

Genetic Diversity in Microsatellite Regions of *Cypripedium parviflorum*

Tiffany Sun

I. Abstract

Many of the 58 species of the *Cypripedium* orchids that are located all around the world are increasingly approaching extinction. The first step to ensure preservation is to assess the genetic diversity of populations to ensure that they will survive even under sudden environmental stresses. Genetic diversity was assessed by examining microsatellite regions of their genomes, which are short repetitive sections of the DNA that typically repeat 5-50 times. A goal of this research is to test the effectiveness of the primers developed by another group for the *Cyp. tibeticum* lady's slipper orchid on *Cyp. parviflorum* (yellow lady's slipper) samples, to determine which primers bind to the DNA successfully. The primers that produce positive results yield amplified DNA that can be sequenced to assess genetic diversity among *Cyp. parviflorum* samples. All samples undergo processes of DNA extraction, PCR, gel electrophoresis, and DNA cleanup, to prepare the sample for sequencing. Ten primer sets were tested and among them, two produced clean chromatograms. This does not yet give us enough genomic data to draw reliable conclusions about the relatedness among individuals. Of the number of samples tested, many of the sequences produced with the same primers were almost identical. We will re-test many primer sets that did not produce consistent results, and then we can further assess diversity among plants in different areas. We may also test more primer sets and collect more samples to extend this study.

II. Introduction

Cypripedium is a genus of 58 species of hardy orchids. It is one of five genera that compose the subfamily of lady's slipper orchids, which can be found in many ranges of climates worldwide. Despite being very hardy and able to tolerate temperate climates, some of the *Cypripedium* species have become endangered, including species in New England.

These orchids reproduce by two means, one being sexual reproduction through pollination and the second, vegetative reproduction, through underground rhizomes. For several reasons including low germination rate and an 8-year time from germination to flowering, sexual reproduction is often unsuccessful (1). In addition, *Cypripedium* flowers do not produce nectar and depend on small insects such as syrphid flies, beetles, and mason bees for pollination. A single seed pod can produce around 200,000 seeds per seedpod seeds, all of which are very lightweight with no endosperm, and are dispersed typically by wind. Since very few (less than 0.1% of them will germinate and reach a flowering state, sexual reproduction is very ineffective (7). On the other hand, vegetative reproduction is where a plant reproduces through the underground stems called rhizomes, the plant grows into another plant with identical genes. Usually, underground rhizomes are about 2-4 centimeters underneath the soil, with roots growing out horizontally. The rhizomes produce genetically identical plants of the parent plant.

Enough genetic variation is a fundamental principle of species fitness and for a species survival so that natural selection can select the fittest individuals from the population, adapting the population to their

environment to survive. Even so, some conserved sequences that mutate slower are needed because there are codes that are very important and change is likely to be detrimental to the organism. Other genomic regions mutate more rapidly since they are less critical to the immediate survival of the plant.

An essential step in the conservation of this plant is to assess the genetic diversity in samples taken from the *Cypripedium* plants. Because the genus is rapidly reaching extinction, it has become a topic of interest in our lab, and we are constantly striving to find ways in which we can preserve it. To ensure preservation, we must understand the genetic diversity among plants which can be done by analyzing microsatellite regions, sections of repetitive DNA where short segments are repeated.

Our analysis will be done using samples from the *Cypripedium parviflorum* which is an endangered species in the northeastern US. Samples were collected from a single forest of *Cyp. parviflorum* in Strafford VT with two from clumps that were about 70 feet apart, and three from clumps that were 3 feet apart. For both locations, samples were collected from plants in clumps separated by large distances to compare genetic diversity in plants close to each other and plants farther apart. We tested 10 primers on the *Cyp. parviflorum* samples to find out whether or not they are effective at binding to the DNA and used the ones that worked to analyze the genetic diversity among the plants' genomes. The microsatellite primers sequences used were 10 out of 20 sets developed for another *Cypripedium* species, *Cyp. tibeticum* (2).

Our current hypothesis is that plants that are nearby to each other will show less genetic diversity than plants that are a farther distance apart (in different clumps). This conjecture was based on the understanding that plants that are farther apart may have been spread by seed germination which will most likely possess more diversity. Plants that are nearby to each other (1-2 feet) may have grown from vegetative reproduction with the same genes as the parent plant, resulting in less diversity.

III. Materials and Methods

Lady's Slipper Samples

We collected five samples from five different clumps, two of which (Y1, Y2) were from around 70 feet apart, and three of which (Y3, Y4, Y5) were from clumps three feet apart (Figure 1).

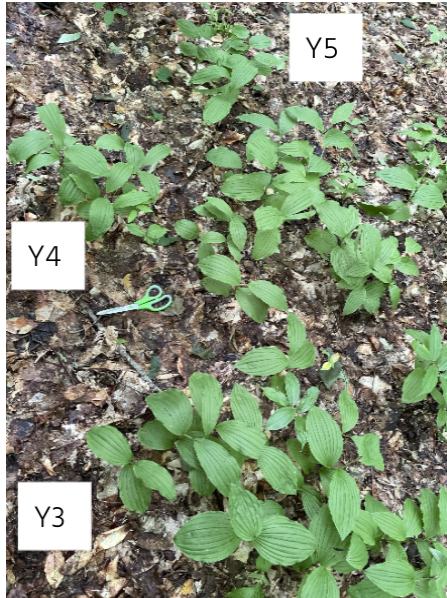


Figure 1. Three clumps of *Cyp. parviflorum* approx. 3 feet apart.

Procedures

These samples underwent processes of DNA extraction, PCR, gel electrophoresis, and DNA cleanup. The purpose of all these steps is to amplify and prepare the samples to sequence them. The purified PCR samples are then sent to Eton Biosciences for DNA sequencing, and come back as a chromatogram which we will analyze to assess for genetic diversity.

Materials

Common materials include a centrifuge, vortex mixer, micropipettes, various sized tips, tube racks for 1.5mL tubes and PCR tubes, and safety supplies such as goggles, gloves, etc. For DNA extraction we will use the Monarch Genomic DNA Purification Kit from New England Biolabs (3) to discard all unwanted proteins and fibers in the plant so that we only have the DNA left. Additionally, we will crush the leaf sample with a micro pestle before following the ‘tissue’ protocol to help break down the cells and fibers. We will skip the step ‘Incubate at 56°C in a thermal mixer with agitation at full speed (1400 rpm)’, and instead use a water bath, and occasionally a vortex mixer to mix the contents.

PCR was performed on a total of 55 tubes, 5 samples x 11 primers (10 *Cyp. tibeticum* primer sets (2), and 1 Chloroplast primer set (5)). Each tube consists of 25µL New England BioLabs Master Mix (OneTaq Quick Load 2X Master Mix with Standard Buffer), 22µL sterile water, and 2µL forward and reverse primers added to 1µL of DNA. The tubes were then put into either a Bio-Rad or a miniPCR thermal cycler to perform the PCR process. PCR was also performed using a GC Buffer instead of the Standard Buffer, to see if any of the samples worked better with a buffer that targets more GC-rich templates. The purpose of the PCR is to make multiple copies of a segment of DNA to amplify the DNA enough so that we can analyze it. However, most of the PCR processes were done with 11 primers and 5 DNA samples, so an “Ultimate Mix” was created that could be used for all of the tubes, excluding the primers, which were later added to the tubes individually. The Ultimate Mixes usually consisted of 1.1n the amounts required for single reactions to account for pipetting error. For example, in an Ultimate Mix used for 11 reactions, 12.1 times the amount of each element was added.

The thermal cycler was set in the following cycle, and the steps denaturation, annealing, and extension were repeated 35 times:

Initial Denaturation: 95°C, 120 seconds

Denaturation: 95°C, 30 seconds

Annealing: 49°C, 30 seconds

Extension: 68°C, 30 seconds

Final Extension: 68°C, 60 seconds

After performing PCR, gel electrophoresis using 2% gels was performed on all of the PCR samples to identify which primers will successfully bind to the specified regions of DNA. A BlueGel gel electrophoresis system was used. In addition to the materials included, agarose, running buffer, GelGreen DNA stain, New England BioLabs 100bp ladder, scale, weigh boats, and a 250 mL Erlenmeyer flask was used. The PCR and gel electrophoresis were repeated three times to ensure consistency of results.

Finally, the DNA Cleanup process is very similar to the extraction process, and its purpose is to prepare the samples for sequencing by removing everything except the amplified DNA. We use the Monarch PCR & DNA Cleanup Kit from New England Biolabs and followed all protocols (6). After this step was completed, samples were sent to Eton Biosciences for sequencing.

Primers

The first ten primer sequences are the *Cyp. tibeticum* microsatellite primers published in Li et al., 2017 (2). The last two chloroplast primers are from Taberlet et al., 1991 (5)

M112 forward: AAGACTGGAAGTAGCGGAGAT

M112 reverse: GAAGTCTACATCCACTGTTCCT

M130 forward: TTCAGCAGATTCAGATTTAGCT

M130 reverse: CTTCTGACAGATCTTCTCCAAT

M136 forward: TCCAACCTCTCTATCCTCGAT

M136 reverse: GATCTGAGGAGGATGAAGGT

M139 forward: GAGGGAGGATTCTGACAAC

M139 reverse: GTGGTTCGACTTCGTGAAAT

M142 forward: TTGTGCATCGTCATCAAGGT

M142 reverse: GATCACAAACATGCTGGGAT

M233 forward: ACACGCTACAGACCGAGCT

M233 reverse: CTTGTTTGCAAGATGAATGCT

M370 forward: GGAGATTCAGTTATCAGAGCTG

M370 reverse: AGTCTACATCCACTATTCCTTGT

M681 forward: ATGCCAGACGGCAGAGGT

M681 reverse: GGCTTTGAGCTAATGGCTCT

M880 forward: GGAGAACAGGAAAAGTGTTCAT

M880 reverse: AGCCGATTACAAAGTCACTGT

M886 forward: CGAAATTCACCTGGGATTCAGT
M886 reverse: TCTCAGGATTACAACAGTAGCT
Chloroplast forward CGAAATCGGTAGACGCTACG
Chloroplast reverse GGGGATAGAGGGACTTGAAC

IV. Results

Gel Electrophoresis

Results from testing primers were obtained by looking for visible bands that appear during gel electrophoresis. 10 primer sets were tested (M112, M130, M136, M139, M142, M223, M370, M681, M880, M886), and used on five *Cyp. parviflorum* samples (the 'CL' represents our chloroplast primer, which is used as a control to ensure that any issues were with primers binding to DNA, and not with procedures or machines). Only six of the ten primers showed positive results. Below are images of the bands from all five samples (Figure 2, 3, 4).

Figure 2. Gel results for samples Y1 and Y2, the leftmost lane is the 100bp ladder, as is the rightmost lane on both gels. The primer sets in red are the ones we sent for sequencing, and the ones in green are the ones that gave clean chromatograms. For each primer set, the first lane is with Y1 DNA and the second lane is with Y2 DNA.

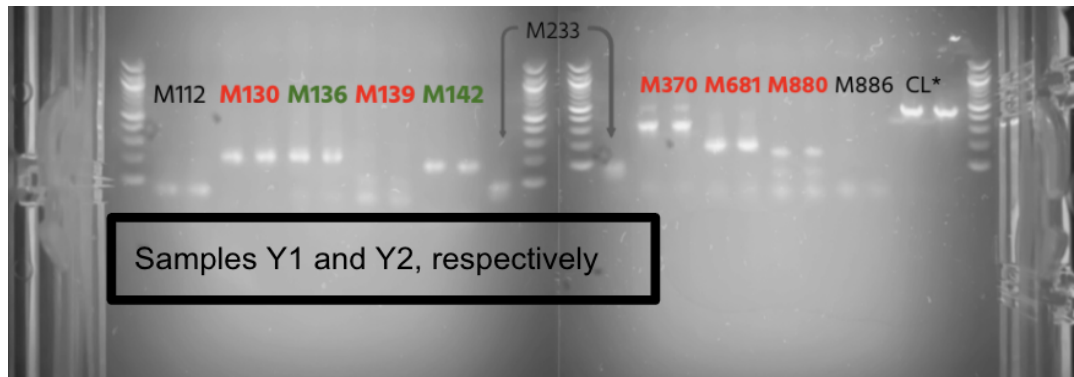


Figure 3. Gel results for samples Y3 and Y4, 100bp ladder is on the left and right sides of the gel.

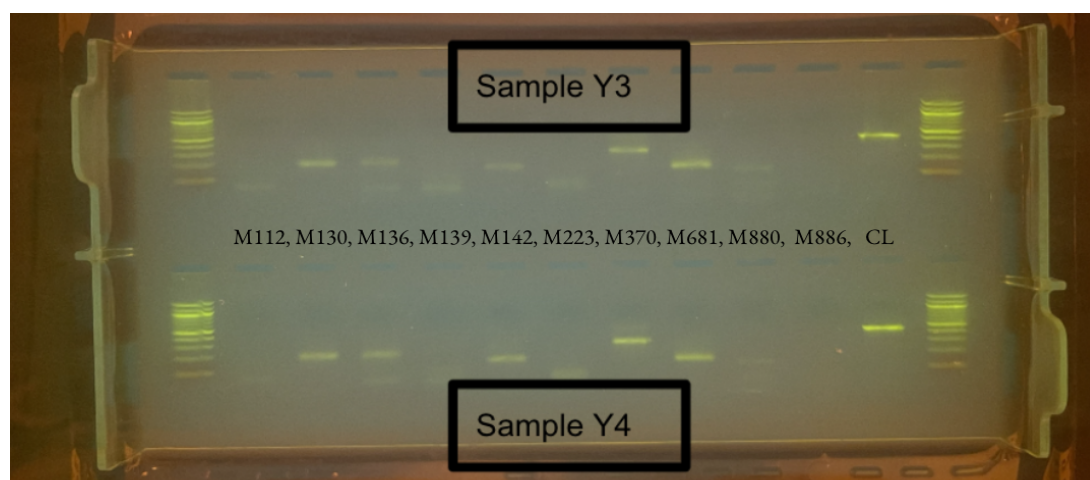
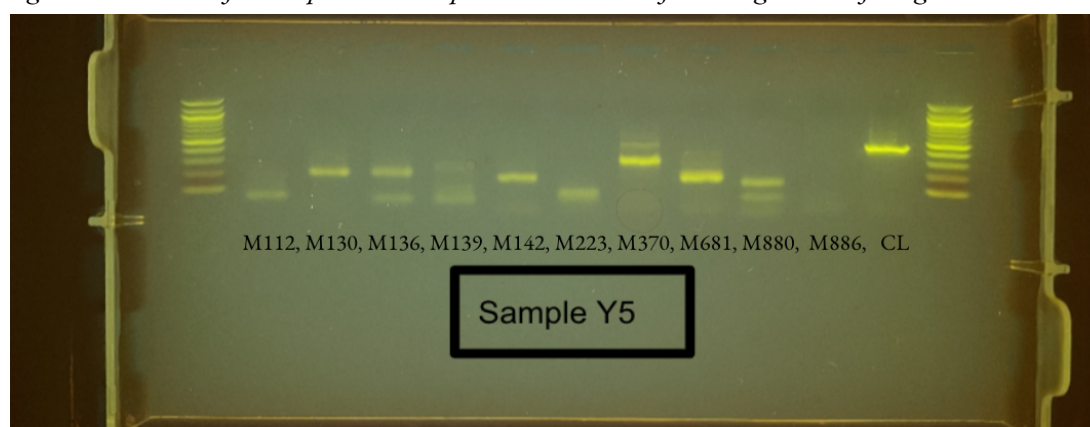


Figure 4. Gel results for samples Y5, 100bp ladder is on the left and right sides of the gel



Sequencing Results

All samples labeled in the format: Sample_Primer

Figure 5. All 5 chromatograms of the M136 primer with all 5 samples, and the alignment of the sequences. The sequence is decent with few double peaks and it looks like a microsatellite region based on ~110-140bp.

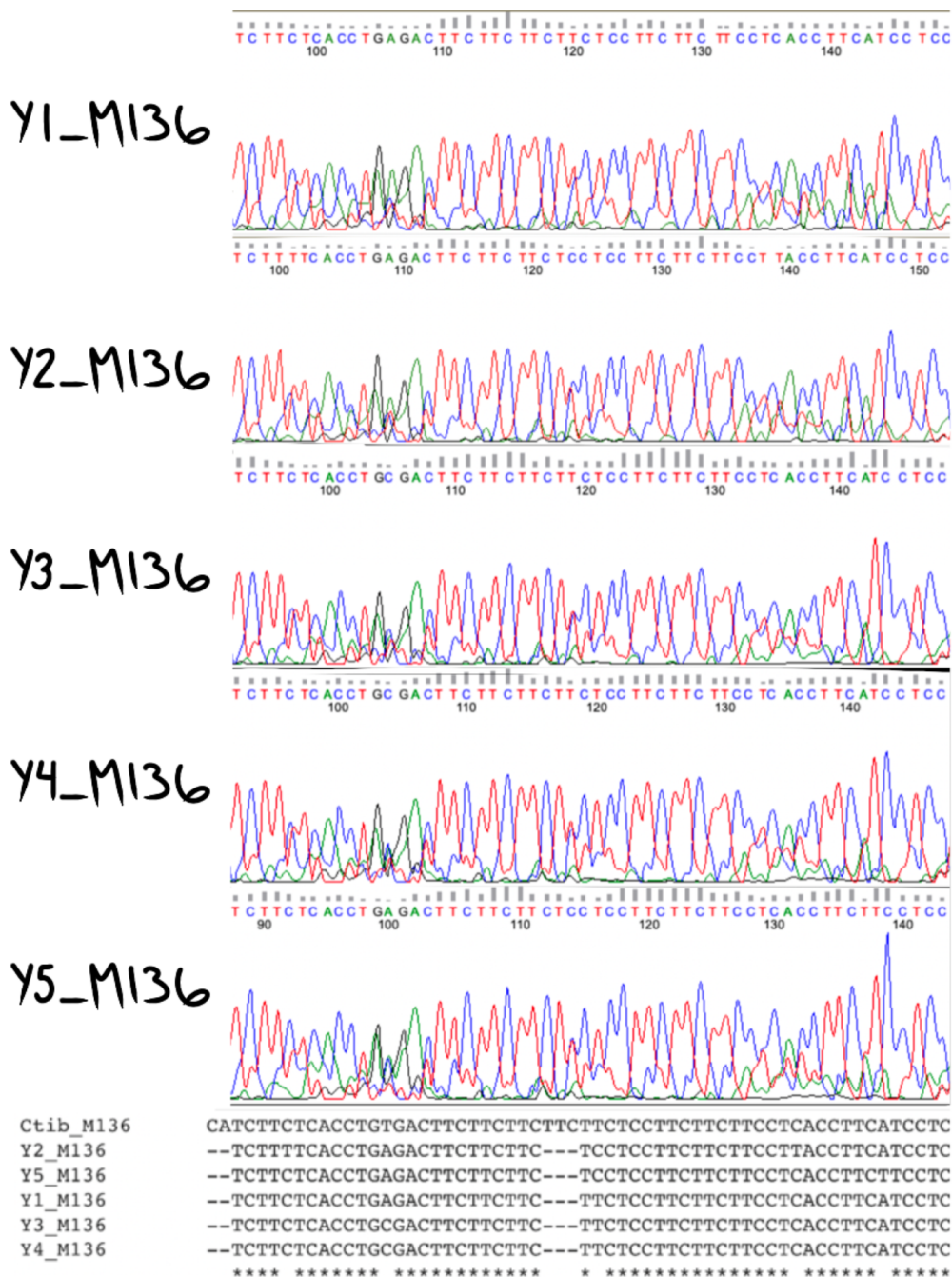
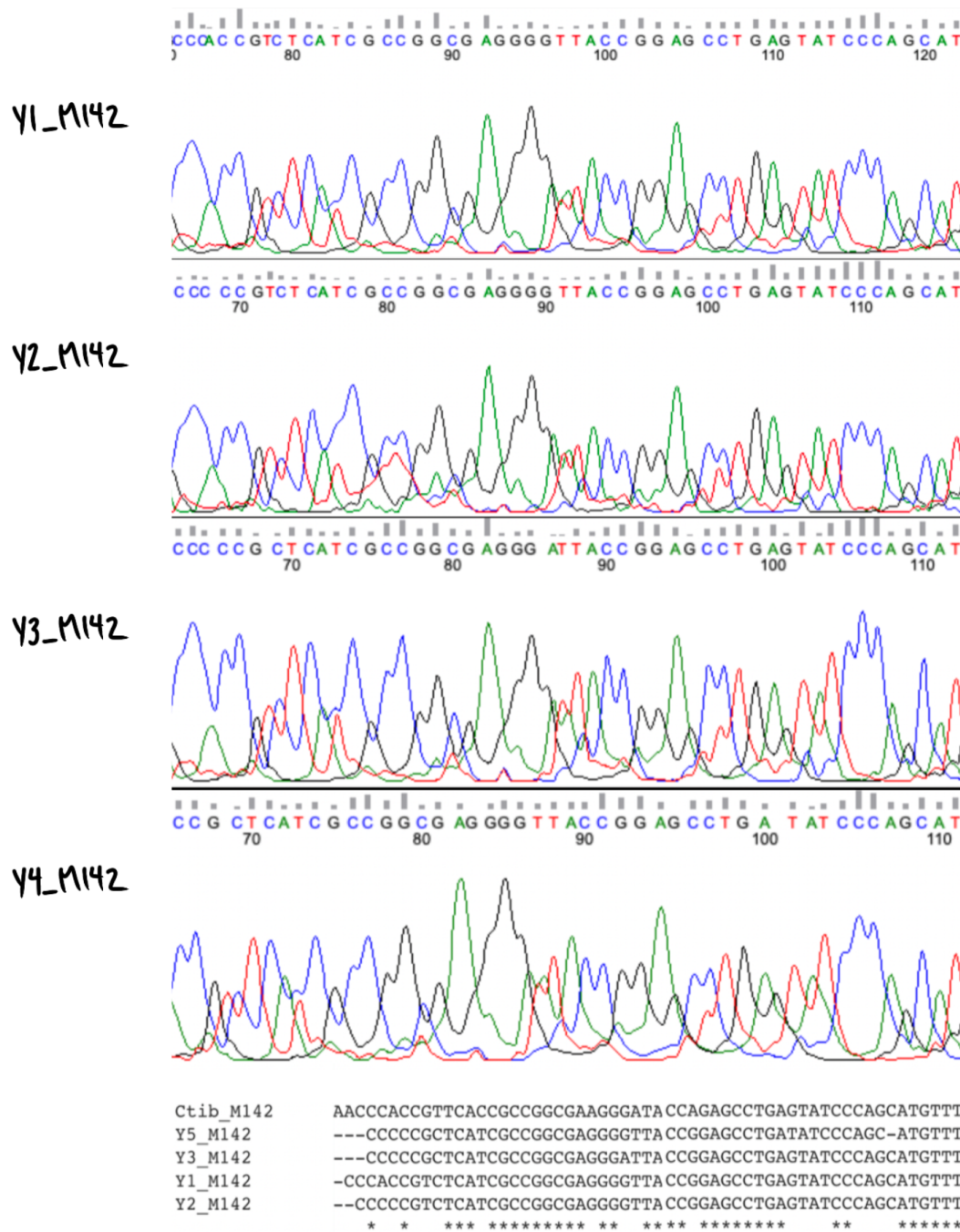


Figure 6. 4 of the clean chromatograms of the M142 primer with only Y1, Y2, Y3, and Y5, and the alignment of the sequences. These sequences are not as good with many more double peaks and messier.



The samples that were sent for sequencing and were unusable due to messy chromatograms are:

All samples, M130

All samples, M139

Y4_M142

All samples, M370

All samples, M681

All samples, M880

V. Discussion

Primer Tests

The primer testing on the 10 primers from Li et al (2017) on *Cypripedium parviflorum* was successful. Seven of the primers, when tested on 5 samples each, all displayed positive results during gel electrophoresis: M130, M136, M139, M142, M370, M681, M880 (Figures 2, 3, 4). These can be used as a base for future research and studies to analyze genetic diversity in *Cyp. parviflorum* populations.

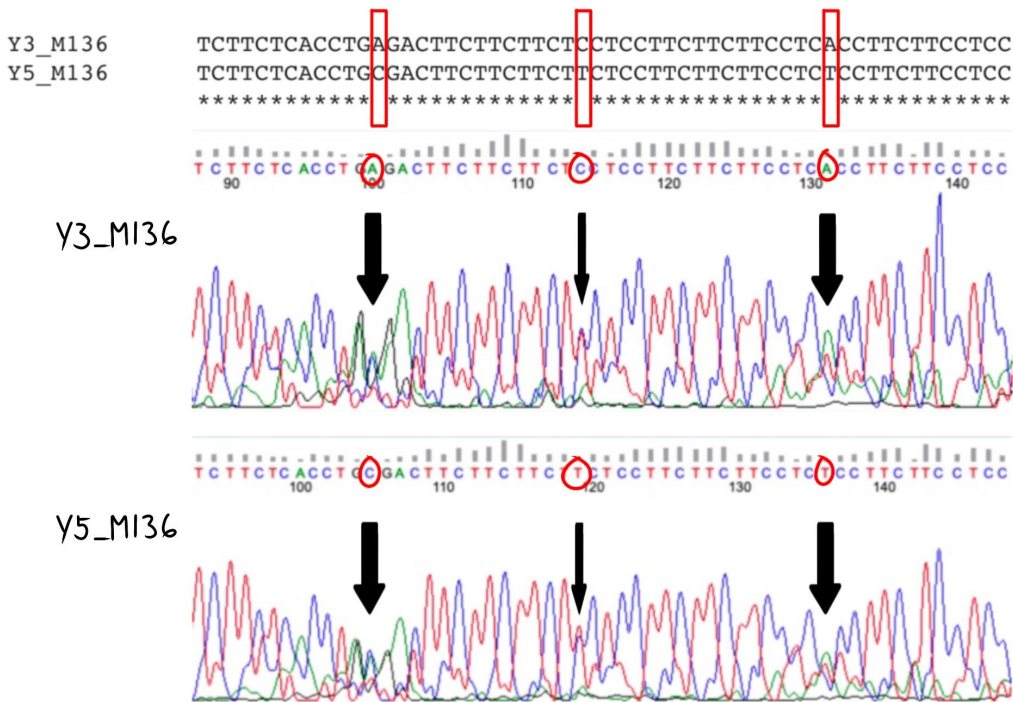
In the future, we will use these results to analyze the chromatograms from PCR and retest primers that yielded bands but did not give clean chromatograms. In addition, more primers from the same paper (2) may also be tested as we only tested ten of the primer sets which leaves another ten that were published in that paper that we did not test due to time constraints this summer though these were being followed up on by another student.

Genetic Diversity

Of the 33 samples sent to Eton Biosciences across all five yellow samples used here, only 9 of them returned with clean results, and only two of the primers sent out had usable chromatograms: M136, M142, which both matched to the *Cyp. tibeticum* microsatellite regions. With very little usable genetic data, it is hard to draw any reliable conclusions and would require more experiments, although these results can be used for future research with the same samples, or samples from different clumps.

Although we did not get enough data to thoroughly assess the genetic diversity, we can compare multiple chromatograms using the same primers, but different plant samples. The samples Y1_M136, Y2_M136, Y3_M136, Y4_M136, Y5_M136 are almost identical (Figure 5), and Y1_M142, Y2_M142, Y3_M142, Y5_M142 are also very similar (Figure 6). Below is an image of the alignment of the Y3_M136 and Y5_M136 chromatograms and sequence (Figure 7).

Figure 7. Comparison between Y3_M136 and Y5_M136 alignment and chromatograms, looking for double bands.



These two chromatograms are almost identical with clear evenly spaced peaks, and some baseline noise. The black arrows show bases that were marked as different on the alignment and as we refer back to the chromatogram, these all had double peaks. This only occurred 3 times in this section, and they are both almost identical. The region from roughly, 110-140 is highly repetitive, showing signs of microsatellite regions. Another thing we look for in chromatograms are double peaks. These occur when there are two or more bases present, and the absence of an asterisk on the alignment is a sign of possible double peaks. Looking back at the chromatograms and looking closer into differences is essential for achieving accurate results. These positions are likely heterozygous in both plants, meaning that different copies of the microsatellite region have different bases in each plant. In this case, all positions missing asterisks show double peaks on both chromatograms (Figure 7). The sequencing software may identify one base or the other as the sequence, but the plants show identical heterozygosity at all highlighted positions.

Future Directions

We will be re-testing samples with unsuccessful samples in these rounds of sequencing, and continue to experiment with more primer sets, and samples, so we can draw accurate conclusions about the diversity. Additionally, we can perform gel extraction on the primers that showed double bands during the gel electrophoresis stage (for example primer M370 in Figure 2). Double bands signify that there may have been two or more DNA templates present, and these need to be separated for sequencing. Performing gel extraction can then prepare the samples again for sequencing to receive more reliable results.

VI. Acknowledgements

I would like to thank Dr. Kelly Salmon and Dr. Peter Faletra for their mentorship, and the New Hampshire Academy of Science Lab for making this program possible and for providing all the materials used in this research.

VII. References

- (1) Sokolski, K., A. Dovholuk, L. Dovholuk, and P. Faletra. "Axenic Seed Culture and Micropropagation of *Cypripedium Reginae*." *Selbyana* 18.2 (1997): 172-182.
- (2) Li, J., Luo, Y. and Xu, L., 2021. "Development of Microsatellite Markers for *Cypripedium tibeticum* (Orchidaceae) and their Applicability to Two Related Species."
- (3) Monarch Genomic DNA Purification Kit Protocol:
<https://www.neb.com/protocols/2019/04/30/quick-protocol-for-extraction-and-purification-of-genomic-dna-using-the-monarch-genomic-dna-purification-kit-neb-t3010>
- (4) MiniPCR Gel electrophoresis Setup Protocol:
<https://www.youtube.com/watch?v=beVVuyTcPAc>
- (5) Taberlet P, Gielly L, Pautou G, Bouvet J. Universal primers for amplification of three non-coding regions of chloroplast DNA. *Plant Molecular Biology*. 1991;17:1105–1109.
<https://pubmed.ncbi.nlm.nih.gov/1932684/>
- (6) Monarch PCR & DNA Cleanup Kit Protocol:
<https://www.neb.com/protocols/2015/11/23/monarch-pcr-and-dna-cleanup-kit-protocol>
- (7) White, Caren. "How To Grow Lady'S Slippers, A Native Plant". Dengarden, 2021,
<https://dengarden.com/gardening/How-to-Grow-Ladys-Slippers-a-Native-Plant#:~:text=Once%20fertilized%2C%20seed%20pods%20are,growth%20of%20their%20underground%20rhizomes.>