ecology. Definitions for each of these molecular markers can be found in Table 1 (Beebee & Rowe, 2003).

Table 1: Molecular marker definitions

Molecular Marker	Definition
Allozymes	Distinct forms of an enzyme encoded by different alleles at the same locus (Bader, 1998). To study genetic variation using allozymes, a protein solution is electrophoresed through a cellulose acetate gel. Since different forms of an enzyme have different charges, the allozymes will travel through the gel at different rates and can be visualized to identify individual variation.
Restriction Fragment Length Polymorphism (RFLP)	Molecular marker technique that digests DNA sequences through the use of restriction endonucleases. DNA variation is detected in restriction sites by different patterns of fragment lengths. (Beebee & Rowe, 2003).
Randomly Amplified Polymorphic DNA (RAPD)	Markers that use an arbitrary primer (about 10 bases long) to amplify many different sites throughout the genome. Fragment size variation is visualized and scored through gel electrophoresis (Jarne & Lagoda, 1996).
Minisatellites (DNA Fingerprinting)	A molecular marker technique that quantifies variation in variable number tandem repeats (VNTRs) using purified DNA that has been cut using restriction enzymes. The fragments are loaded into a gel and electrophoresed in order to separate larger fragments from smaller ones. Afterwards a blotting technique renders the dsDNA to ssDNA so autoradiography can be performed. (Beebee & Rowe, 2003).

The primary reason why microsatellites are preferred over other genetic markers is their specificity. Jarne and Lagoda (1996) compared and contrasted the uses of microsatellites against both allozymes and randomly amplified polymorphic DNA (RAPDs). Ultimately, they concluded that allozymes and RAPDs cannot offer the level of detail that microsatellites do.

Microsatellites offer a wealth of information that other genomic components cannot. Queller *et al*. (1993) detailed some of the reasons for this. DNA can be difficult to extract or degraded to a point where other molecular marker tests become unreliable. By employing the power of microsatellites, even ancient DNA can be analyzed due to the presence of microsatellites throughout the entire genome (Queller *et al.*, 1993). More specifically, microsatellites are found in large numbers in the noncoding region of the genome and sometimes in the coding region. Therefore, even a small amount of DNA is likely to contain a microsatellite locus.

The quality and quantity of information provided by microsatellites allows for consistent scorable bands compared to other molecular markers. A consistent and scorable band is a piece of DNA that will appear in a gel in the same relative spot with the same intensity over multiple trials. With other molecular markers, such as minisatellites and RAPDS, bands may appear in multiple places with varying intensity, causing some scientists to discount a faint band (Queller *et al.*, 1993). In microsatellites, the presence of any band (after primers and contamination have been ruled out) theoretically indicates a microsatellite locus.

Another challenge that arises when using molecular markers is comparing electrophoresis bands. Employing a standard on a gel allows for general quantification of the band size, which helps identify different alleles. However, if there are several alleles that only vary in size by a few nucleotides, identification becomes more difficult, and a more accurate analysis method is necessary (e.g. sequencing or fragment analysis). Minisatellites and RAPDS share this difficulty (Queller *et al.*, 1993). Allozymes and RFLPs have few alleles to analyze, so it can be assumed that the same allele is duplicated when multiple bands appear (Queller *et al*., 1993).

Another issue that must be addressed when deciding between molecular markers is detecting codominance. Codominance is a biological condition where two alleles in a gene are fully expressed in a heterozygote. Minisatellites and RAPDS only display bands of dominant alleles. When codominance is a factor, bands will appear to be homozygous, which means the two bands will only look like one on the gel (Queller *et al.*, 1993). Conversely, microsatellites, allozymes, and RFLPs are all able to easily display codominant alleles, making them the superior choice for molecular markers when codominance is of significance. Other markers however, can only show the dominate allele.

Lastly, molecular markers also vary regarding the quality and volume of DNA necessary. If DNA is not fresh, or the sample of DNA is very small, some molecular markers may not be effective. This is an issue for minisatellites (DNA fingerprinting), as minisatellite loci are found together in a specific region of alleles (Queller *et al.*, 1993). Other molecular markers run into a similar problem; both RAPDS and VNTRs need long target sequences before amplification can be performed, leading to questionable results when minimal high quality DNA is available (Queller *et al.*, 1993). However, it has been shown that DNA from an 1850-year-old Egyptian mummy can be amplified using microsatellite primer sequences (Queller *et al.*, 1993). This demonstrates that microsatellites can be successful with smaller amounts and lower quality of DNA.

2.1.2 Common challenges with Microsatellites

Despite their broad capabilities, there are inherent challenges specific to microsatellites. Zane *et al*. (2002) points out one major concern with microsatellites: there is currently no way to discover new microsatellite loci in a species without an extensive analysis of its entire genome (Zane *et al.*, 2002). This can be an expensive and time-consuming process. Furthermore, as generations are analyzed, the original primer for one microsatellite locus might not be as effective due to mutations in the noncoding region. Since the nucleotide substitution rate is higher in the noncoding region (where most microsatellites are found) than the coding region, microsatellite loci are likely to be mutated (Zane *et al.*, 2002); however, these mutation rates are still comparatively low, and thus do not frequently affect the binding capability of the primer.

In recent studies, there have been strides in rapid and cost effective analysis of new genomes. Csencics *et al.* (2010) were able to create 17 polymorphic microsatellite markers and

test them against 20 individual endangered dwarf bulrush (*Typha minima*) in less than 6 weeks for about $5000 US. In another study by Abdelkrim *et al*. (2009), high-throughput genomic sequencing technology and bioinformatics toolsets were used to sequence millions of base pairs. From this mass sequencing, 13 polymorphic microsatellites were identified for a species which had little genetic information available (Abdelkrim *et al.*, 2009). Overall, as microsatellite techniques improve, more cost effective options will be found to design primers that work on previously under-researched subjects or those who have mutated too much for the original primer to work.

The next concern for microsatellites is poor amplification results, whether from allelic dropout, contamination, or poor optimization of the primers. Allelic dropout occurs when an allele is known to be present in genome but fails to produce in the products. This is thought to be caused by too much or too little DNA, a mutation in the primer binding site, or the allele size being unexpected (Beebee and Rowe, 2003). Bonin *et al*. (2004) agree that while the science behind microsatellites can lead to error, human factors can also lead to large amounts of laboratory errors. In their study, they examined amplified fragment length polymorphisms (AFLPs) and microsatellites and common errors associated with them. From their research, it was concluded that improvement would require better precautions against contamination and technical artifacts, an addition of blind sampling and/or automation, training and rigor for all work and scoring methods, and a systematic way to report error rate in population genetic studies (Bonin, 2004). While these seem like obvious solutions to many problems with laboratory experiments, implementing a system would increase positive results and decrease wasted resources. In addition, Bonin *et al.* also report that most of the allelic dropout discovered was due to the quality and quantity of the DNA from noninvasive samples and technical challenges in

amplification of invasive samples. Overall, it was agreed upon that the best method to prevent this was accuracy of genotyping, whether through repeated genotyping of the same/similar samples or use of a different genetic marker (Bonin, 2004).

Microsatellites also face technical challenges in addition to the experimental challenges described. Dewoody *et al*. (2006) highlighted three common causes of error that can affect downstream analysis of the data: stuttering patterns, large allele dropout, and null alleles. Another error which Dewoody did not mention was false alleles, or alleles that are not present in the genome but appear in the PCR products (Broquet & Petit, 2004).

Stutter bands appear in gels as several close but distinct bands. They may indicate a dissociation of the *Taq* Polymerase from the template strand, nonspecific primer site sequence, or poor quality of the DNA (Fazekas *et al.*, 2010). Walsh *et al.* (1996) define stutter bands more precisely as the amplification of a tetranucleotide short tandem repeat, yielding a product that is usually 4 base pairs shorter than the corresponding allele. Shinde *et al.*, (2003) explored how often *Taq* DNA polymerase slips by measuring it through PCR and quasi-likelihood analysis. They discovered that the rate of insertions and deletions for "CA" repeats was 3.6×10^{-3} mutations/repeat/PCR cycle, and contractions were 14 times greater than expansions. Meanwhile, "A" repeats had 1.5×10^{-2} mutations/repeats/PCR cycle with 5 fold more expansions than contractions (Shinde *et al*., 2003).

As shown in Figure 2, Lima *et al.* (2003) encountered stutter bands in their microsatellite examination of polymorphism in wheat. They attributed the stutter bands to very similar alleles which only varied by a single repeat unit.

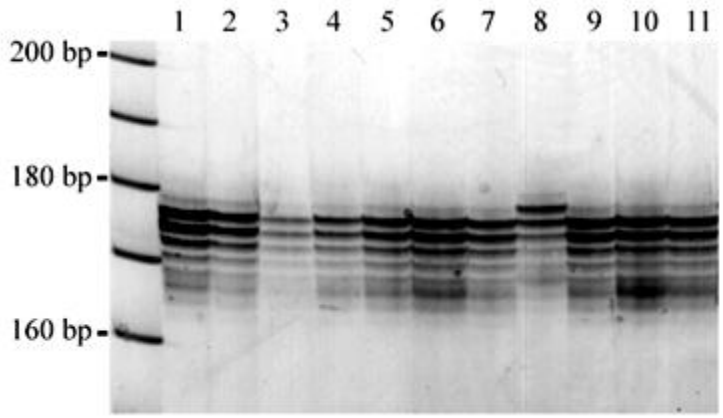

Figure 2: Example of microsatellite stutter bands (Lima *et al.*, 2003)

As shown in Figure 2**Error! Reference source not found.**, stutter bands are a common issue with microsatellites, and there are currently no known ways to remove them from the PCR product. The stutter patterns vary across loci; some markers may produce a small stutter pattern, while others produce large stutter peaks (Zane, 2002; Jarne & Lagoda, 1996; Dewoody *et al.*, 2006). The type of microsatellite sequence can affect the stutter pattern as well. There appears to be a positive relationship between stutter products and uninterrupted core repeat units; as the number of tandem repeats increases, so does the chance of stutter band product (Walsh *et al.*, 1996). When a stutter pattern occurs, it can lead to misinterpretation of the data in two distinct ways. First, it can cause two close but separate alleles to appear as one on the gel, which causes an increase in homozygosity (Dewoody *et al.*, 2006). Conversely, if a homozygous allele stutters, it can be interpreted as heterozygous.

Researchers have been actively looking for a resolution to this issue. Miller and Yuan (1997) were able to successful predict the shape of a stutter pattern both theoretically and experimentally. They accomplished this task by making the assumption that there is a chance to lose or gain a microsatellite repeat during every PCR cycle; a conclusions which they came after realizing *Taq* polymerase either can skip a nucleotide doublet, the PCR product increases due to

non-templated addition to the 3' terminus or a slippage of the strands during PCR (Miller *et al.*, 1997). From this they further deduced that the cumulative effect of losing a repeat is greater than gaining one as cycle number increases; ergo a band appears to stutter as the product shrinks. The model can predict stutter bands on heterozygous and homozygous alleles as they tested several known microsatellites pairing over four different cycle patterns. However, their solution does not offer way to remove stutter bands and further research is required to find a way to prevent stuttering.

Another challenge in using microsatellites is large allele dropout. This occurs when one allele in a heterozygous genotype is preferentially amplified. This will make a heterozygous individual appear homozygous when the data is analyzed, leading to potential parentage errors (Jones, 2010). Current research shows that low-quality DNA or allele size may be the main causes of large-allele dropout (Dewoody *et al*., 2006; Wang *et al.*, 2012). To show that allele dropout has occurred, data will need to be compared to reference genotypes from other sources of DNA or repeated analysis of the original sample (Broquet & Petit, 2004). Therefore, when allelic dropout occurs, it will cause overestimation and omission of alleles from data (Dewoody *et al.*, 2006). Figure 3 (below) displays the two types of allelic dropout that can occur. This issue can be solved by repeated replication of the protocol while minimizing experimental error until the allele reappears (Wang *et al.*, 2012; Broquet & Petit, 2004). Wang *et al.* (2012) also developed a method to correct for allelic dropout in microsatellite data with no replicate genotypes. By using a formula derived from the Hardy-Weinberg equilibrium, they can re-produce missing data, estimate model parameters, and correct the negative bias observed heterozygosis (Wang *et al.*, 2012).

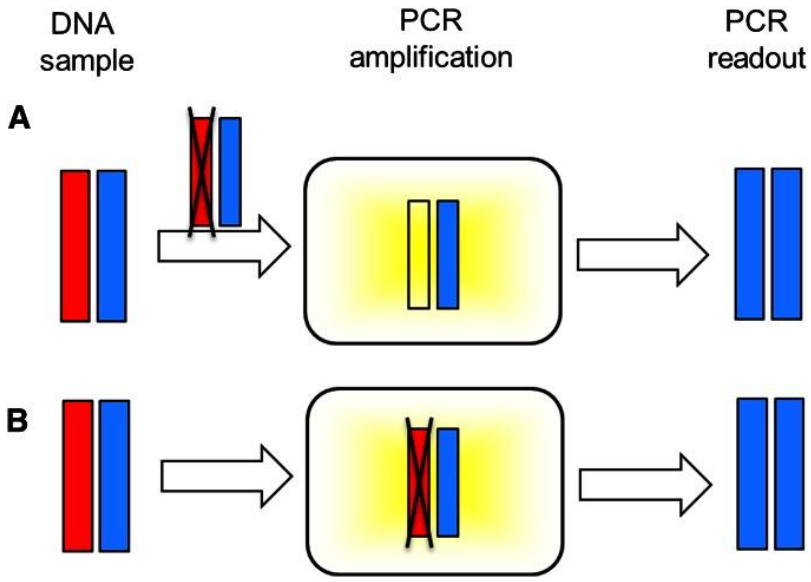

Figure 3: Two stages of allelic dropout (Wang *et al.* 2012)

Figure 2 displays the two different stages of allelic dropout and drawn by Wang *et al.* (2012). As described, the red and blue bars are different alleles of a locus in a heterozygous individual. The black X indicates allelic dropout. In individual A there was an issue with the quality or quantity of the DNA, leading to a dropout before PCR amplification, while in individual B there may be a low binding affinity for the red allele. Overall, both issues will result in an apparent homozygous genotype (Wang *et al.*, 2012).

False alleles are, in some ways, the opposite of allelic dropout. A false allele is an allele that was not initially present in the organism, but was created during the PCR process due to slippage (Broquet & Petit, 2004). New alleles can appear as a result of contamination from other organisms during the PCR process (Gerber, 2000). False alleles usually occur early in the stages of the PCR cycles and therefore can be prevented by minimizing experimental error and contamination (Broquet & Petit, 2004). These unnatural alleles lead to mistakes in analyzing homozygous individuals, as a false allele may be mistaken for a second allele (Broquet & Petit, 2004).

Lastly, null alleles occur when a product completely fails to form. Due to the nature of null alleles, it is very difficult to detect whether a null allele exists, or whether the target sequence did not exist in the sample. True null alleles occur when there is a mutation in the primer site. When a null allele is present in a data set but undetected, data will be biased towards visible alleles being overestimated, increased homozygosity and increased apparent levels of inbreeding (Dewoody *et al*., 2006). As alleles disappear, homozygosity levels will increase, which inflates the perceived degree of inbreeding (Girard, 2011). Therefore, to counter null alleles, techniques similar to combating large allele dropout must be employed (which might include reordering/redesigning primers).

2.2 DNA Fragment Analysis

While microsatellites can be amplified and quickly visualized on a gel, this method has limited resolution and is very prone to error. Considerably more information can be garnered through determining the nucleotide sequence of each sample. However, sequencing is costly and time-consuming. A third method is fragment analysis, which quickly and accurately measures fragment size, though without providing other information about the locus that sequencing might provide. By labeling the amplified product with a fluorescent primer and running the results through a capillary, the overall size of each fragment can be determined without reading every nucleotide (Comparative Genomics Center, 2012). Those lengths can then be compared to learn more. For example, two fragments of different lengths from the same sample indicate heterozygosity; variations in fragment length across several samples might indicate a wider variety of alleles throughout the species. Fragment analysis is much faster than sequencing, but

less error-prone than gel electrophoresis, making it the most commonly-used method of basic microsatellite analysis.

Fragment analysis is not as comprehensive as complete sequencing, but it is fast, accurate, and inexpensive. This makes it ideal for studying microsatellites. Because mutations in microsatellites usually appear as an increase or decrease in the number of repeats, resulting in a change in its length, measuring the length provides the necessary information to differentiate between alleles. The lower cost and greater speed allow for larger sample sizes and faster quantification than would be possible with sequencing, making fragment analysis the best option for studying microsatellites in many situations.

2.3 Our Project

Microsatellites are found universally in eukaryotic species. This broad applicability, as well as their cost-effectiveness and efficiency, make them a prime choice for studies in population genetics. Additionally, with the wealth of literature about them, there are many opportunities to follow-up on previous studies and learn more about how these markers work. For these reasons, and the advantages described previously, this study used microsatellites instead of another molecular marker.

For a variety of reasons, we chose to focus on microsatellites in crayfish. North America contains 70% of the world's 540 recognized crayfish species (Fetzner & Crandall 2002). Of these 350 species, this project will focus on three species of crayfish that are found locally in central Massachusetts: *Procambarus clarkii*, *Orconectes limosus*, and *Orconectes quinebaugensis*. In this study, we hope to create an effective technique that will allow us to use the same primers across multiple similar species.

Belfiore and May (2000) studied the red swamp crayfish, *P. clarkii*, in attempt to better understand their characterization to other species. In total, they optimized 23 variable microsatellite loci from their original library where 18 of the clearest markers were tested against two other species. While not every primer worked in each species, alleles were reported from each primer pair.

Zhu and Yue (2008) isolated eleven polymorphic microsatellite loci from enriched CA and GA repeat regions, and were able to characterize them in 48 individual *P. clarkii.* Of the eleven which displayed polymorphism, there was a decrease in the number of alleles from previously researched *P. clarkii* loci. The authors explained this as a natural change of environment due to inbreeding within the invasive *P. clarkii* population in China (Zhu & Yue, 2008).

A similar study by Hulak *et al.*, (2010) attempted to amplify existing microsatellite loci from other crayfish in *O. limosus*. Of the microsatellite loci they used, 10 successfully amplified loci in *O. limosus*. Data was then analyzed to discuss heterozygosity levels, alleles per locus, population structure, and paternity studies among other population genetics.

The purpose of our project was to assay 20 primers used for *P. clarkii,* and to determine whether these primers could also be used with DNA from two other species of crayfish found in Massachusetts. If optimized for use with all three species, these primers might then be utilized in subsequent investigations regarding phylogeography, evolutionary ecology, and paternity and mating behavior on both native and invasive crayfish species.

3. Methodology

3.1 Sample Collection

A total of 53 crayfish of three different species was collected from two sites in central Massachusetts: Institute Pond in Worcester, and East Brimfield Dam in Sturbridge. Of these crayfish, 6 were from *Procambarus clarkii*, 32 from *Orconectus limosus*, and 15 from *Orconectus quinebaugensis*. Live *P. clarkii* were also ordered from Carolina Biological Supply Company. Eighteen additional samples of *P. clarkii* DNA from 2009 were also included.

All live crayfish were frozen upon collection, and muscle tissue from the leg, claw or tail was removed. DNA was then extracted from the tissue using the Gentra Puregene Tissue Kit following the manufacturer's instructions. The resulting samples were visualized on a 1.2% agarose gel in TBE buffer with 0.1% of 500mg/ml ethidium bromide to assay the DNA for quantity and quality. Known amounts of undigested lambda DNA and Bioline Hyperladder I were utilized to estimate the amount of DNA in each sample.

Although extractions were attempted on all crayfish collected, not all extractions yielded DNA signals. Samples that failed to be successfully visualized were not used in our study. Table 2 indicates the species, location, and time of collection for all DNA samples, as well as the number of successful extractions.

Table 2: Summary of crayfish DNA sample collection

Year	Source Location	Species	# Crayfish	# Successful Extractions	Sample Names
2009	Institute Pond	*P. clarkii*	18	18	PA 1-18
2013	Institute Pond	*P. clarkii*	6	4	PC3-6
2013	East Brimfield Dam	*O. limosus*	32	6	OL 27-32
2013	East Brimfield Dam	*O. quinebaugensis*	15	9	OQ 1-8, 14-15
2014	Carolina Biological Supply	*P. clarkii*	12	12	PC17-28

3.2 Initial PCR

In order to practice PCR technique and verify that the DNA samples were of reasonable quality, PCR was run on six DNA samples (*P. clarkii* and *O. quinebaugensis*) using the 16S-1472 and 16S-L2 primer pair, which amplifies a locus found in the mitochondrial gene for the 16S ribosomal subunit (Noda *et al.*, 2001) The 16S locus has a previously been used in phylogenetic studies of various species, including many crustaceans (Bracken *et. al.*, 2009). Because it is known to exist in *P. clarkii*, weak results would indicate poor technique or DNA quality, rather than lack of locus. PCR was run on a BioRAD DNAEngine Peltier Thermal Cycler. The samples were heated to 95˚C for two minutes to initiate the reaction. They then underwent 40 cycles of denaturation (95˚C for 30 seconds), annealing (43˚C for 30 seconds), and elongation (72˚C for one minute). Finally, the samples were held at 72˚C for ten minutes.

Table 3 below indicates the concentrations of reagents in each PCR sample. The results were visualized on a 1.2% agarose gel with 0.1% of 500mg/ml ethidium bromide.

Table 3: Reagent concentrations for PCR

Reagent	Concentration
10X ThermoPol Reaction Buffer	1X
10mM dNTPs	200uM
10uM Forward Primer	0.2uM
10uM Reverse Primer	0.2uM
Taq DNA Polymerase	0.5 units/20uL PCR

3.3 Primer Testing PCR

Previous research was reviewed in an effort to compile a list of primers that amplified *P. clarkii* microsatellite loci. In total, 20 primer pairs were identified, including the 16S pair, based on research done by Zhu and Yue (2008) and Belfiore and May (2000). Table 4 displays these 20 primer pairs. All primers were ordered from Integrated DNA Technologies.

Table 4: Selected primer pairs from previous research. Melting temperature is the temperature at which the primer dissociates from anything it has bound. Annealing temperature is the temperature at which the primer binds to complimentary DNA, and was calculated by subtracting five from the lowest melting temperature in that pair.

Locus Name	Number of Bases	Sequence	Product size range (bp)	Melting Temperature (˚C)	Annealing Temperature (˚C)
PCL02*	19	GAA GAC GGG ACA CCA CGA G	245-271	56	45
EF564119*	24	ATC AAA TCA AAC GAA GCA AGA AAG		50	
PCL06*	23	GTG TCC ACC TTC CTC CAT AGT CA	224-266	57	52
EF564120*	21	AGT CAG CCT CCA CCA CAT CAC		58	
PCL11*	20	GGT TCC TGA GCC TAC TGG AC	156-186	56	47
EF564121*	20	AGC CGC ACC TTA ACA ATC TT		52	
PCL17*	21	TTC ACT ACC GCC CAC AGG ATG	370-422	58	52
EF564122*	24	TAG GCA CCG CAC TTA TTA GAC CAG		57	
PCL24*	26	CGT TAT TCT CTT TAT GTG GGC TTC AG	130-184	54	48
EF564123*	25	TTT CAA TTT TCC TCC GTG GTC TAA C		53	
PCL25*	25	GGG GAG ACA TGA TCA CAA CCT ACA A	172-178	57	52
EF564124*	21	GCT GCT ACG GGC TGC TTC CTA		61	
PCL28*	24	CCT ACC AGA GAA CCC AAA ACA GAA	229-239	55	50
EF564125*	22	GTC AGC CTC CAC CAC ATC ACT T		58	
PCL29*	21	CAC TCA AGC CTG CCC TCA CTC	170-214	59	53
EF564126*	23	GTC TCT TCC TCC CCC ATT CTC AC		58	

(continued below)

Locus Name	Number of Bases	Sequence	Product size range (bp)	Melting Temperature (˚C)	Annealing Temperature (˚C)
PCL33*	21	ACT CCT GTC CCA TTT CAC TAC	132-158	52	44
EF564127*	23	ACA ACT AAC TGC AAC TCA TTC TA		49	
PCL47*	21	ACT CTG CCC ATT GTT TCT CGG	324-344	56	51
EF564128*	21	AGC CCT TGG ACC CCG CCT ATC		63	
PCL50*	24	AAG CGC TGA AAT GCA CAA ACA AGA	424-524	56	51
EF564129*	21	CAA GCC CCG AGG TCA AAG GTC		63	
PclG-03F**	20	CTC TCC ACC AGT CAT TTC TT	216-420	49	38
PclG-03R**	25	AAG CTT ACA ATA AAT ATA GAT AGA C		43	
PclG-07F**	24	CCT CCC ACC AGG GTT ATC TAT TCA	100-160	56	51
PclG-07R**	20	GTG GGT GTG GCG CTC TTG TT		59	
PclG-08F**	24	ACG ATA AAT GGA TAG ATG GAT GAA	148-220	49	44
PclG-08R**	20	CCG GGT CTG TCT GTC TGT CA		57	
PclG-16F**	20	CTC GGA ATG TCC ACC TGA GA	80-160	54	39
PclG-16R**	24	TCA TTA TGG ATT TTG TCA ATC TAT		44	
PclG-26F**	22	ATA TAG CCT CGC CCT TTT ACC C	210-300	55	50
PclG-26R**	22	TCG TGT TCA CAT CAG CAG GAG A		56	
PclG-32F**	22	CCC CCA CTC GTC TCT GTG TAT G	150-250	58	53
PclG-32R**	19	TGT GCT TGC GGG AGT GAG C		59	
PclG-33F**	24	TTC GAG GCG TTG CTG ATT GTA AGT	120-180	57	52
PclG-33R**	24	CAA GGA AGC GTA TAG CCG GAG TCT		59	
PclG-37F**	24	TAA ATA AGT GGC GTG TAA GAC GAG	80-180	52	47
PclG-37R**	24	TAA CTA AGC CAG GGT GGT CTC CAG		59	
16S-1472	18	AGA TAG AAA CCA ACC TGG	550	45	39
16S-L2	20	TGC CTG TTT ATC AAA AAC AT		44	

*Indicates primers that were used from Zhu & Yue (2008)
**Indicates primers that were used from Belfiore & May (2000) and Hulak *et al.* (2010).
Melting Temperature was reported by the manufacturer of the primers and it is based on when the primer will dissociate and become a single strand DNA.
Annealing Temperature was calculated using several difference resources such as NFSTC and New England Biolabs Calculators. Overall, Annealing Temperature was determined to be 5 C° lower than the lowest Melting Temperature.

Three *P. clarkii* samples (PC3, PC5, and PC6) were selected by analyzing the results of the initial 16S PCR, and identifying the *P. clarkii* samples that showed the strongest PCR products. These three samples were used to test each of the 19 primer pairs through several iterations of PCR to confirm their efficacy with *P. clarkii.*

PCR was carried out for each primer pair using the same concentrations and PCR conditions used for the 16S locus. A negative control using autoclaved deionized water in place of DNA was also included for each primer pair. From these results, eight were selected to be re-run on a 2% Amresco 3:1 High Resolution Blend agarose gel to achieve better resolution. The reactions were then run twice more on the same DNA samples, this time using an optimum annealing temperature specific to each primer pair (see Figure 9, Figure *10*, and Figure *11*) during the annealing stage instead of a single annealing temperature for all primers. This was done to optimize the PCR conditions for each primer pair, with the hopes of producing stronger and more distinct results. The resulting PCR samples were visualized on 2% agarose gel with 0.1% of 500mg/ml ethidium bromide.

The results of these three iterations of primer-testing PCR were analyzed, and the primers were placed into three different categories based on both the strength and specificity of the products. We evaluated "strength" according to the brightness of bands; a bright, solid band was considered very strong, while a faint or fuzzy band was considered weak. Meanwhile, we evaluated "specificity" based on the number of products, and how distinct each one was. One or two distinct products indicated high specificity, whereas multiple products, products blurring together, or noticeable streaking on the gel indicated poor specificity, and were considered inconclusive. Figure 4 below illustrates examples of strong, specific results compared to poor, inconclusive results.

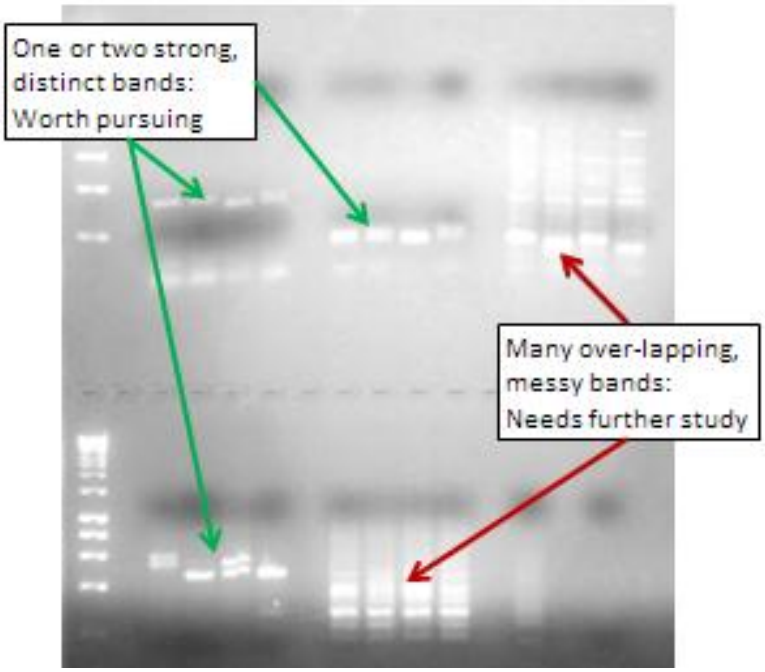

Figure 4: Comparison of strong, specific results to weak, non-specific results

Eliminated primers were those that showed very weak or inconclusive results. Potential Primers of Promise were partially successful, showing moderately strong results, or strong results that were inconclusive; they required more testing to determine whether they were worth pursuing. Finally, true Primers of Promise showed very strong, distinct results with only one or two bands, and were deemed to be worth testing on a wider range of DNA samples.

3.4 Gradient Testing on Potential Primers of Promise

Eight pairs of Potential Primers of Promise underwent further analysis to determine whether they were worth pursuing. Specifically, the goal was to optimize PCR using these primers to produce strong and specific results, such as the true Primers of Promise showed. For each pair, sixteen PCR reactions were prepared, half using the DNA from sample PC5 and half using the DNA from sample PC6. These were run on a temperature gradient such that each

sample underwent a different annealing temperature, with the lowest temperature at about 2˚C below the original annealing temperature and the highest at 5˚C higher than the original annealing temperature. This was done to determine the annealing temperature which would produce the best results. Too high of an annealing temperature would render the primer unable to anneal to the DNA, resulting in no product; too low of an annealing temperature would allow for nonspecific annealing, resulting in nonspecific, inconclusive products. The resulting samples were visualized on a 2% agarose gel with 0.1% of 500mg/ml ethidium bromide. The two highest annealing temperature samples from three primer pairs (PCL-G37 F/R, PCL-29/EF564126, and PCL-G32 F/R) were also run on a 4% 3:1 Amresco High Resolution Blend agarose gel for better resolution. Each pair of Potential Primers of Promise was re-categorized to either Eliminated or Primers of Promise according to the criteria defined in Section 3.3.

3.5 Primers of Promise Tested on All *Procambarus clarkii* DNA Samples

After confirming that the Primers of Promise successfully amplified the three *P. clarkii* samples used in original primer testing (PC3, PC5, PC6), the next step was to confirm that these pairs would successfully amplify other *P. clarkii* samples. To accomplish this, PCR was run for each pair on 9 *P. clarkii* samples, using the annealing temperature that we had identified as optimal for that pair (based on initial primer testing and gradient testing on Potential Primers of Promise). To achieve cleaner results, one unit of PicoMaxx High Fidelity PCR System polymerase (Agilent Technologies) was used per 20ul reaction, and the concentration of primers was increased to 0.3uM. The PicoMaxx system combines two types of polymerases (*Taq* and Pfu) with ArchaeMaxx polymerase-enhancing factor for high sensitivity to produce strong results. It achieves high specificity by inhibiting polymerase activity until cycling begins

(Agilent, 2014). This improved sensitivity and specificity of the enzyme would allow for better differentiation between strong, specific results and those that were inconclusive.

The results were visualized on a 2% agarose gel with 0.1% of 500mg/ml ethidium bromide. Again, the primer pairs were further categorized based on the strength and specificity of the results. Those with weaker or inconclusive results were set aside to be further explored in future studies, while those with stronger results were carried over to the next step.

3.6 Primers of Promise Tested on *Orconectes limosus* and *Orconectes quinebaugensis* DNA Samples

After identifying the primers that gave the strongest results in *P. clarkii,* the next step was to test whether those primers would amplify similar loci in *O. limosus* and *O. quinebaugensis.* PCR was performed on all *O. limosus* and *O. quinebaugensis* samples using the Primers of Promise and the same conditions that were used for *P. clarkii* (as described in Section 3.5). The results were visualized on a 2% agarose gel to determine which samples were successful.

3.7 PCR Product Fragment Analysis

Once it was confirmed that the three Primers of Promise successfully amplified products from *O. limosus* and *O. quinebaugensis* as well as *P. clarkii,* PCR was run on a two DNA samples from each species using primers labeled with fluorescent tags. All fluorescent primers were ordered from IDT. PclG33F was modified with a 5’ 6-5FAM (519 nm). Pcl50 was modified with a 5’ HEX (556 nm). PclG26F was modified with 5’TET (536 nm). This PCR was

run using optimal annealing temperatures for each primer pair (as determined in Section 3.4), using PicoMaxx and the concentrations described in Section 3.5. The resulting samples were visualized on a 2% agarose gel with 0.1% of 500mg/ml ethidium bromide to confirm that PCR was successful. Further optimization of conditions for using fluorescent primers, included doubling the amount of primer and increasing the number of cycles to 45. After running a test gel to confirm the presence of product and lack of contamination 1uL of PCR product and 0.2 uL of LIZ 600 size standard was added to a 96 well plate; that is 1 uL per PCR product and 0.2 uL of LIZ 600 size standard was placed in each well amounting up to the number of samples that were prepared. Wells containing samples were brought up to 10 uL using formamide while all other wells had 10 uL of formamide added. The 96 well plate was sent to the Cornell University Institute of Biotechnology for automated fragment analysis. Applied Biosystems PeakScanner software 1.0 was used to analyze the results. 4. Results

4. Results

This section outlines our efforts to confirm the effectiveness of 20 microsatellite primers previously developed for *P. clarkii*. Our ultimate goal was to test the ability of these primers to amplify microsatellites in two other species of crayfish: *O. limosus* and *O. quinebaugensis*.

Table 5 provides a summary of all gels included in this section, their purpose, and their significance.

Table 5: Summary of all gels and their significant findings

Figure	DNA Samples*	Primers	Optimization features	Purpose	Key Findings
Figure 5	PC1 OQ1 OQ5 OQ9	16S	None	PCR technique confirmation Reagent quality control	16S successfully amplified samples of *P. clarkii* and *O. quinebaugensis* DNA, confirming PCR technique and reagent quality
Figure 6	PC3 PC5 PC6	Pcl50 PclG03 PclG07 PclG08 PclG16 PclG26 PclG32 PclG33 PclG37	None	Initial testing of primer pairs from previous studies on *P. clarkii* DNA	PclG37 and Pcl50 produced strong results. Pcl32 and Pcl33 also produced results.
Figure 7		Pcl25 Pcl28 Pcl29 Pcl33 Pcl47 PclG26 PclG32 PclG33 PclG37	None		Pcl25, Pcl28, Pcl29, and Pcl47 produced strong results. Pcl33 also produced results.
Figure 8		Pcl02 Pcl03 Pcl07 Pcl08 Pcl16 Pcl11 Pcl17 Pcl50	None		Pcl50 produced strong. Pcl16, Pcl02, and Pcl11 also produced results.
Figure 9	PC3 PC5 PC6	Pcl06 Pcl16 Pcl25 Pcl28 Pcl32 Pcl33 Pcl47 Pcl50 PclG26 PclG33	Annealing temperatures specific to each primer pair	Primer testing using optimal annealing temperatures to clarify the results of previous testing	Pcl25, Pcl16, Pcl28, and Pcl47 produced strong results.
Figure 10		Pcl50 PclG03 PclG07 PclG08 PclG16 PclG26 PclG32 PclG33 PclG37			PclG26, PclG33, PclG37, and Pcl16 produced results.
Figure 11		Pcl02 Pcl06 Pcl11 Pcl17 Pcl24 Pcl25 Pcl28 Pcl29 Pcl33 Pcl47			Pcl25, Pcl28, Pcl29 and Pcl47 produced results.

(Continued below)

Figure	DNA Samples*	Primers	Optimization features	Purpose	Key Findings
Figure 12	PC5 PC6	Pcl29 PclG32	Annealing temperature gradient	Re-test Primers of Promise using an annealing temperature gradient to determine whether a particular annealing temperature would yield stronger results	No significant results for either primer pair.
Figure 13		PclG37 Pcl08			PclG37 successfully amplified DNA at all temperatures. Pcl08 failed to amplify.
Figure 14		Pcl25 Pcl47			Both primers yielded product, but no difference was observed between temperatures
Figure 15		Pcl28 Pcl50			Pcl28 yielded results. Pcl50 failed to produce strong results.
Figure 16	PC5 PC6 PA4 PA6	Pcl50 Pcl28 PclG26 PclG33 PclG37	Annealing temperatures specific to each primer pair	Testing of Primer of Promise Samples to see what would be the best candidates for fragment analysis	PclG26, PclG33, Pcl50 displayed the cleanest bands out of all the products where heterozygosity can be assumed due to separate bands. PclG37 and Pcl28 need further optimization for cleaner results.
Figure 17	PC5 PC6 PA1 PA2 PA4 PA6 PA7 PA9	Pcl50 PclG26 PclG33	Annealing temperatures specific to each primer pair	Testing of all viable *P. clarkii* samples across successful primers	PclG33 showed the strongest bands across all primers. Pcl50 then PclG26 displayed strong bands as well.
Figure 18	OQ5 OQ6 OQ8 OL28 OL29 OL30	PCl 50 PclG26 PclG33 16S	Annealing temperatures specific to each primer pair	Testing whether *O. quinebaugensis* and *O. limosus* contained microsatellite loci similar to *P. clarkii*	PclG26 showed weak bands for both *O. quinebaugensis* and *O. limosus.*
Figure 19	PC5 PC6	Pcl50 PCl50Fl** PclG26 PclG26Fl PclG33 PclG33Fl	Annealing temperatures specific to each primer pair	Demonstrating quality difference between Fluorescent Primers and regular primers	Fluorescent primer product quality was less than that of regular primers.

*PC and PA= *P. clarkii*; OQ= *O. quinebaugensis*; OL= *O. limosus*
**Fl= Fluorescent primer

Because the 16S locus is known to exist in *P. clarkii* and other crustaceans, our research began by testing the 16S primers on *P. clarkii* and *O. quinebaugensis* samples to become familiar with using PCR for microsatellite analysis and verify effective PCR technique. The results of this PCR can be seen in Figure 5 below.

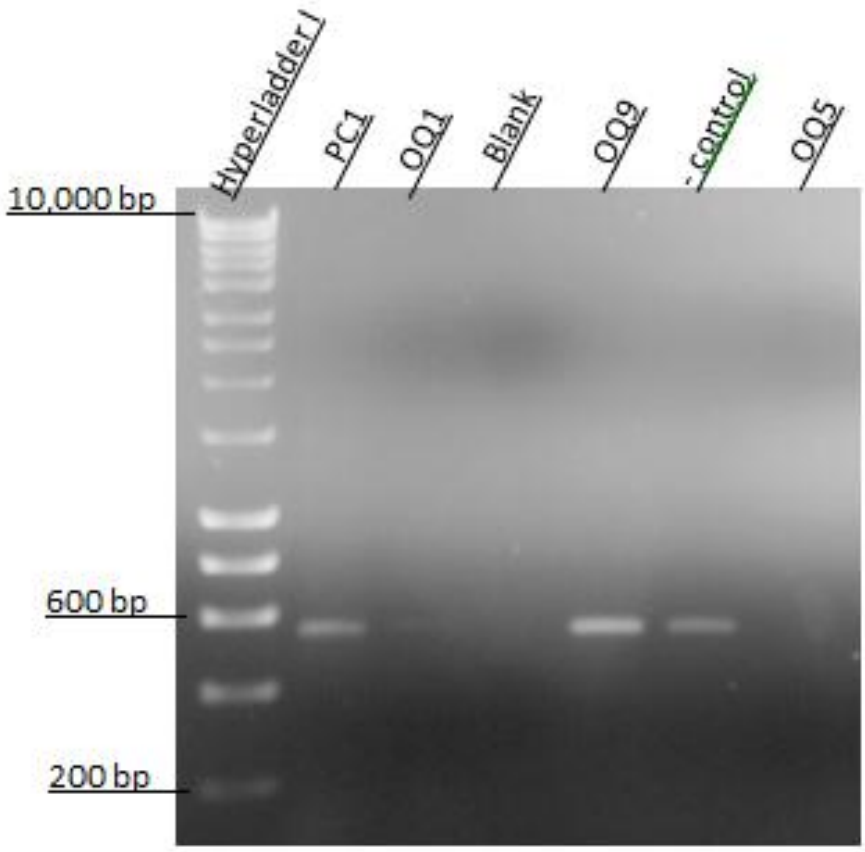

Figure 5: Amplification of *P. clarkii* and *O. quinebaugensis* DNA with 16S primer. Each lane is labeled with the name of the DNA sample used.

Figure 5 demonstrates that our technique was successful and effective. Three of the four DNA samples used produced visible results. The well labeled OQ9 has a particularly bright, clear signal, supporting the success of our technique. From this point forward, the 16S primer was used as a positive control with *O. quinebaugensis* DNA in all PCR reactions. In all gels throughout the rest of our research, the 16S product in the positive controls matched the expected size of 550bp.

The negative controls in all gels included the primer of interest with no DNA. Since there was nothing to amplify, the well was expected to show no product in the absence of contamination. In this particular gel, the negative control gave a strong band, indicating likely

contamination. In attempt to address this issue, a new stock sample of sterile water was created and used in future experimentation.

The next step in our project was to test each of the primers previously designed for *P. clarkii* (see Table 4) against our *P. clarkii* DNA samples to verify their effectiveness in the population we used. In the primer testing phase, PCR was done using each primer pair on three *P.clarkii* DNA samples (PC3, PC5, and PC6), with a separate negative control for each primer pair. This process was completed twice to compare results. In both tests, the same annealing temperature was used for all primers. Figure 6 throughFigure *8* (shown below) illustrate some of the results of this initial testing. The results were analyzed for strength and specificity as described in Section 3.3

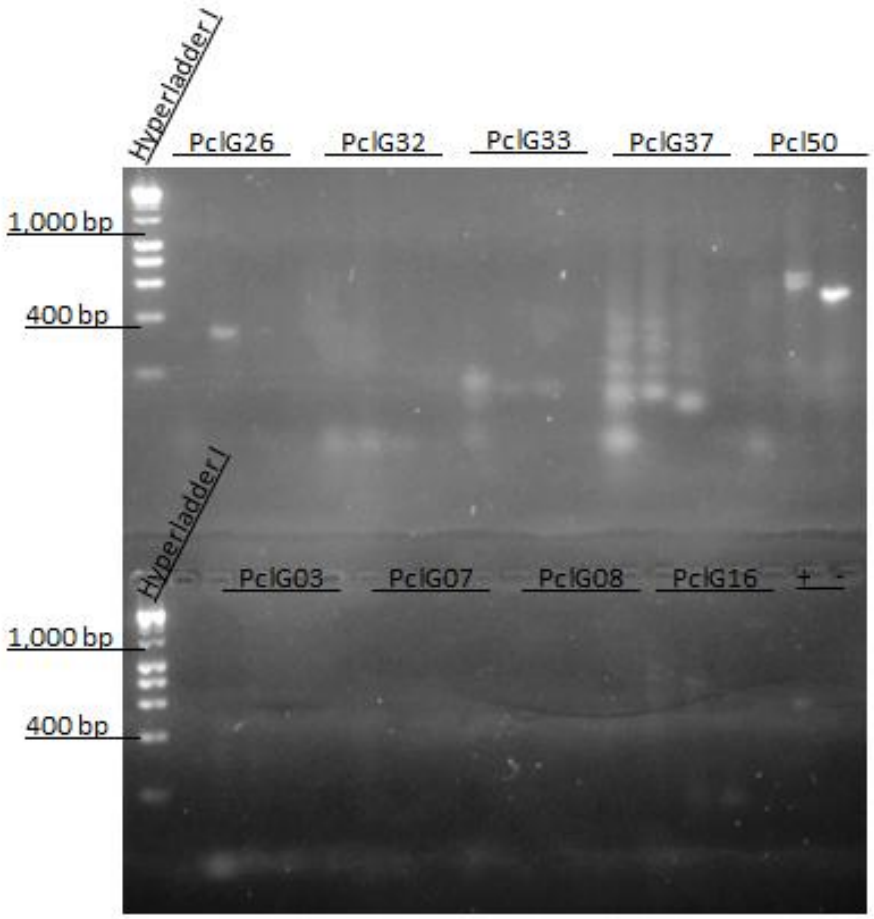

Figure 6: Initial primer testing on nine primer pairs. Each primer pair was tested on three *P. clarkii* samples and one negative control. Loci are labeled above the wells with the name of the forward primer used. The last lanes on the bottom row show a positive control using 16S and the OQ8 DNA sample, and an additional negative control using all primer pairs.

Figure 6 above shows the results of the initial test of nine primers. Most primers produced a faint band, and the size was less than 200 base pairs long (compared to Hyperladder I). Based on the small size and weakness of the bands, and the consistent presence across most samples, these bands were presumed to be primer dimers. Similar bands in subsequent gels were similarly identified as primer dimers and disregarded for purposes of analysis. As seen in the figure, PclG37 and PclG50 formed the strongest product. The presence of product is confirmed by the position of the bands on the gel. As shown in Table 4, Zhu and Yue (2008) predicted a fragment length of 424-524bp for the Pcl50 primer, which is the approximate size of the bands

on the gel. Likewise, Belfiore and May (2000) state that the fragment produced by PclG37 will be 80-180bp, which was also observed. Despite its strength, however, PclG37 showed considerable streaking and multiple products which indicated relatively weak specificity.

Products from PclG32 and PclG33 were also visible in expected size ranges, but faint in comparison. The remaining 5 primers produced little or no product on this gel. The positive control was created by using the 16S primer to amplify *O. quinebaugensis* DNA, since this combination previously gave strong results (Figure 5). This lane yielded product, which indicates no error with the PCR reaction. The negative control showed no product, illustrating lack of contamination.

Meanwhile, Figure 7 (below) shows the results of the initial test on the other ten primers. This gel shows bands in Pcl25, Pcl28, Pcl29, Pcl33 and Pcl47. It was noted that Pcl29, Pcl33, and Pcl47 showed a lack of specificity, with multiple products and streaking; further study was necessary on those primers to determine whether they were worth pursuing. The results of Pcl25 also warranted further study due to the unexpectedly large size of the bands. The remaining primer pairs gave inconclusive results, with only very faint product if any at all.

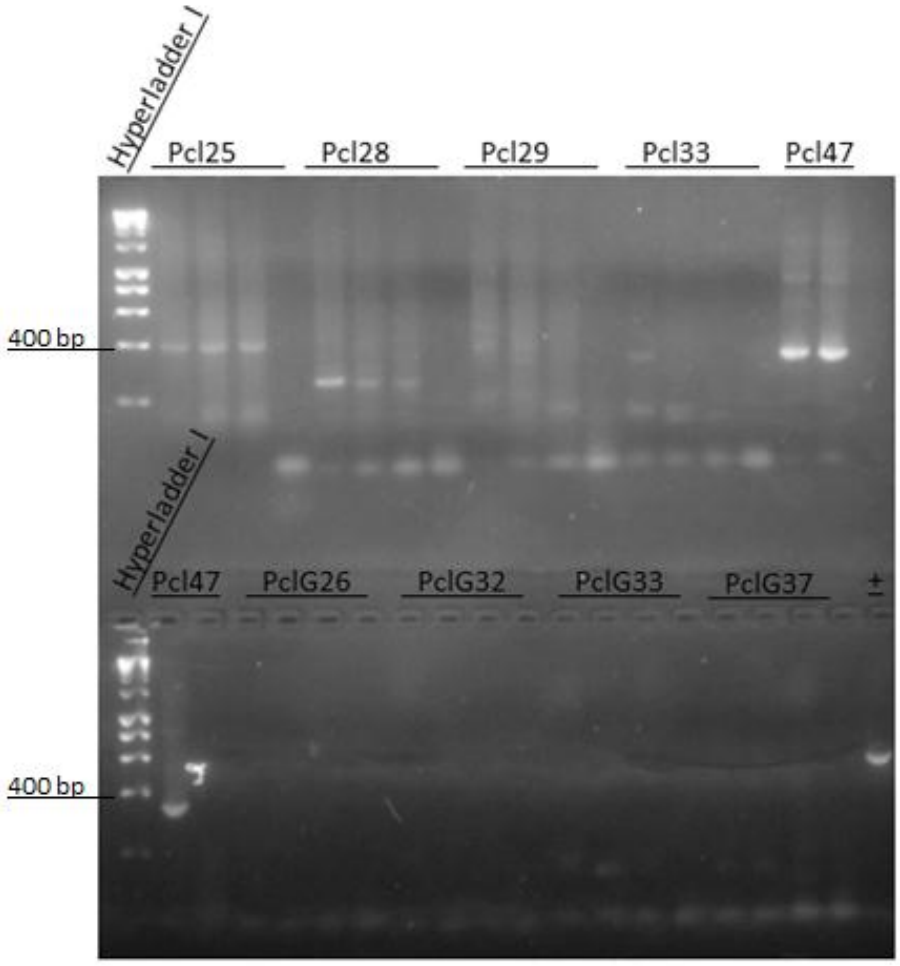

Figure 7: Initial primer testing on ten primer pairs. Each primer pair was tested on three *P. clarkii* samples and one negative control. Loci are labeled above the wells with the name of the forward primer used. The last lane on the bottom row show a positive control using 16S and the OQ8 DNA sample.

The second test of the 19 primer pairs showed many similar results, but also contained some noteworthy differences. Figure 8 shows results of the second test from nine primer pairs, five of which were initially shown in Figure 6, and the other four of which were initially shown in Figure 7. Pcl50 once again showed strong results, though with some lack of specificity. Certain bands are seen well below the expected size range of 424-524, indicating possible amplification of an artefact or possibly unspecific primer binding. Pcl02, and Pcl11 produced stronger products in this secondary test than they did initially, and the bands were in the expected

size range. While Pcl16 showed product, this product was much larger than expected, making the results questionable. The remaining primers on this gel continued to give weak or no product.

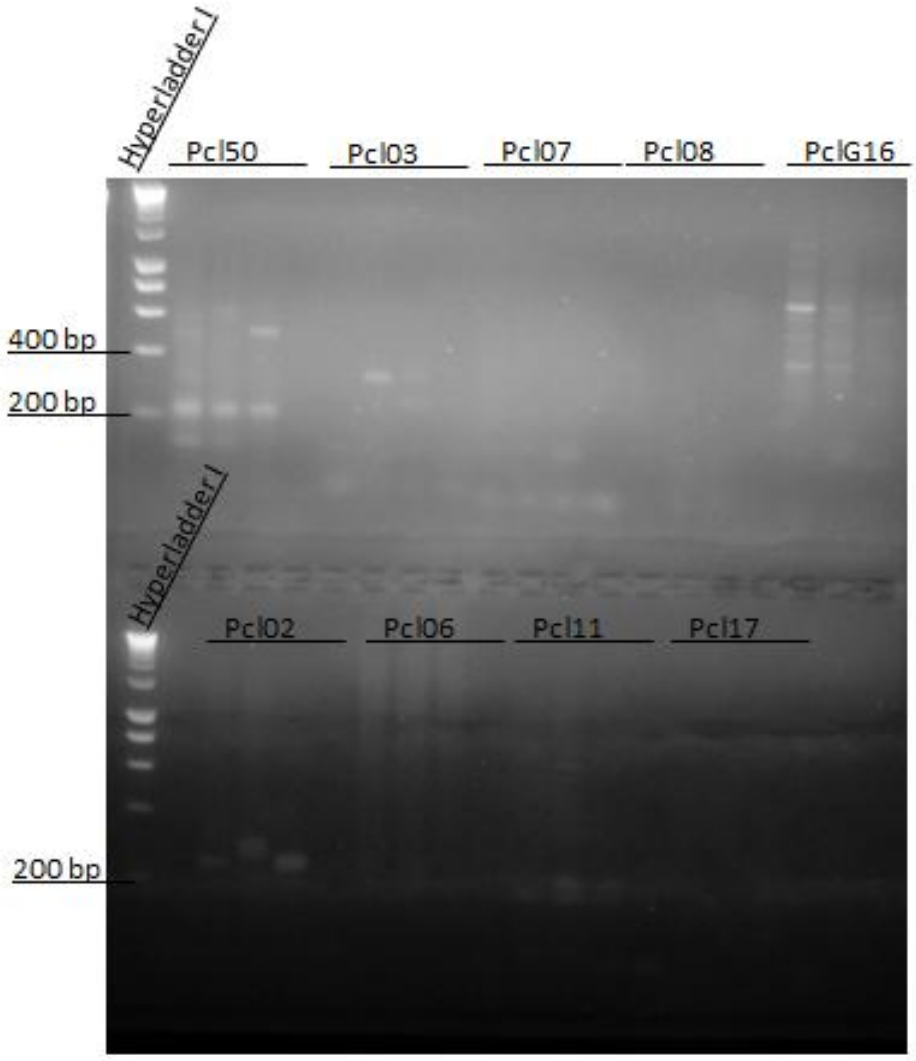

Figure 8: Subsequent primer testing on nine pairs. Each primer pair was tested on three *P. clarkii* samples and one negative control. Loci are labeled above the wells with the name of the forward primer used.

The remaining 10 primers were tested and visualized on another gel (not shown), and showed a similar mix of variation from the initial test. Because there was such variation between these two tests, further tests were performed using the same primers and DNA samples, employing optimization techniques in an attempt to clarify ambiguous results. Specifically, a different annealing

temperature was used for each primer, according to the calculation method described previously in Table 4. The results of these tests are shown below in Figure 9 through Figure *11*.

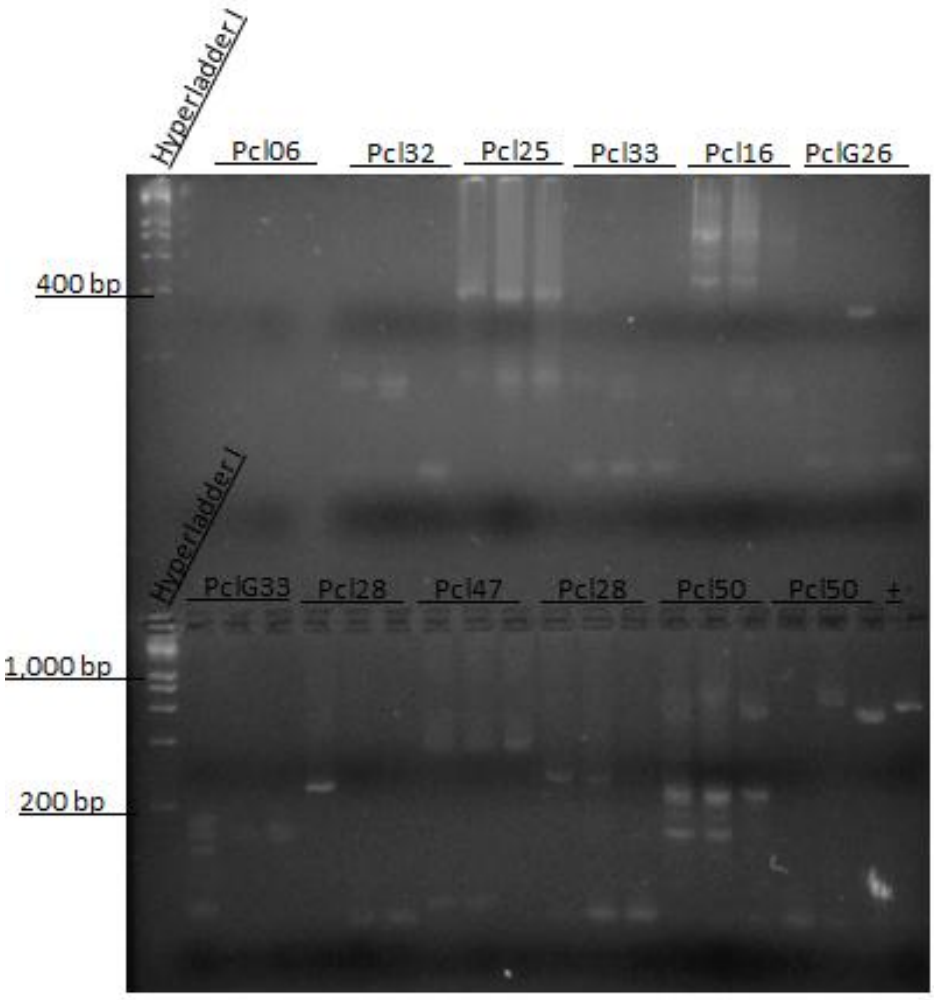

Figure 9: Optimized primer testing on 11 primer pairs using calculated specific annealing temperatures. Each primer pair was tested on three *P. clarkii* samples. Loci are labeled above the wells with the name of the forward primer used. The last lane on the bottom row show a positive control using 16S and the OQ8 DNA sample.

The use of separate, optimal annealing temperatures for each primer pair appeared to be effective. In Figure 9 above, all 10 primer pairs tested showed some product, and some of those that produced results in the initial trials showed stronger results here. In particular, PclG26 showed weak but noticeable results in this test, where it showed no results in the initial tests. Because Pcl50 showed strong results in both initial trials, it was run twice in this test to confirm

that it would be successful across a wider range of trials; both produced strong results, though the one on the right indicates higher specificity than the one on the left. Overall, Pcl16, Pcl28, and Pcl47 showed results using optimized annealing temperatures, while Pcl06, Pcl32, Pcl33, and PclG26 produced weak but still notable results.

Interestingly, Pcl25 once again produced bands much larger than the expected size according to Zhu and Yue (2008). Since this large product is the same size as in previous experiments (Figure 7), it is speculated that there may be contamination. On the other hand, since the product is clean and distinct, it is also a possibility that the primer is not amplifying the region originally proposed by the authors.

Figure 10 below depicts a subsequent primer testing using optimal annealing temperatures. PclG26, PclG33, and PclG37 produced visible results. Pcl16 gave product, but the results were once again indistinct, implying contamination in the primer (compared to results in Figure 8 and Figure *9*). Since this primer had given the same results in many separate trials, this primer was not considered for further study.

PclG26 looked much better than it did in previous trials (Figure 8); conversely, Pcl50 produced significantly weaker results than every previous trial. Since this primer had always produced strong products, it was rerun again in a later trial to ensure that it was not simply an experimental error. Additionally, it is worth mentioning that some of the negative control wells show visible results, indicating possible contamination. This is another reason that further studies were performed on this primer set.

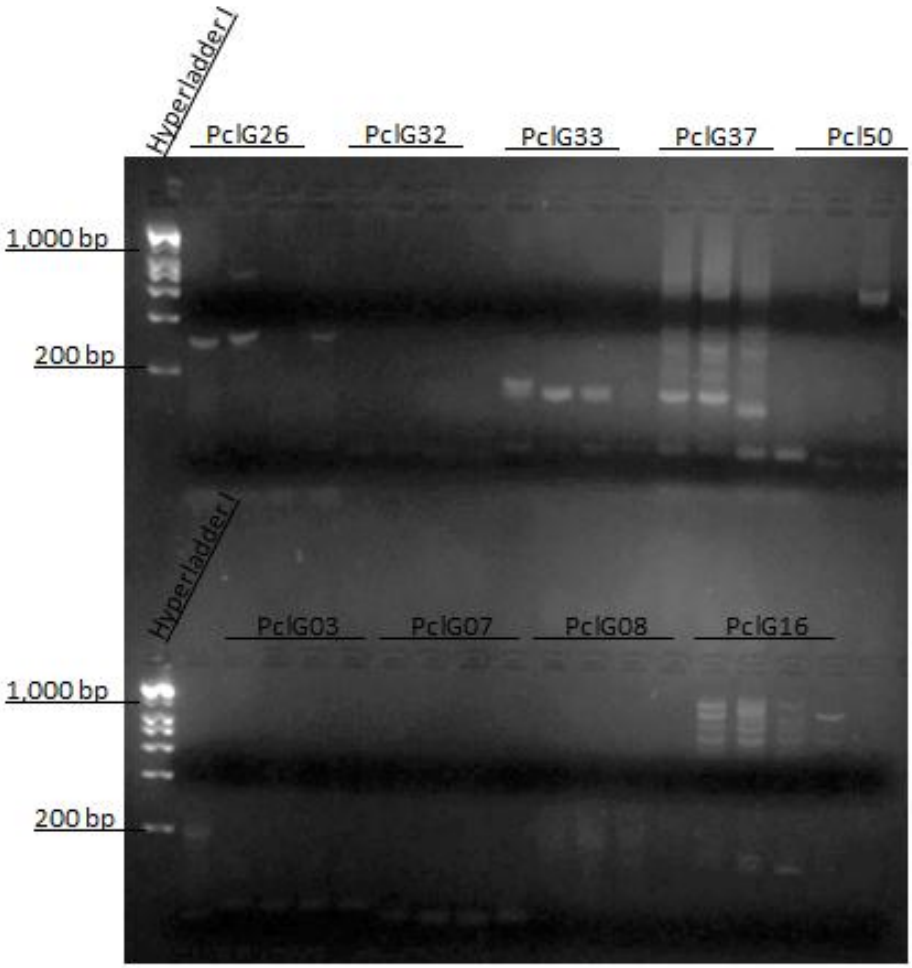

Figure 10: Optimized primer testing on 9 primer pairs using calculated specific annealing temperatures. Each primer pair was tested on three *P. clarkii* samples and one negative control. Loci are labeled above the wells with the name of the forward primer used.

Again, these tests were done twice for comparison of results. Figure 11 below shows further results from these annealing temperature-optimized tests. These results indicate that Pcl25, Pcl28, Pcl29 and Pcl47 produced promising product once more.

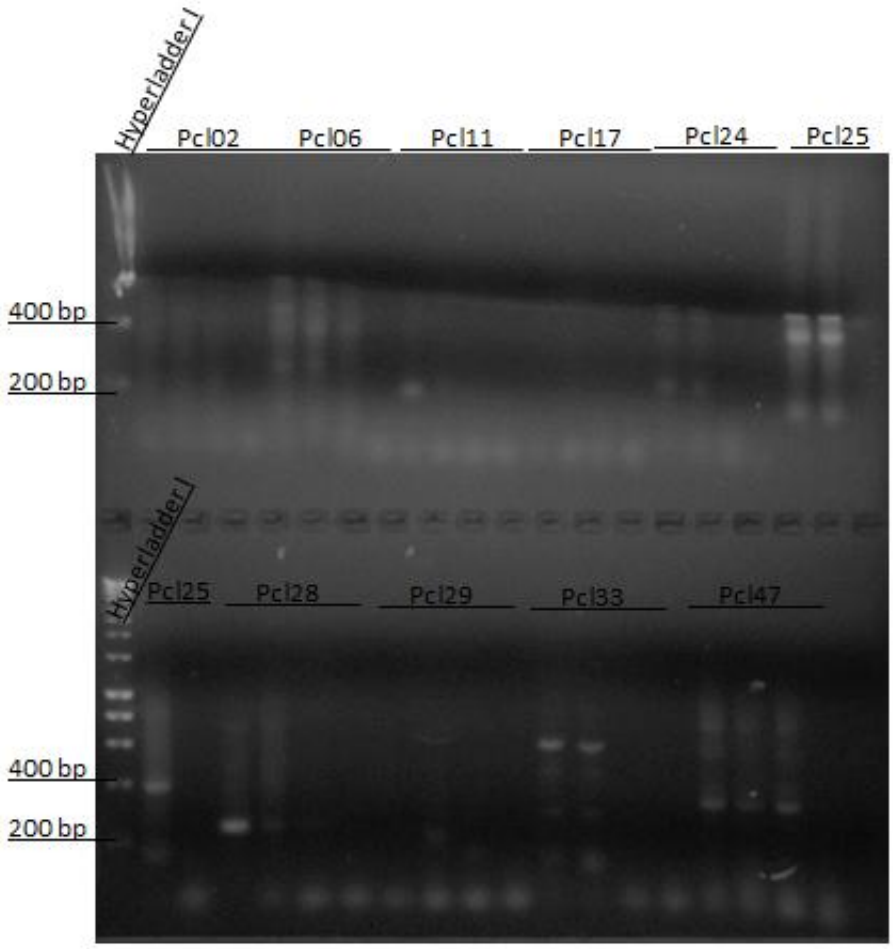

Figure 11: Optimized primer testing on 10 primer pairs using calculated specific annealing temperatures. Each primer pair was tested on three *P. clarkii* samples and one negative control. Loci are labeled above the wells with the name of the forward primer used.

The results of all these tests, as well as the initial tests, were compared to identify trends in consistency of the strength and specificity of results. A total of eight primer pairs which repeatedly showed strong results were identified as Potential Primers of Promise. These included Pcl25, Pcl28, Pcl29, Pcl32, Pcl47, PclG08, PclG32, and PclG37. Two primer pairs (PclG26 and PclG33) were deemed sufficiently consistent to use in further studies due to their clean results of expected size in (Figure 10). These primers were identified as Primers of Promise and would undergo further tests to try to produce strong, specific products with greater consistency than the previous tests had shown. The remaining eight primers were considered Eliminated; further

exploration beyond the scope of our study would be needed to determine why they were not successful under the conditions we used.

The next tests attempted to further optimize PCR for the eight Potential Primers of Promise, this time by running a set of eight identical samples at different annealing temperatures ranging from 2˚C below the previously calculated optimal annealing temperature to 5˚C above this temperature. The goal was to determine whether a given temperature would enable stronger or more specific results. This was done twice for each primer pair, using the two DNA samples that most consistently showed results (PC5 and PC6). Figure 12 throughFigure *15* show the results of these tests.

The first of these gradient tests, shown in Figure 12 below, used two of the weaker primers from the previous gel (Pcl29 and PclG32). The annealing temperature gradient did not seem to have any effect on the results; all visible bands were of the same approximate strength. Furthermore, only very faint results could be seen in the expected range; the strong bands were small enough that they most likely indicated primer dimers rather than product. Because of the lack of apparent results, these primers were considered Eliminated, and underwent no further testing.

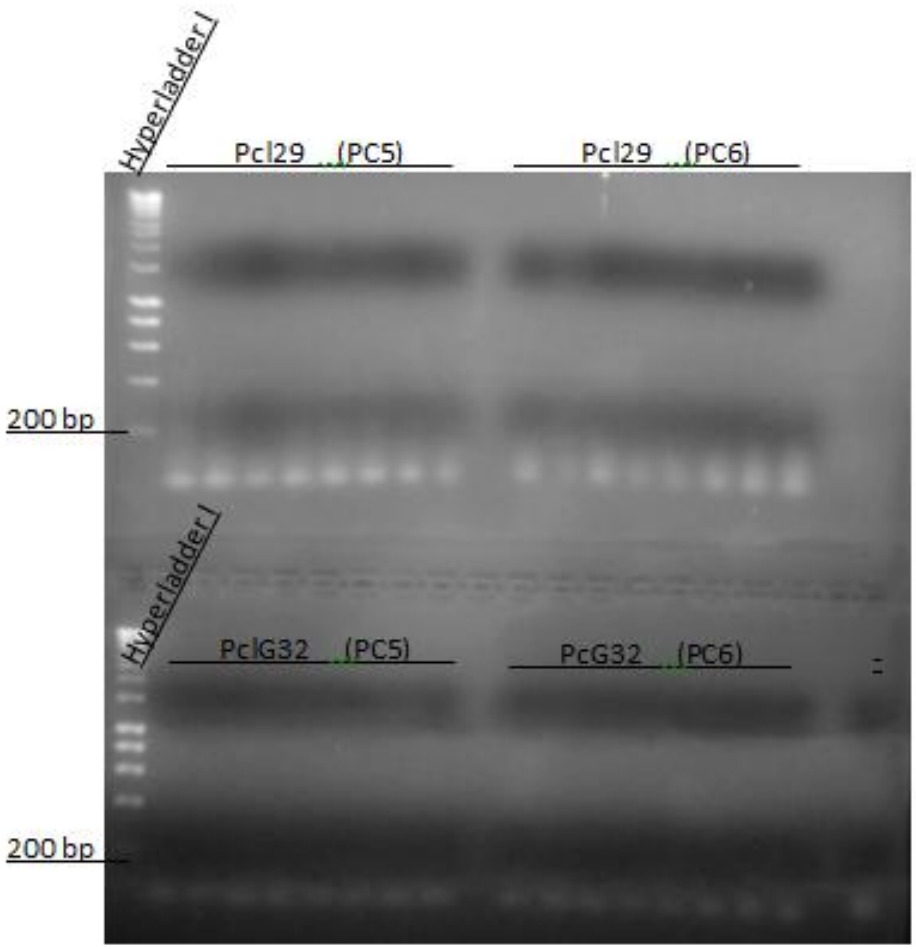

Figure 12: Testing of two Primers of Promise with annealing temperature gradient. Both primer pairs were tested on two *P. clarkii* samples, with eight identical PCR samples each. Samples were run on a gradient to determine the best annealing temperature to produce clean results. The last lane in the bottom row shows the negative control, containing both primers.

The next primers tested on a gradient were PclG37 and PclG08. As shown in Figure 13 below, once again the gradient didn't appear to have a noticeable effect. In this case, PclG37 showed a distinct product of expected size in addition to the presumed primer-dimer bands. Although the gradient did not improve the results, the presence of bands indicated that PclG37 was worth further pursuit. At this point, Pcl08 was considered Eliminated.

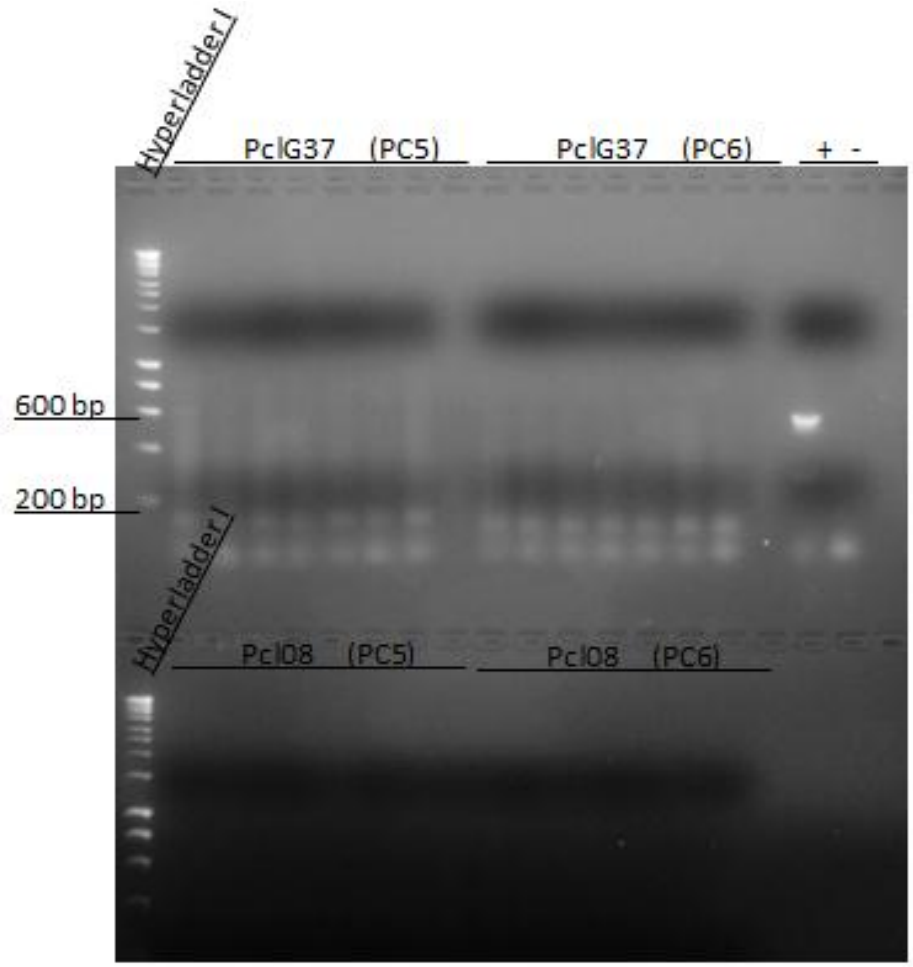

Figure 13: Testing of two Primers of Promise with annealing temperature gradient. Both primer pairs were tested on two *P. clarkii* samples, with eight identical PCR samples each. Samples were run on a gradient to determine the best annealing temperature to produce clean results. The last lanes in the top row show the positive control using the 16S primer and OQ8 DNA, and the negative control containing both tested primers.

Pcl47 and Pcl25 were the next primers tested on a gradient of annealing temperatures. Again, the gradient appeared to have no effect. As shown in Figure 14 below, both produced notable product. However, in both cases, there was considerable streaking, which indicated a likely lack of specificity. Due to the low specificity, both were considered Eliminated.

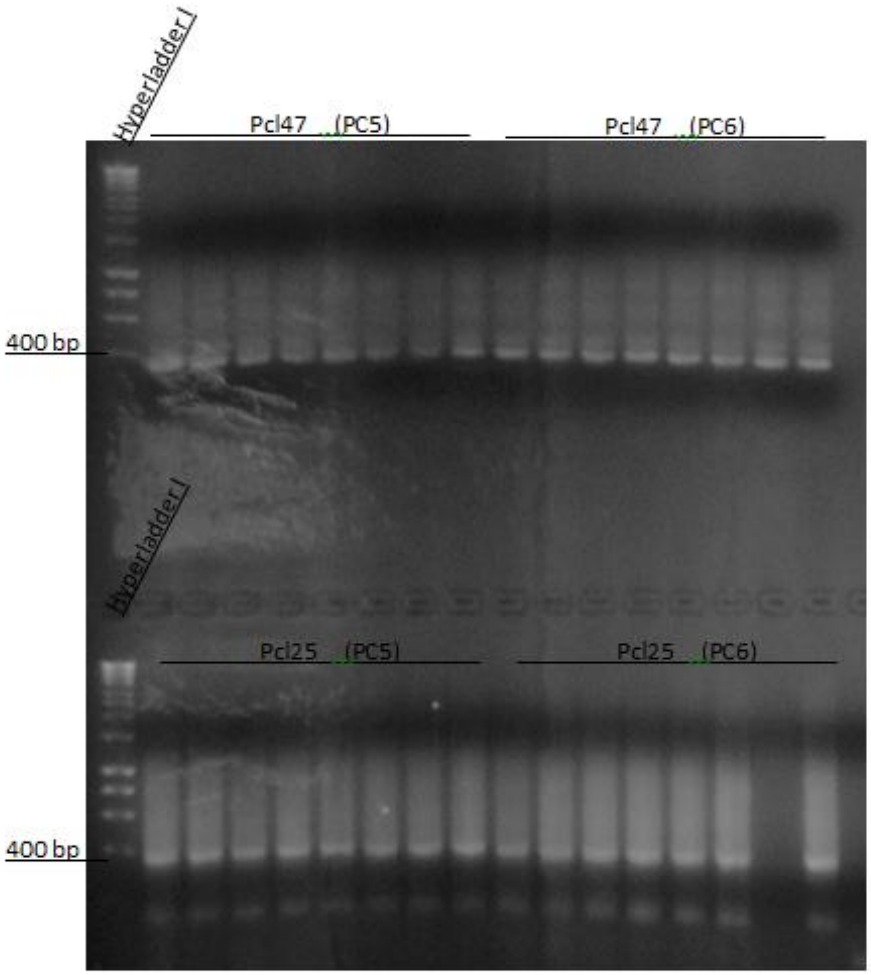

Figure 14: Testing of two Primers of Promise with annealing temperature gradient. Both primer pairs were tested on two *P. clarkii* samples, with eight identical PCR samples each. Samples were run on a gradient to determine the best annealing temperature to produce clean results.

The final PCR to test Potential Primers of Promise on an annealing temperature gradient was run with Pcl28 and Pcl50. These results are shown in Figure 15 below. Pcl28 showed strong bands large enough to indicate product rather than primer dimers, although the lack of clarity in the Hyperladder makes it difficult to distinguish exact size. Pcl50 also showed consistent products in the samples using PC5 DNA, with some products using PC6 DNA as well. Both were considered worth pursuing further.

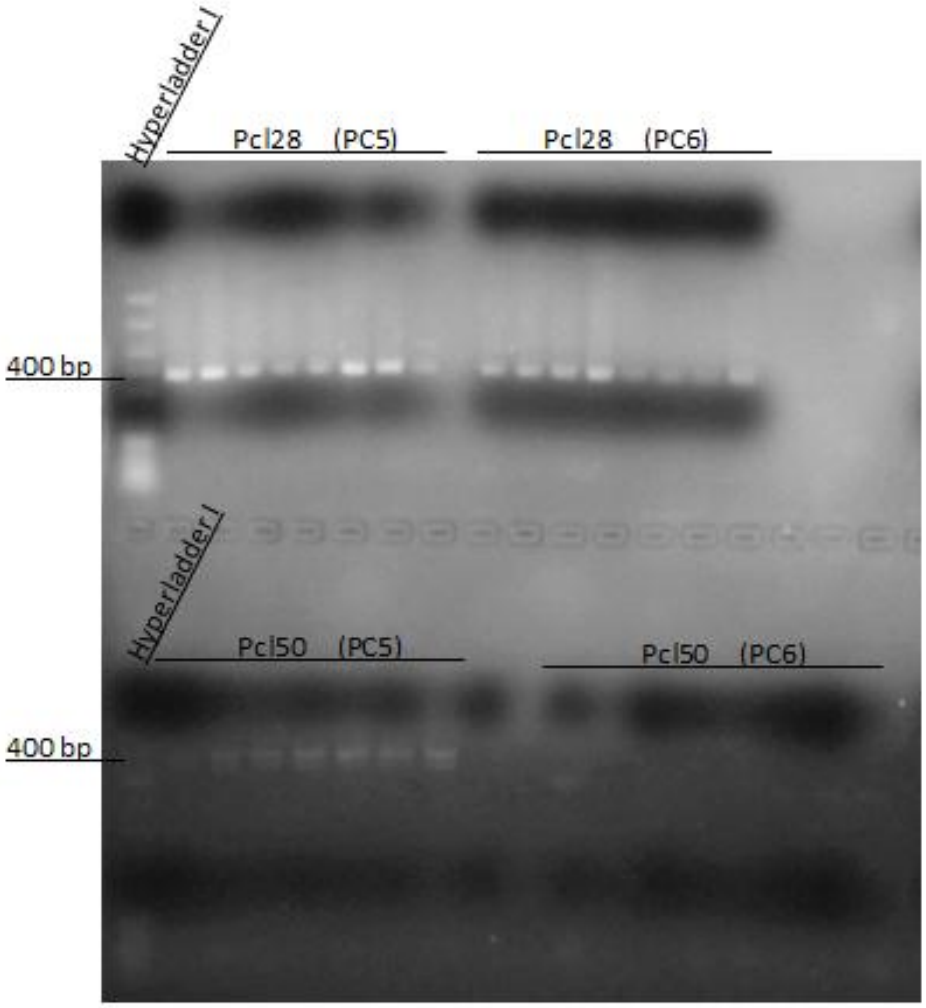

Figure 15: Testing of two Primers of Promise with annealing temperature gradient. Both primer pair were tested on two *P. clarkii* samples, with eight identical PCR samples each. Samples were run on a gradient to determine the best annealing temperature to produce clean results.

After narrowing down the primers of promise to five pairs (PclG26, PclG33, PclG37, Pcl28, and Pcl50), a PCR was run using the enzyme PicoMaxx to determine whether this would improve the results. This was done in order to finalize the list of primers to be used in future experimentation. Figure 16 below shows the results of this final test.

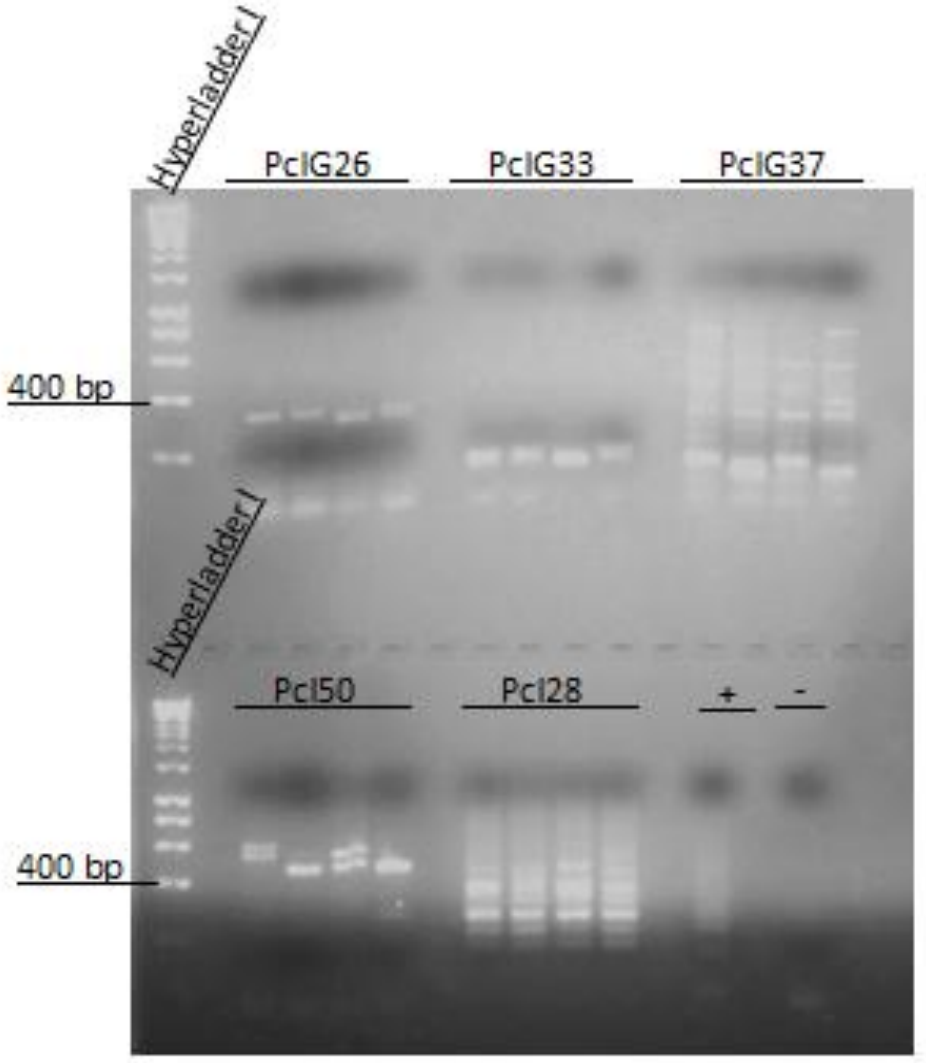

Figure 16: Final Primers of Promise test. Each primer pair was tested on four *P. clarkii*. Loci are labeled above the wells with the name of the forward primer used. The last lanes on the bottom row show a positive control using all primers and OQ8 DNA, and a negative control using all primers.

All five primer pairs produced considerable amounts of product. PclG37 and Pcl28 both showed a large amount of streaking, as well as multiple distinct bands which might indicate multiple loci. Due to that lack of specificity, those primers were Eliminated. However, there were one or two strong, well-defined bands each for PclG26, Pcl33 and Pcl50. Because of the high strength and specificity shown by these primers, as well as their appropriate product size, they became our official Primers of Promise for further experimentation.

After selecting the three Primers of Promise, they were tested on all *P. clarkii* samples using the calculated optimal annealing temperature to ensure that they successfully amplified microsatellites across the majority of individuals. Figure 17 shows the result of this PCR.

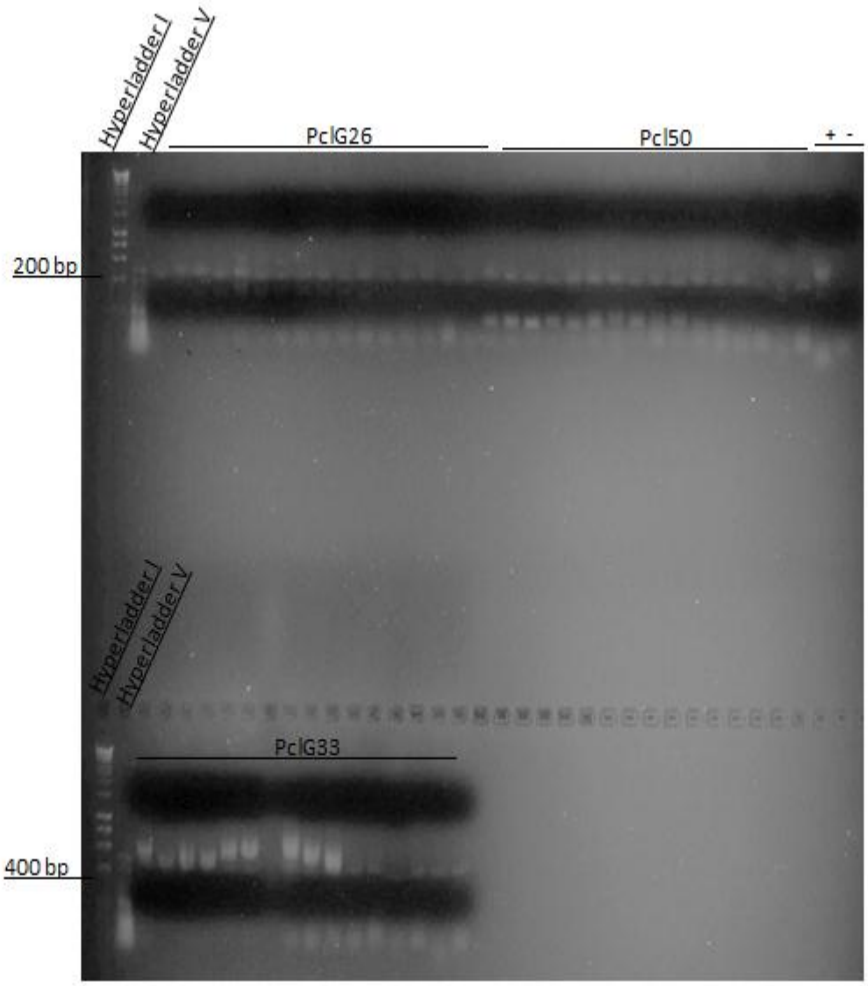

Figure 17: Primers of Promise testing across all *P. clarkii* samples. Each primer pair was tested on sixteen *P. clarkii* samples. Loci are labeled above the wells with the name of the forward primer used. The last lanes on the top row show a positive control using 16S and the OQ8 DNA sample, and a negative control using all primers.

All primers showed product, indicating their ability to amplify microsatellites across several *P. clarkii* DNA samples. However, it was noted that Pcl50 showed two distinct products in addition to the presumed primer dimers. This may indicate a lesser specificity than was

thought based on previous tests. In spite of this, there was consistency within the Pcl50 results, and it was still considered worth using.

Because the results were promising in *P. clarkii,* the primers were used in an attempt to amplify DNA from two additional species: *O. limosus* and *O. quinebaugensis.* The results of this PCR are seen in Figure 18 below.

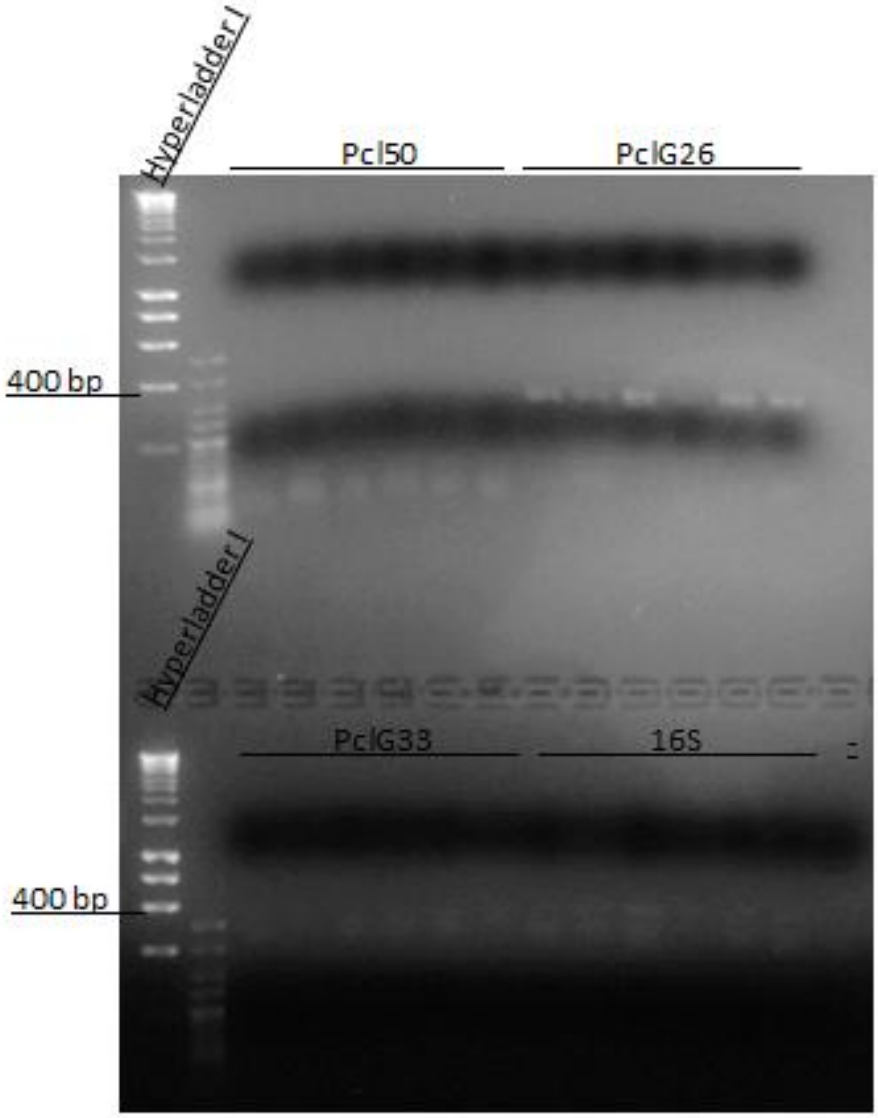

Figure 18: Primers of Promise testing on *O. quinebaugensis* and *O. limosus* DNA. Each primer pair was tested on three *O. quinebaugensis* samples (first three lanes in each set) and three *O. limosus* samples (last three lanes of each set). Loci are labeled above the wells with the name of the forward primer used. Each sample was run using 16S as a positive control. The last lane on the bottom row shows a negative control using all primers.

The first three samples under each primer pair were *O. quinebaugesis,* and the last three samples were *O. limosus*. The products of these reactions were not as strong or clearly defined as they were in *P. clarkii*. Due to time constraints, the primers were not pursued further in these species.

In order to prepare samples fragment analysis, fluorescent primers were ordered. The fluorescent primers were run with *P. clarkii* DNA samples. In early trials following the protocol used with normal primers, none of the fluorescent primers produced clear results (gels not shown). In order to increase the PCR product, the quantity of fluorescent primers was doubled. Figure 19 below shows the results of PCR using doubled fluorescent primers.

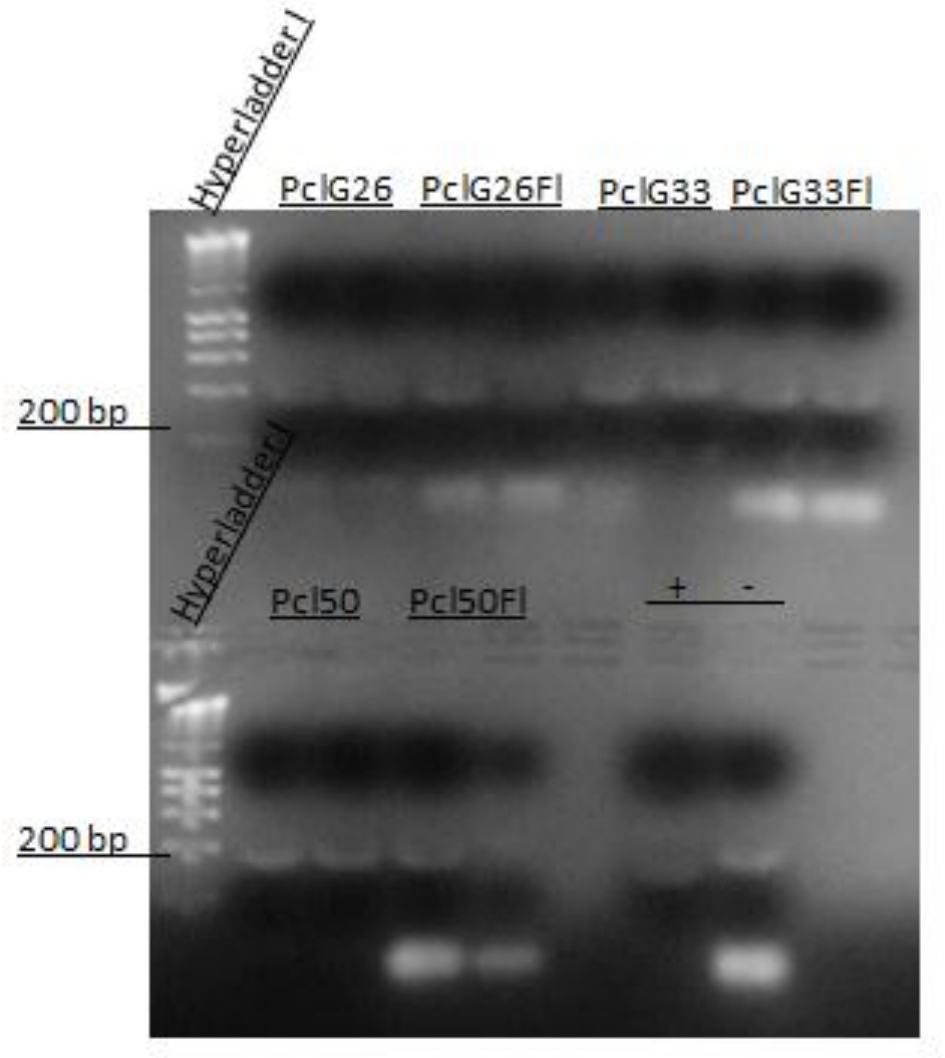

Figure 19: Fluorescent primer testing against non-fluorescent primers for quality control. Each primer pair was tested on two *P. clarkii* samples. Loci are labeled above the wells with the name of the forward primer used. Fluorescent primers are labeled with Fl. The last lanes on the bottom row show a positive control using 16S and the OQ8 DNA sample, and a negative control using all primers.

Very weak products were yielded, in spite of the larger amount of primers used. The bands for both fluorescent and normal primers were approximately 400bp. However, the regular primers produced bands of about 80ng/uL, while the fluorescent primers produced bands with a much lower concentration. A weak positive control product (using 16S primer and OQ8 DNA) can be seen in the bottom row. There also appeared to be contamination in the negative control at just

above 400bp, causing some concern regarding the validity of other results on this gel. However, as the primary purpose of this test was to confirm the fluorescent primers were able to successfully amplify DNA, it was determined to have accomplished that goal in spite of the contamination.

The products of this gel were sent to Cornell University Institute of Biotechnology for fragment analysis. The file was analyzed, and a screenshot of the analysis of PclG26 is shown in Figure 20 below.

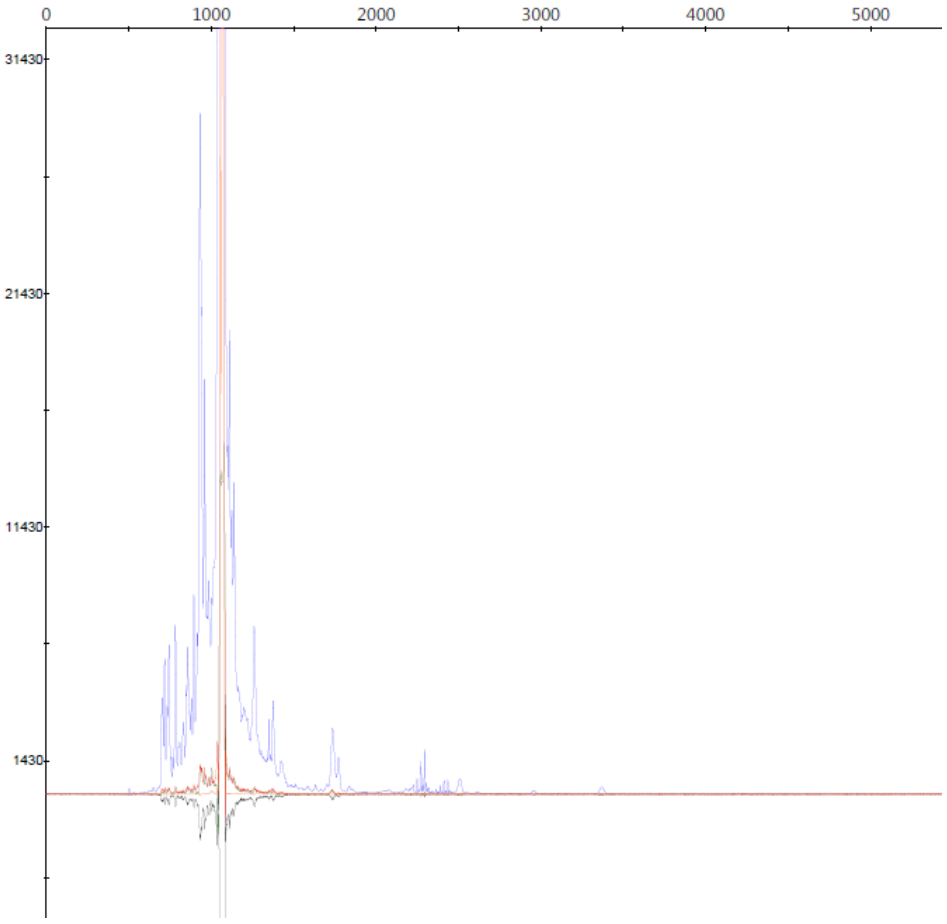

Figure 20: Fragment analysis of PclG-26 from P. Clarkii samples. Fragment analysis was received from Cornell University and viewed using Peak Scanner software v1.0 from Applied Biosystems. The results shown above show colors from blue, green, yellow, red, and orange dyes indicating a contamination issue as the Liz600 dye is colored orange while PclG-26 with a 5' Tet modification emission range was 535 nm.

Overall, it is difficult to distinguish what this data means due to difficult with the software and possible contamination.

5. Discussion

The eventual goal of our research was to apply successful *P. clarkii* primers to amplify microsatellite loci in *O. limosus* and *O. quinebaugensis*. Using primers that have already been developed for *P. clarkii* would save the time and cost of designing and manufacturing new primers for the other species. Furthermore, the presence of matching primer sites and microsatellites in different species might provide information regarding the evolutionary relationship between the species. For example, highly conserved microsatellite regions may indicate a relatively recent common ancestor.

In order to achieve that goal, we first needed to confirm that the primers worked with *P. clarkii*. This proved to be an extremely problematic step; only three of the nineteen primers produced consistent, strong, specific results in *P. clarkii*, in spite of the fact that they all reportedly worked in previous studies. This demonstrates the sensitive nature of microsatellite loci, which can produce widely varied results depending on a variety of factors that will be discussed later in this section. Using those three Primers of Promise, we were able to move forward in testing these *P. clarkii* primers in *O. limosus* and *O. quinebaugensis*. Our results showed limited preliminary success with the three successful primers used with these two species. To determine whether those three primer pairs are viable for consistent use with these two species, additional research would be required.

There were many areas of this project that introduced challenges and delayed obtaining successful results. The first involved obtaining fresh DNA samples from recently collected crayfish. Genomic DNA extraction is a difficult process. There are many different methods that produce a

wide range of results regarding the amount and quality of the DNA obtained. Specifically, where DNA is obtained from an organism may affect the overall quality of the DNA (Li *et al.*, 2011). However in this study, DNA that was successfully extracted appeared clearly on the gel with little indication of degradation. While the DNA extraction gels are not shown, samples were tested with a 16S primer that target mitochondrial microsatellite loci which produced strong and distinct bands as shown in Figure 5.

There was also considerable variation among our individual samples due to experimental and environmental factors both during and after the extraction process. From the time the tissue is extracted, it must be kept on ice to prevent degradation of the sample. This makes extraction a time-sensitive process and storage of samples an important consideration. During this project, several samples were lost when the freezer in which we stored our DNA samples and crayfish tissue malfunctioned. The extended exposure to temperatures above freezing may have degraded the samples, rendering some unusable or less than optimal; the tissue, which would have been used for future extractions, began to rot and had to be discarded entirely. Additional challenges arose during the extraction process itself; the tissues were physically difficult to extract, and some samples may have been cleaner or contained more tissue than others. The centrifugation and washing processes provide more room for experimental error and human contamination. For example, when discarding the supernatant, it is easy to accidentally disturb the pellet and decrease the quality of the sample. All of these factors may have caused variation among the samples that might influence the results of PCR.

Because there were only a few samples that consistently produced strong and clear results, we diluted those samples after a few successful iterations of PCR to ensure there would be enough DNA to complete the project. The effort to ration the DNA extracted was primarily due to two separate incidents in lab that resulted in the complete loss of frozen crayfish samples and degradation of DNA. That change in dilution may have affected our PCR results, due to the smaller amount of

DNA in each sample; however, there was no noticeable change in PCR results after the dilution, so we believe that any amplification problems were due to other factors. Currently there is no clear answer to why the primers failed to isolate the microsatellite loci previously reported in reported by Belfiore and May (2000), Zhu and Yue (2008) and Hulak *et al.*, (2010). From the numerous attempts to isolate the specific loci and positive reinforcement of both the positive and negative controls found throughout most of the figures shown, any answer to why this happened is purely speculation. However, there are some solutions that will increase the probability of obtaining cleaner results in future experimentation. From personal communications with Professor Destin Heilman, Professor Lauren Mathews, and Professor Michael Buckholt, genomic DNA extraction is one of the harder techniques to master. DNA was extracted from several difference sources on the crayfish, which may have led to inconsistencies with the quantity of DNA extracted. Therefore, in effort to improve DNA quality, DNA extraction location and amount should be standardized. There should be some time spent investigating other techniques and reagents for DNA extraction to see if there is a way to improve the quality of DNA as well. Specifically, the protocol described by Glenn and Schable (2007), should be attempted as they were able to improve microsatellite amplification by running DNA fragments that contain microsatellite through a Dynabead column. This method alone would help remove all other DNA fragments and ensure the purity of the amplification product during PCR.

The 19 primer pairs studied in our research were identified based on two studies on *P. clarkii* microsatellites: one in China in 2008 (Zhu & Yue, 2008), and one in southern USA in 2000 (Belfiore & May, 2000). Although those studies indicated that all 19 primer pairs should produce microsatellite products in our *P. clarkii* samples, we needed to verify their results, as well as understand the range of efficacy among the primer pairs. Specifically, in using these primer pairs to amplify microsatellite loci in *O. quinebaugensis* and *O. limosus*, we would expect that the primers

that produced the strongest product in *P. clarkii* may be more likely to work in the two other species, providing a starting point for cross-testing. To that end, several iterations of PCR were run in order to identify the primers which consistently produced the best results.

Of the nineteen primers we studied, only three showed consistent, strong, and specific results over the various iterations. There are a few potential explanations for this deviation from the previous studies. First, they likely used different PCR conditions. Adjusting the annealing temperature, number of cycles, enzyme, or concentrations of reagents might produce different results. Secondly, the previous studies were performed on populations of crayfish that were separated by both time and distance from the central Massachusetts crayfish used in our research. We do not know the history of the invasive population in Institute Pond, and it is possible that, due to that time and distance, the primer sites in our crayfish had mutated to the point where the primers could no longer adequately anneal to the DNA during PCR, rendering those primers ineffective. This supports the logic that the strongest *P. clarkii* primers would be more likely to succeed in *O. limosus* and *O. quinebaugensis.* The primer binding sites that successfully worked on our crayfish would be more conserved throughout various *P. clarkii* populations, and therefore potentially in other species as well.

Not all variation in our results, however, can be attributed to population differences. Throughout the course of the experiment there were cases of substantial inconsistency in the visualized PCR results from one iteration to the next. This was often attributed to user error, and the iteration repeated to achieve reliable results. Even when the gels displayed consistent results, it was often difficult to distinguish product bands from primer-dimer bands or, in some cases, potential contamination. These experimental errors are demonstrated in Figure 10 and Figure 19, as well as several other gels not pictured. In an effort to improve the quality of the product, steps were taken to optimize the PCR process. These steps included adjusting the annealing temperatures of the primers and using a different enzyme.

In optimizing the annealing temperature, we used a method called Touchdown (TD) PCR (Hecker *et al.*, 1996). In TD PCR, only a single variable is changed in the PCR procedure. In our case, PCR samples were placed across a gradient of temperatures during the annealing phase. The goal was to determine at which temperature the strongest and most distinct product was formed. Higher temperatures would theoretically diminish nonspecific primer binding, forming product with higher specificity. Conversely, because lower temperatures allow for less specific binding, a larger volume of product may be produced. If the annealing temperature is too high, the primer will be unable to bind to the DNA at all, as it will immediately denature, resulting in the formation of no product.

Figure 12 throughFigure *15* show the attempts to optimize the primers. As shown, there are no noticeable differences as the temperature ranges above and below the original annealing temperature (Hecker *et al.*, 1996). This indicates that the primers were already highly specific to their binding sites and therefore any increase or decrease in annealing temperature would, if anything, only weaken the product. Due to the size and number of the samples run, it is difficult to discern qualitatively if there is a difference between each individual product. Further analysis would need to be done to determine the amount of product and whether there were subtle variations at different temperatures. For the sake of our research, any such variation would be negligible, and we continued using the original annealing temperatures from that point forth.

Next the decision was made to use a different polymerase enzyme which might produce stronger and more specific results. We switched from *Taq* Polymerase to the higher-quality PicoMaxx High Fidelity PCR System (2.5 U/uL). According to Agilent's website, PicoMaxx is a blend of *Taq* and *PfU* DNA polymerases as well as ArchaeMaxx polymerase-enhancing factor, which increases the success rate of the PCR enzyme. Figure 16 displays the first use of the

PicoMaxx system. It shows brighter, stronger, and more distinct bands than any previous PCR gels. After several more trials it was clear that PicoMaxx consistently resulted in cleaner bands.

Lastly, in effort to achieve a higher resolution band separation, a finer agarose was used to separate the bands. Specifically, Amresco 3:1 High Resolution Blend agarose was employed. According to the manufacturer, agarose is specially made for small PCR products that are less than 1,000 bp in length. As all our results indicate our PCR products to be less than 300 bp, this was used to help distinguish single (homozygous) bands from sets of heterozygous bands that were very close in size. Overall, it was qualitatively determined that while the gel did provide finer resolution, it was difficult to work with and did not provide sufficiently improved information to continue using consistently (Amresco, 2014).

Fragment analysis came back inconclusive from Cornell University and the size of number of the fragments were unable to be discovered (Figure 20). From the PCR products produced, illustrated in (Figure 19), there is some doubt of whether or not the primers isolated the target microsatellite loci. When we compared the sizes of our products with those reported by Belfiore and May (2000), Zhu and Yue (2008) and Hulak *et al.*, (2010) we found that the analysis revealed different allele sizes than predicted. Specifically, on average, the products from the analysis were 300 base pairs in size; the literature states that the fragment sizes for only one of our primers (PclG26) fits this criteria.

There are several steps which could be taken to build upon this project in the future. First, sequence analysis could be performed on our successful PCR results. The sequence of the microsatellites amplified in our project could be compared to the original microsatellite sequences amplified by Belfiore and May (2000) and Zhu and Yue (2008). Some discrepancy in the sequence would be expected, as it would indicate that the microsatellite has mutated normally. If the sequence shows a different pattern than the expected microsatellite, however, the

discrepancy may indicate that the primer no longer amplifies the region it was designed to amplify. The original primers were developed several years ago and in different areas of the world. Just as in some cases the primer sites may have mutated such that the primers are no longer complimentary, other sites may have mutated to become compatible with the primers, yielding new sequences and confounding the results.

Another future step that could be taken is to further optimize the primers for better results in *P. clarkii*, as well as for use in *Orconectes limosus* and *O. quinebaugensis*. Designing primers is costly and time-consuming, so it would be beneficial for future researchers to adapt the same primer pairs to work across multiple species. This could involve changing annealing temperatures, changing magnesium concentrations, improving reagent quality, or a number of other technical adjustments. An entirely new primer may need to be designed if these adjustments are not successful. When designing these primers, one would need to ensure that it only amplified one piece of the genome. One way to accomplish this would be to use a longer primer sequence to reduce the chances that it would match more than one piece of the genome.

This project lays the groundwork for future research in the field of population genetics. Once the primers are optimized to produce consistent results, topics like inheritance, gene flow, genetic diversity, reproduction habits, and more will right at researchers' fingertips.

6. References

Abdelkrim, J., Roberston, B.C., Stanton, J. L., and Gemmell, N.J. (2009). Fast, cost-effective development of species-specific microsatellites genomic sequencing. *BioTechniques. 46*(3), 185-192.

Andrew, R., Bernatchez, L., Bonin, A., Buerkle, C.A., Carstens, B., Emerson, B., Garant, D., Giraud, T., Kane, N., Rogers, S., Slate, J., Smith, H., Sork, V., Stone, G., Vines, T., Waits, L., Widmer, A., & Rieseberg, L. (2013). A road map for molecular ecology. *Molecular Ecology*, *22*(10), 2605-2626.

Agilent Technologies (2014). Retrieved on 30 April 2014. http://www.home.agilent.com/agilent/home.jspx?&cc=US&lc=eng

Amresco Biochemicals and Life Science Research Products (2014). Retrieved 30 April 2014. http://www.amresco-inc.com/AGAROSE-3-1-E776.cmsx

Avise, J.C. (1994). Molecular markers, natural history, and evolution. *Springer*.

Bader, J. M. (1998). Measuring genetic variability in natural populations by allozyme electrophoresis. *Tested studies for laboratory teaching*, *19*, 25-42.

Beebee, T.J.C., & Rowe, G. (2008). An introduction to molecular ecology. *Oxford University Press*.

Belfiore, N.M., & May, B. (2000). Variable microsatellite loci in red swamp crayfish, *Procambarus clarkii*, and their characterization in other crayfish taxa. *Molecular Ecology*, *9*(12), 2231-2234.

Blouin, M.S. (2003). DNA-based methods for pedigree reconstruction and kinship analysis in natural populations. *Trends in Ecology & Evolution*, *18*(10), 503-511.

Blouin, M.S., Yowell, C.A., Courtney, C.H., & Dame, J.B. (1998). Substitution bias, rapid saturation, and the use of mtDNA for nematode systematics. *Molecular Biology and Evolution*, *15*(12), 1719-1727.

Bonin, N., Bellemain, E., Eidesen, P.B., Pompanon, F., Brochmann, C., & Taberlet, P. (2004). How to track and assess genotyping errors in population genetics studies. *Molecular Ecology*, *13*(11), 3261-3273.

Bracken, H.D., Toon, A., Felder, D.L., Martin, J.W., Finley, M., Rasmussen, J., Palero, F., & Crandall, K.A. (2009). The decapod tree of life: compiling the data and moving toward a consensus of decapod evolution. *Arthropod Systematics & Phylogeny*, *67*(1), 99-116. Broquet, T., & Petit, E. (2004). Quantifying genotyping errors in noninvasive population genetics. *Molecular Ecology*, *13*(11). 3601-3608.

Clementino, C.S., Barbosa, F.J.V., Carvalho, A.M.F., Costa-Filho, R.A.R., Silva, G.R., Campelo, E.G., Britto, F.B., & Diniz F.M. (2010). Microsatellite DNA loci for population studies in Brazilian chicken ecotypes. *International Journal of Poultry Science*, *9*(12), 1100-1106.

Comparative Genomics Center (2012). Automated Fluorescent DNA Fragment Analaysis in the CGC. *University of Washington*. http://depts.washington.edu/genomelb/

Dewoody, K., Nason, J.D., & Hipkins VD. (2006). Mitigating scoring errors in microsatellite data from wild populations. *Molecular Ecology*, *6*(4), 951-957.

Estoup, A., Jarne, P., & Cornuet, J. M. (2002). Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology*, *11*(9), 1591-1604.

Fazekas, A.J., Steeves, R., & Newmaster, S.G. (2010). Improving sequencing quality from PCR products containing long mononucleotide repeats. *BioTechniques*. *48*(4), 277-285.

Fetzner, J.W., & Crandall, K.A. (2001). Genetic Variation. *Biology of Freshwater Crayfish*, 291-327.

Gerber, S., Mariette, S., Streiff, R., Bodenes, C., & Kremer, A. (2000). Comparison of microsatellites and amplified fragment length polymorphism markers for parentage analysis. *Molecular Ecology*, *9*(8), 1037-1048.

Girard, P. (2011). A robust statistical method to detect null alleles in microsatellite and SNP datasets in both panmictic and inbred populations. *Statistical applications in genetics and molecular biology*, *10*(1), 1-10.

Glenn, T. C., & Schable, N. A. (2005). Isolating microsatellite DNA loci. *Methods in enzymology*, *395*, 202-222.

Gouin, N., Grandjean, F., & Souty-Grosset, C. (2006). Population genetic structure of the endangered crayfish Austropotamobius pallipes in France based on microsatellite variation: biogeographical inferences and conservation implications. *Freshwater Biology*, *51*(7), 1369-1387.

Grant, P., & Grant, R. (2008). How and why species multiply: the radiation of Darwin's finches. *Princeton University Press*, 994-995.

Harrison, H.B., Sa Enz-Agudelo, P., Planes, S., Jones, G.P., & Berumen, M.L. (2013). Relative accuracy of three common methods of parentage analysis in natural populations. *Molecular Ecology*, *22*(4), 1158-1170.

Hearne, C.M., Soumitra, G., & Todd, J.A. (1992). Microsatellites for linkage analysis of genetic traits. *Trends in Genetics*, *8*(8), 288-294.

Hecker, K. H., & Roux, K. H. (1996). High and low annealing temperatures increase both specificity and yield in touchdown and stepdown PCR. *Biotechniques*, *20*(3), 478-485.

Hulak, M., Kašpar, V., Kozak, P., Buric, M., Filipova, L., & Petrusek, A. (2010). Cross-species amplification of microsatellite markers in the invasive spiny-cheek crayfish (*Orconectes limosus*): assessment and application. *Journal of Applied Genetics*, *51*(1). 73-78.

Jarne, P., & Lagoda, P.J.L. (1996). Microsatellites, from molecules to populations and back. *Trends in Ecology & Evolution*, *11*(10), 424-429.

Jones, A.G., Small, C.M., Paczolt, K.A., & Ratterman, N.L. (2010). A practical guide to methods of parentage analysis. *Molecular Ecology*, *10*(1), 6-30.

Li, Y.C., Korolo, A.B., Fahima, T., Beiles, A., & Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology*, *11*(12), 2453-2465.

Li, Y., Wang, W., Liu, X., Luo, W., Zhang, J., & Gul, Y. (2011). DNA extraction from crayfish exoskeleton. Indian *Journal of Experimental Biology*, *49*(12), 953-957

Lima, V.L.A., Seki, H.A., & Rumjanek, F.D. (2003). Microsatellite polymorphism in wheat from Brazilian cultivars; inter- and intra-varietal studies. *Genetics and Molecular Biology*, *26*(3), 349-353.

Miller, M.J., & Yuan, B. (1997). Semiautomated Resolution of Overlapping Stutter Patterns in Genomic Microsatellite Analysis. *Analytical Biochemistry*, *251*(1), 50-56.

Monsen-Collar, K. J., & Dolcemascolo, P. (2010). Using Molecular Techniques to Answer Ecological Questions. *Nature Education Knowledge, 1*(1).

Nachman, M.W., & Crowell, S.L. (2000). Estimate of the Mutation Rate per Nucleotide in Humans. *Genetics. 156*(1). 297-304.

Noda, R., Kim, C. G., Takenaka, O., Ferrell, R. E., Tanoue, T., Hayasaka, I., Ueda, S., Ishida, T., & Saitou, N. (2001). Mitochondrial 16S rRNA sequence diversity of hominoids. *Journal of Heredity*, *92*(6), 490-496.

Queller, D.C., Strassmann, J.E., & Hughes, C.R. (1993). Microsatellites and Kinship. *Trends in Ecology & Evolution*, *8*(8), 285-288.

Selkoe, K.A., & Toonen, R.J. (2006). Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, *9*(5), 615-629.

Shinde, D., Lai, Y., Sun, F., & Arnheim, N. (2003). *Taq* DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: $(CA/GT)_n$ and $(A/T)_n$ microsatellites. *Nucleic Acids Research*, *31*(3), 974-980.

Stenerson, M. 2002. All About Mutations. Retrieved April 29, 2014 from HOPES Stanford University: https://www.stanford.edu/group/hopes/cgi-bin/wordpress/2011/02/all-about-mutations/

Walker, D., Porter, B. A., & Avise, J. C. (2002). Genetic parentage assessment in the crayfish *Orconectes placidus*, a high-fecundity invertebrate with extended maternal brood care. *Molecular Ecology*, *11*(10), 2115-2122.

Walsh, P.S., Fildes, N. J., & Reynolds, R. (1996). Sequence Analysis and Characterization of Stutter Products at the Tetranucleotide Repeat Locus vWA. *Nucleic Acids Research*, *24*(14), 2807-2812.

Wang, C., Schroeder, K.B., & Rosenberg, N.A. (2012). A Maximum-Likelihood Method to Correct for Allelic Dropout in Microsatellite Data with No Replicate Genotypes. *Genetics*, *192*(2), 651-669.

Wieczorek, A.M., & Geber, M.A. (2002). Microsatellite loci for studies of population differentiation and range expansion in *Solidago sempervirens* L. (Asteraceae). *Molecular Ecology*, *2*(4), 554-556.

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

Zhu, Z.Y., & Yue, G.H. (2008). Eleven polymorphic microsatellites isolated from red swamp crayfish, *Procambarus clarkii*. *Molecular Ecology Resources*, *8*(4), 796-798.