THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOLOGY

DNA FINGERPRINTING AND GENOTYPING IN PRUNUS SEROTINA

DANIEL J. DEVINE
Spring 2010

A thesis
submitted in partial fulfillment
of the requirements
for a baccalaureate degree
in Biology
with honors in Biology

Reviewed and approved* by the following:

John E. Carlson
Professor of Molecular Genetics
Thesis Supervisor

Richard J. Cyr
Professor of Biology
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

Abstract:

The main goal of this study was to perform DNA fingerprinting and identify full siblings in a population of black cherry (*Prunus serotina*) seedlings, an economically and ecologically valuable hardwood tree species native to the eastern United States and invasive in Europe. Fingerprinting was performed on 600 samples from the ozone sensitive R-14 family of the black cherry. To fingerprint, SSR markers previously identified for a black cherry EST library and found to be polymorphic were used. DNA was extracted from the 600 samples and was used to perform PCR followed by genotyping. At this point in time, 2 markers (131 and 654) have been found to be polymorphic in black cherry. Once an additional 4 to 8 polymorphic markers are found, fingerprinting for the 600 samples will ensue. This data will then be used to establish linkage maps for quantitative trait loci related to ozone in black cherry.

TABLE OF CONTENTS

Introduction:

The Black Cherry Tree

Prunus serotina, more commonly known as the black cherry tree, is found in the eastern United States and Canada (Liu, Anderson, & Pijut, 2009). Though indigenous to North America, the black cherry was introduced to European Forests where it has taken over as an invasive species. *P. serotina* are able to out-compete native species in poorer soils and have thus taken over entire forests. Researchers believe the species ability to have a single year's seed crop germinate over a period of three years enhances the species fitness over its natural, European counterparts (Phartyal, Godefroid, & Koedam, 2009). Because of the high quality and valuable hardwood the black cherry tree produces, demand is ever increasing (Liu, Anderson, & Pijut, 2009).

Mature trees grow to be between 60 to 90 feet in height and have an oval silhouette. Leaves and twigs of the black cherry tree contain hydrocyanic acid, which, when consumed in large quantities, can poison livestock and other animals. Aside from its prized wood, the black cherry tree's fruit is often used by industry to make jams and liqueurs (Gilman & Watson, 1994). Seeds are believed to be distributed by birds and small mammals which consume the fruit. In terms of environmental stressors, *P. serotina* is known to be sensitive or moderately tolerant to ozone and drought resistant, as its roots explore large volumes of soil. Because of an expansive root structure, the black cherry tree does not fare well in urban environments (Gilman & Watson, 1994).

Ozone

In this study, two families of *P. serotina*, R-14 and MO-21, were analyzed. Previous research has shown samples from the R-14 family are more tolerant to ozone exposure than 13

other families (Lee, Steiner, Zhang, & Skelly, 2002). Previous research has also shown the family MO-21 to be less tolerant to ozone exposure than 13 other families (Lee, Steiner, Zhang, & Skelly, 2002). A goal of this project is to examine leaf samples from 600 black cherry trees in more-ozone-tolerant R-14 family. Ozone is believed to be a contributor to the increase in oxidative stress of many tree species in Pennsylvania and much of the Allegheny Plateau during the later part of the twentieth century.

Tropospheric ozone is a greenhouse gas and air pollutant that affects both the health of humans and vegetation. The effects and susceptibility to ozone varies between tree species (Ashmore, 2005). Although the variation in ozone response on a physiological and genetic basis is not yet understood, it is known that important implications exist for ozone tolerance and sensitivity within a species. The loss of more sensitive genotypes to genotypes with enhanced fitness due to a tolerant genotype can lead to adverse natural selection and thus the elimination of trees with sensitive genotypes (Karnosky, 1980).

P. serotina is known to be particularly sensitive to ozone and is sometimes used as a bio-indicator for ozone pollution. This, in conjunction with high levels of anthropogenic ozone found to be present in Pennsylvania, threatens the viability of a sustainable black cherry tree population in the Allegheny Plateau region (Lee, Steiner, Zhang, & Skelly, 2002).

Current Research on Black Cherry Genomics

Studying the genomic affects of ozone and identifying other genes of economic importance (wood quality, pest resistance, germination, etc.) requires the mapping of the *P. serotina* genome as well as construction of genetic linkage maps. Currently, an EST library for the black cherry tree has been developed and will be used to find microsatellite markers for

further genomic studies, including the present one. A framework for study is provided by previous genomic analyses of other *Prunus* species such as the peach and almond trees.

DNA fingerprinting

DNA fingerprinting encompasses a set of techniques that allows the examination of the deoxyribonucleic acid (DNA) and distinguish an individual from the rest of a population. At the University of Leicester in the 1980's, Sir Alec Jeffreys developed DNA fingerprinting as a means to identify an individual. In terms of humans, DNA fingerprinting can be used to positively identify an individual because there are more possible combinations of the nucleotides adenine, guanine, thymine, and cytosine than there are people on earth (Lerner & Lerner, 2004). In a study where a researcher was attempting to study the genetic diversity of Indian populations, DNA fingerprinting was used. For analysis, the researcher used the complete mitochondrial DNA sequences as well as STRs and short nuclear polymorphisms (SNPs) that were found on the Y chromosomes of 14,000 individuals from over 200 ethnic populations. Through this, the researcher was able to trace back the lineage of the populations where it was discovered that most of the populations came from a common source in Southeast Asia (Singh, 2009).

In plants, scientists have used DNA fingerprinting as a means to find improvements for plant breeding. Using DNA fingerprinting techniques, scientists are able to figure out which genes play a role in physiological processes, and are able to figure out which plants possess said genes (Scholl, 2001). To find the "fingerprint", scientists focus on specific areas on the DNA where nucleic acid sequences vary significantly from individual to individual. These areas, known as polymorphisms, are a result of mistakes in the DNA replication process or environmental agents (Scholl, 2001). In the early years of DNA fingerprinting, Jeffreys used

variable number tandem repeats (VNTRs), which are part of the five to ten percent of DNA that contains repeats of the same nucleotide sequences (Lerner & Lerner, 2004). Restriction fragment length polymorphisms (RFLPs) have also been used as a tool for fingerprinting. RFLPs work on the basis that different individuals have different restriction fragment sizes, and by using gel electrophoresis one could discern between two individuals (Lerner & Lerner, 2004). Technological advances over the course of the past decade have yielded new tools for DNA fingerprinting. In modern times, short tandem repeats (STRs) are used for DNA fingerprinting. STRs are DNA sequences, usually between two and nucleotides long, which are repeated sequentially. Most STRs are found in a mostly unused area of DNA of each chromosome called heterochromatin (Beckman, 2003).

SSRs

Microsatellites, also known as simple sequence repeats (SSRs), are highly polymorphic sequences of nucleotides that are inherited co-dominantly. SSRs are widely used in laboratories because of their simplistic methods of development and their high level of observed heterozygosity (Aranzana, Pineda, & al., 2003). In this study, SSR markers derived from an initial EST database for the black cherry genome which is currently used for DNA fingerprinting. An EST, or Expressed Sequence Tag, is unique portion of DNA found within a coding region of a gene that is useful in identifying full length genes and also serves as a signpost for mapping. Polymerase chain reaction (see section below for further explanation) is a technique used to amplify specific DNA sequences of one's choosing. The PCR results are run on an agarose gel using gel electrophoresis to separate the DNA fragments based on size. During gel electrophoresis, shorter fragments of DNA travel further along the gel than larger, bulkier fragments. The fragments from PCR will create banding patterns on the gel which can

be analyzed for specific patterns which may match and indicate full siblings. Gels are used to determine amplifiable SSRs and GeneMapper® Software, a genotyping software package, provides DNA sizing and quality allele identification. With the identification of full siblings through PCR-based DNA fingerprinting, and by studying the frequency by which two genes are transmitted together, estimations can be made to determine how close they are on the chromosome. Information such as this is used to create what is referred to as a linkage map for black cherry.

A linkage map for the genus *Prunus* has been developed using a second generation cross from an almond tree and a peach tree. This reference map was made to be used as a reference for comparison between species as well as a basis for the development of framework and reduced maps. Maps such as the ones previously mentioned will aide in locating major genes of interest and it is believed that such maps are transferable to a certain degree between different species of the genus (Aranzana, Pineda, & al., 2003). Previous studies have used SSR markers from a central *Prunus* map to fingerprint sweet cherry trees (Wunsch & Hormaza, 2004). In this study, SSR markers from the *Prunus* reference map as well as other candidate SSR markers derived from an initial mapping of the black cherry tree genome will be used.

PCR

An important part of DNA fingerprinting is the use of the Polymerase Chain Reaction (PCR). Essentially the PCR process involves three steps: denature, anneal, extend. Before one can amplify a desired portion of the genome, DNA must be isolated from a sample (Polymerase Chain Reaction (PCR), 2009). In this study DNA extraction was performed using ground-up leaf

tissue, a CTAB extraction buffer, and phenol-chloroform (See Materials and Methods section for further explanation). Once isolated DNA is obtained, it is combined with a heat sensitive polymerase, nucleotides, and a desired sequence of DNA (Polymerase Chain Reaction (PCR), 2009). PCR is used to amplify a certain sequence on the sample genome so as it can be analyzed using agarose gel electrophoresis or a bioanalyzer. Depending on which nucleotides attach, the sample can be identified as homozygous or heterozygous. In this study, 6 polymorphic SSR markers will be used in an effort to fingerprint 600 samples of the ozone tolerant R-14 family and identify full siblings within the population.

Genotyping

With the use of PCR and a population of full siblings, one can genotype each sample tree in the population by analyzing SSR allele distribution between parent trees. To complicate analysis, previous research has shown the black cherry tree to be a tetraploid (Pairon & Jacquemart, 2008). Humans are diploid and thus can have up to 2 different alleles for each gene. In contrast, the tetraploid black cherry tree can have up to 4 different alleles for each gene. Though the black cherry tree is believed to be tetraploid, it is not uncommon for plants to be diploid, tetraploid, hexaploid, or even octaploid (De Silva, Hall, Rikkerink, McNeilage, & Fraser, 2005). A central issue with genotyping in the tetraploid black cherry, as is a problem in all polyploidy plants, is that SSR markers sometimes do not reveal the dosage of all of the alleles in a sample. This creates problems with calculating allele frequency (De Silva, Hall, Rikkerink, McNeilage, & Fraser, 2005).

Materials and Methods:

Samples

Black cherry seeds used in this study were obtained from open-pollinated trees found at the Penn Nursery of Pennsylvania's Bureau of Forestry. These trees had been used by two Penn State faculty members, John Skelly and Kim Steiner, in a study on multiple tree species on the effect of ozone pollution. Three families of black cherry were studied by the laboratory based on their varying resistance to ozone. Of the three families, the MO-21 family was tolerant to ozone, the R-14 was sensitive to ozone, and R-12 showed intermediate tolerance (Figure 1).

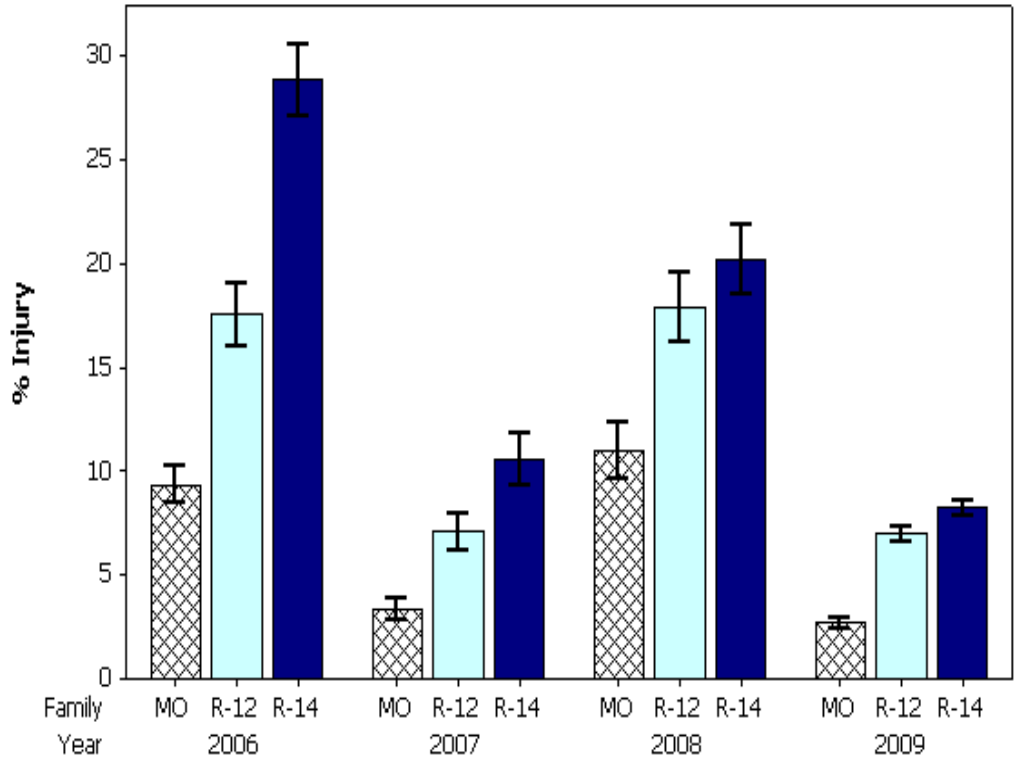

Figure 1: Graph showing percent ozone damage for each of 3 families. R-14 shows the most damage and MO-21 shows the least amount of damage (Lee, Steiner, Zhang, & Skelly, 2002).

The 600 samples used in this project came from the ozone tolerant genotype R-14. After the 600 stratified seeds germinated, they were planted and grown in a greenhouse (Figure 2). After ozone fumigations, leaves from each of the 600 seedlings were collected and stored for DNA extraction.

Figure 2: Photograph of greenhouse where samples were raised. Samples within chambers were exposed to ozone.

DNA extraction

A CTAB extraction buffer was used to isolate DNA based on a modified version of the CTAB (Cetyltrimethylammonium Bromide) extraction protocol developed by Doyle and Doyle. The CTAB buffer was prepared using the formula found in Table 1.

Table 1: Formula for CTAB extraction buffer.

CTAB-extraction for DNA	100 mL
1 M Tris- HCl pH = 8	5 mL
0.25 M EDTA pH = 8	5 mL
5.0 M NaCl	14 mL
1.0 M LiCl	40 mL
20% SDS	10 mL
CTAB (2%)	2.0 g
1.0 M Na_2SO_3	10 mL
PVP	5.0 g
DIECA	0.2 g
Ascorbic Acid	0.2 g
BME	20 µL
dd H_20	To reach final volume

The following protocol was used to complete the extraction:

1. Label 2 mL tubes
2. Add 1 mL extraction buffer to each tube and warm at 65°C
3. Add 20 µL of BME to each tube
4. In liquid nitrogen, grind samples and add to the labeled tubes containing extraction buffer.
    a. Vortex
    b. Keep it at 65°C for 30 minutes
5. Spin in microcentrifuge at full speed for 10 minutes
6. Pipette the supernatant into new tubes (1.5 mL)
    a. Supernatant can be between two layers of crud
    b. Do not disturb the solids
7. Add 2/3 volume phenol/chloroform 1:1 to each tube
8. Mix and centrifuge at max speed for 10 minutes
9. Pipette the supernatant and put it into newly labeled tubes
    a. Do not disturb the bottom layer
10. Add 2/3 volume of 24:1 chloroform : iso-amyl alcohol to each tube
11. Mix by inversion and centrifuge at max speed for 10 minutes
12. Pipette the supernatant into a new tube
13. Repeat steps 12 through 15
14. Add equal volume of cold isopropanol
    a. Can place the tube at -20°C for 20 minutes
15. Spin at max speed for 10 minutes

16. Decant the liquid and wash the pellet with 70% Ethanol
17. Spin at 13,000 rpm for 5 minutes
18. Decant the liquid and dry the pellet
    a. Residual ethanol is destroys the integrity of the DNA
19. Add 75 µL of TE to re-suspend DNA

Following extraction, the re-suspended DNA was analyzed using a spectrophotometer to assess its purity and concentration. Once quantified, the DNA was diluted to the desired concentration of 20-50 nanograms per reaction volume for PCR.

PCR

Polymerase chain reaction was performed using a 10 µL reaction volume. 100 reactions were performed per plate and a hot start Taq polymerase was used. DNA was diluted to 20-30 mg/µL and primers were diluted to 10 µM for PCR. The formula for the PCR reaction is listed in Table 2.

Table 2: PCR formula for each reaction.

	1 Reaction 10 µL
5x GoTaq Buffer	2 µL
25 mM $MgCl_2$	2 µL
10 mM NTP	0.2 µL
Primers	0.2 µL (forward)
	0.2 µL (reverse)
Hot Start GoTaq	0.05 µL
ddH_2O	5 µL
Template DNA	1 µL

The settings for the thermocycler are listed in Table 3.

Table 3: Temperatures and time settings for Thermocycler for PCR

	Time	**Temperature**
	4 minutes	94º C
35 Cycles	30 seconds	94º C
	45 seconds	60º C
	45 seconds	72º C
	7 minutes	72º C

Analysis

Microsatellite markers were screened for amplification using agarose gel electrophoresis. Capillary electrophoresis data from amplifiable SSR markers screened for polymorphisms and for fingerprinting on the 600 R-14 samples were analyzed using GeneMapper software (ABI).

Results:

To date, DNA has been extracted from 600 samples of the R-14 family of black cherry tree. So far, 85 primers have been tested. Primers used in this study came from microsatellite markers obtained from a cDNA library previously established by the Carlson Laboratory (unpublished data). Of the 85 primers tested, 66 have shown amplification and two have been found to be polymorphic. The two markers shown to be polymorphic in black cherry are marker 654 and marker 131. After performing PCR on DNA samples with candidate primers, they were placed on an agarose gel and electrophoresis was performed. If distinct bands were visible, desired amplification had occurred. Figure 3 shows an example of an agarose gel after running on the electrophoresis apparatus, first number represent the marker; the second number represents the expected size of the amplified sequence.

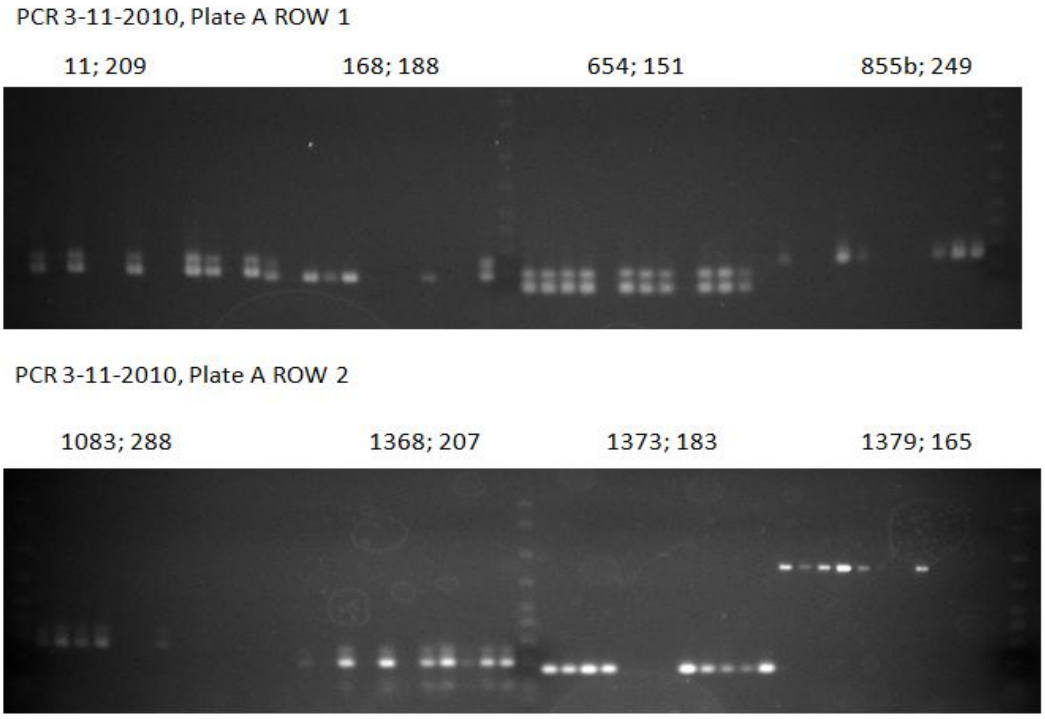

Figure 3: Agarose gel with 8 markers tested. All of which showed amplification.

If a primer was found to have amplified, another PCR plate was prepared with DNA samples and labeled primers. The plate was then placed in a bioanalyzer in an effort to determine polymorphisms. The capillary electrophoresis graphs in Figures 4 and 5 show that markers 654 and 131 both amplify and are polymorphic in black cherry.=, and this could be used for screening of the OP families,

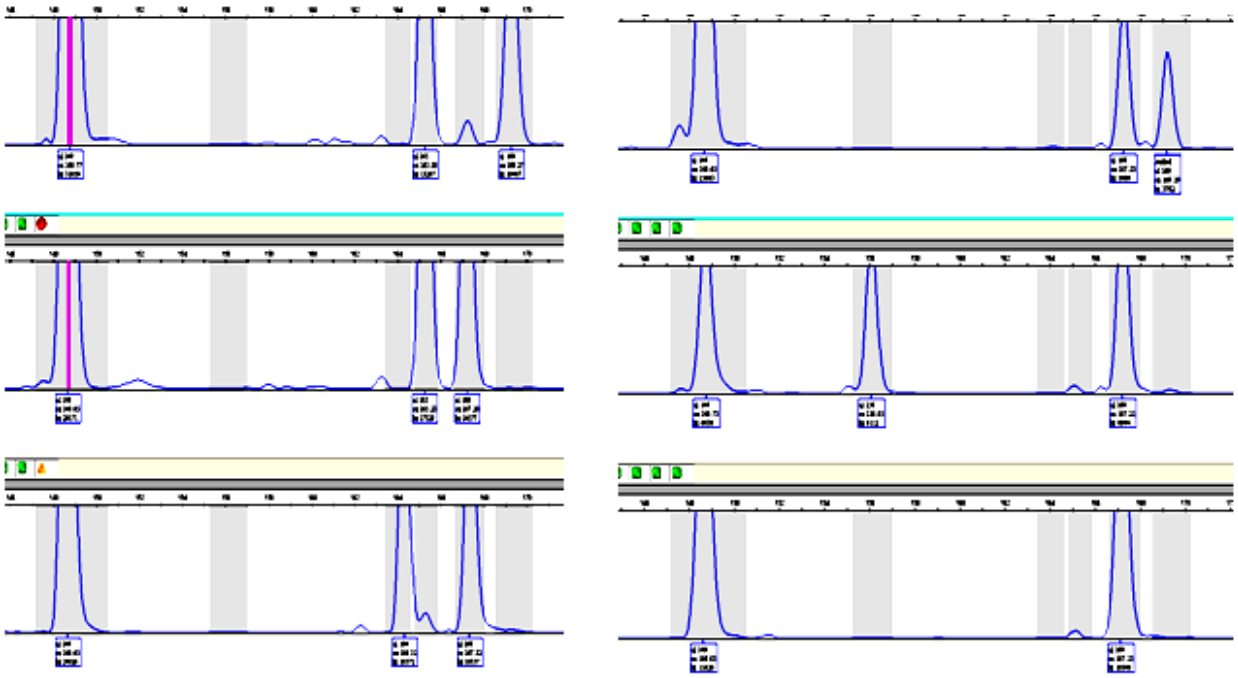

Figure 4: Graph showing heterozygosity for marker 654 in black cherry between 6 samples.

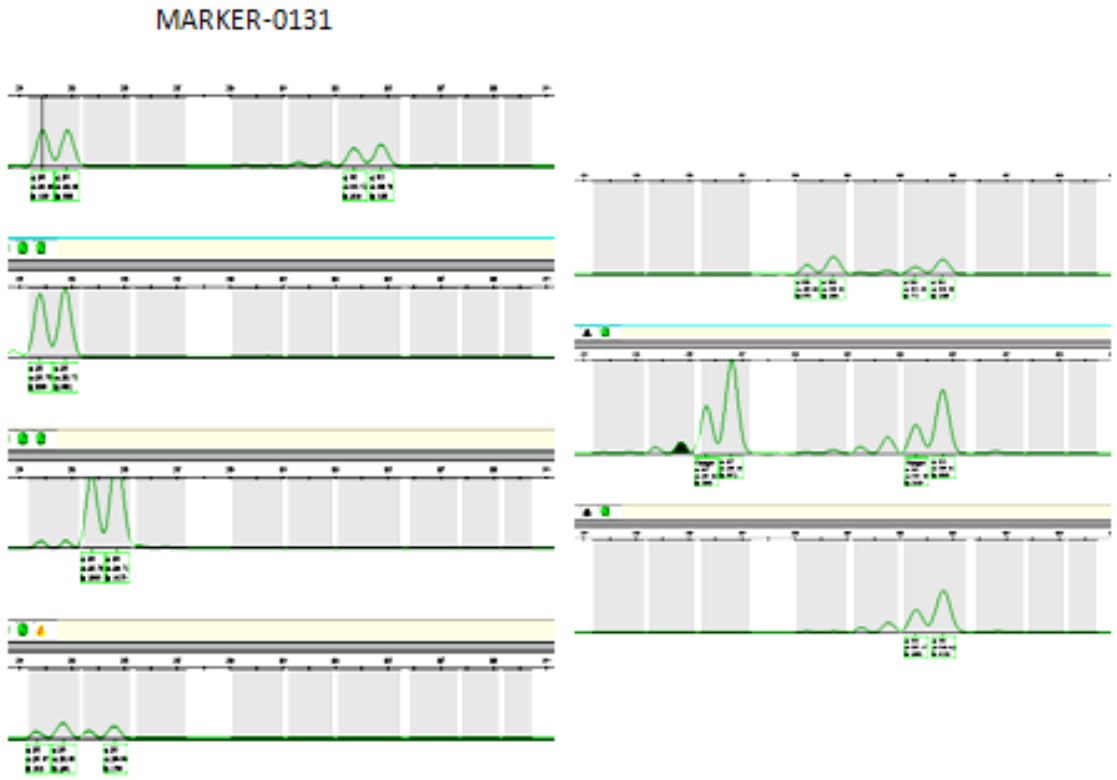

Figure 5: Graph showing heterozygosity for marker 131 in black cherry between 7 samples.

After establishing marker 131 was polymorphic in black cherry (Figures 5 above and 8 in Appendix A), I and another Honors student Peter Breen, screened 180 samples from the R-14 family with said marker in an attempt to develop a rough fingerprint. Of the 180 samples screened, 15% could not be identified due to low amplification and required further troubleshooting of the PCR process. 10% of the samples had an identical allelic profile with the mother tree and were thus clones. Of the remaining samples, 27% had identical profiles and thus had a preferred pollen source. Three other groups of identical samples existed (14 %, 10%, and 10%) which showed three additional pollen sources with similar rates of parentage. Additionally, three more pollen sources had similar rates of parentage with 5%, 6%, and 7% of the screened samples falling into one of these genotypes. The remaining 5% of the samples was composed of allelic profiles from 4 additional genotypes from 4 additional pollen sources.

Additional gels and graphs with data from the screening of other SSR primers can be found in Appendix A. In the appendix, Figures 5 through 10 show photographs of agarose gels. Figures 11 through 13 show graphs of allele distribution between samples.

Discussion:

The preliminary results discussed in the previous section where 180 samples were screened using marker 131 show 12 separate genotypes. Being able to cross reference results such as these will enable the identification of individuals instead of the identification of separate genotypes. Only having two SSRs that are polymorphic in black cherry out of the 85 candidate markers leaves us with no way of positively identifying full siblings. Once an additional 4 to 8

primers are found to both amplify and be polymorphic in *Prunus serotina*, all 600 samples from the R-14 will be screened for full siblings.

The full siblings and polymorphic markers found in this study will then be used to develop genetic linkage maps for the species. Once developed, the linkage maps will be used to identify quantitative trait loci (QTLs) related to ozone stress and resistance. The QTLs can then be used to help develop lines of black cherry trees more tolerant of ozone pollution and thus protect their place in the ecosystem and industry. With the use of linkage maps produced as a result of this study, as well as other genomic tools developed in the greater project, additional QTLs related to other important traits and characteristics in black cherry trees can be identified. For example, wood quality can be improved and a more pest resistant genotype can be developed. The black cherry tree is a valuable source of lumber in the eastern United States of America and an important invasive species being studied in Europe. *Prunus serotina* has been studied extensively because of its economic importance in the United States. By identifying QTLs related to response to ozone stress, we can select trees which are better capable of surviving with more resistance to ozone. This will help protect one of the more important resources in the United States.

In addition to benefits for the black cherry tree, the polymorphic SSRs identified in this study can be used in other species within the *Prunus* genus. Therefore syntenies between black cherry and other Prunus/Rosaceae spp. can be identified.

Acknowledgments:

I thank Teodora Best, John Carlson, Peter Breen, Ashley Yang, and Tyler Wagner for their help on this project. Financial support for this research was provided by USDA Plant Genome Program Grant #2008-35300-19234, and by the Schatz Center for Tree Molecular Genetics Undergraduate Research fund.

Works Cited

Aranzana, M. J., Pineda, A., & al., e. (2003). A set of simple-sequence repeat (SSR) markers covering the Prunus Genome. *Theoretical and Applied Genetics* , 819-825.

Ashmore, M. R. (2005). Assessing the future global impacts of ozone on vegetation. *Plant Cell and Environment* , 949–964. .

Beckman, M. (2003). *DNA profiling.* Retrieved April 13, 2010, from Genetics: http://www.encyclopedia.com/doc/1G2-3406500079.html

De Silva, H. N., Hall, A. J., Rikkerink, E., McNeilage, M. A., & Fraser, L. G. (2005). Estimation of allele frequencies in polyploids under certain patterns of inheritance. *Heredity* , 327-334.

Gilman, E. F., & Watson, D. G. (1994, October). *Prunus serotina Fact Sheet.* Retrieved March 21, 2010, from United States Department of Agriculture Forest Service: http://hort.ufl.edu/trees/PRUSERA.pdf

Karnosky, D. F. (1980). Changes in southern Wisconsin white pine stands related to air pollution sensitivity. *Effects of Air Pollutants on Mediterranean and Temperate Forest Ecosystems* (p. 43). General Technical Report PSW.

Lee, J. C., Steiner, K. C., Zhang, J. W., & Skelly, J. M. (2002). Heritability of ozone sensitivity in open-pollinated families of balck cherry (Prunus serotina Ehrh.). *Forest Science* , 111-117.

Lerner, K. L., & Lerner, B. (2004). *DNA Fingerprinting.* Retrieved March 22, 2010, from Gale Virtual Reference Library: http://find.galegroup.com.ezaccess.libraries.psu.edu/gvrl/infomark.do?&contentSet=EBKS&type=retrieve&tabID=T001&prodId=GVRL&docId=CX3418500716&source=gale&userGroupName=psucic&version=1.0

Liu, X., Anderson, J. M., & Pijut, P. M. (2009). Cloning and Characterization of Prunus serotina AGAMOUS, a Putative Flower Homeotic Gene. *Plant Molecular Biology* , 193-203.

Pairon, M., & Jacquemart, A.-L. (2008). Detection and Characterization of Genome-specific Microsatellite Markers in the Allotetraploid Prunus serotina. *Journal of the American Society of Horticultural Science* , 390-395.

Phartyal, S. S., Godefroid, S., & Koedam, N. (2009). Seed development and germination ecophysiology of the invasive tree Prunus serotina (Rosaceae) in a temperate forest in Western Europe. *Plant Ecology* , 285-294.

Polymerase Chain Reaction (PCR). (2009, October 23). Retrieved March 23, 2010, from National Human Genome Research Institute: http://www.genome.gov/10000207

Scholl, R. (2001). *Molecular Plant Genetics.* Retrieved March 22, 2010, from Gale Virtual Reference Library: http://find.galegroup.com.ezaccess.libraries.psu.edu/gvrl/infomark.do?&contentSet

=EBKS&type=retrieve&tabID=T001&prodId=GVRL&docId=CX3408000205&source=gale&userGroupName=psucic&version=1.0

Singh, L. (2009). Development and Exploitation of DNA Fingerprinting Technology in India. *Biotechnology Journal* , 335-341.

Wunsch, A., & Hormaza, J. I. (2004). Molecular evaluation of genetic diversity and S-allele composition of local Spanish sweet cherry (Prunus avium L.) cultivars. *Genetic Resources and Crop Evolution* , 635-641.

APPENDIX A

Additional Figures

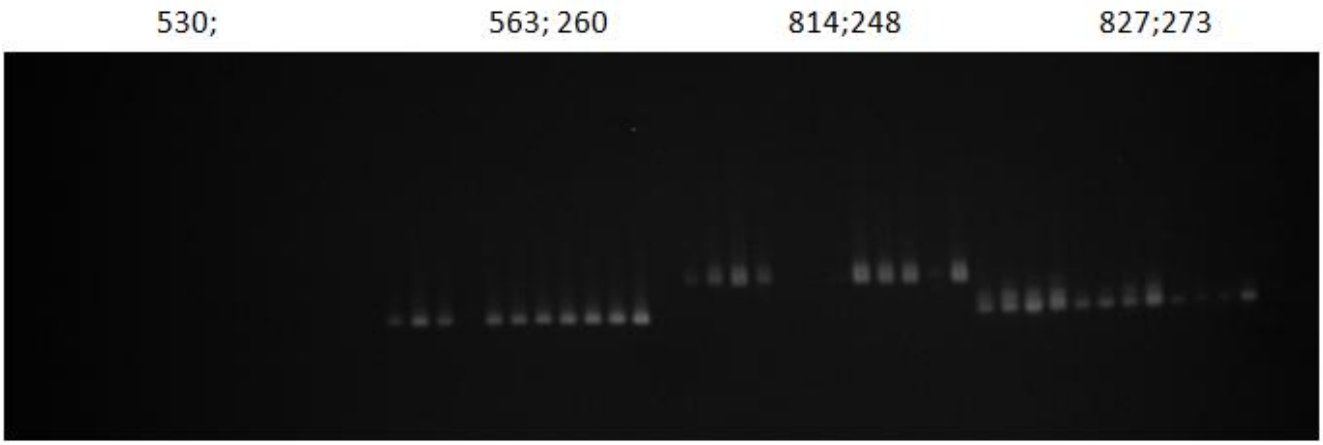

Figure 6: Agarose gel with 4 markers tested. Marker 530 showed no amplification.

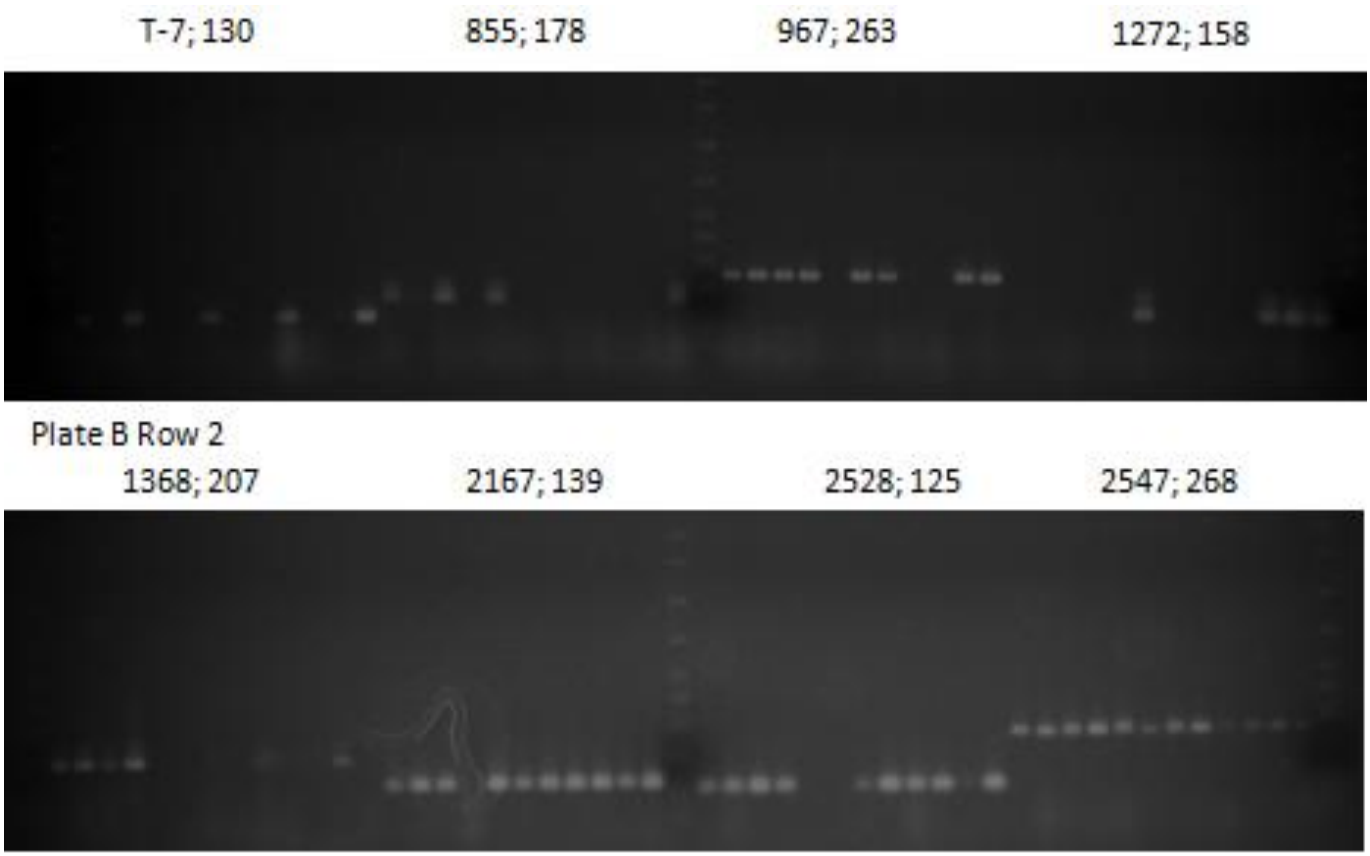

Figure 7: Agarose gel with 8 markers tested. All of which showed amplification.

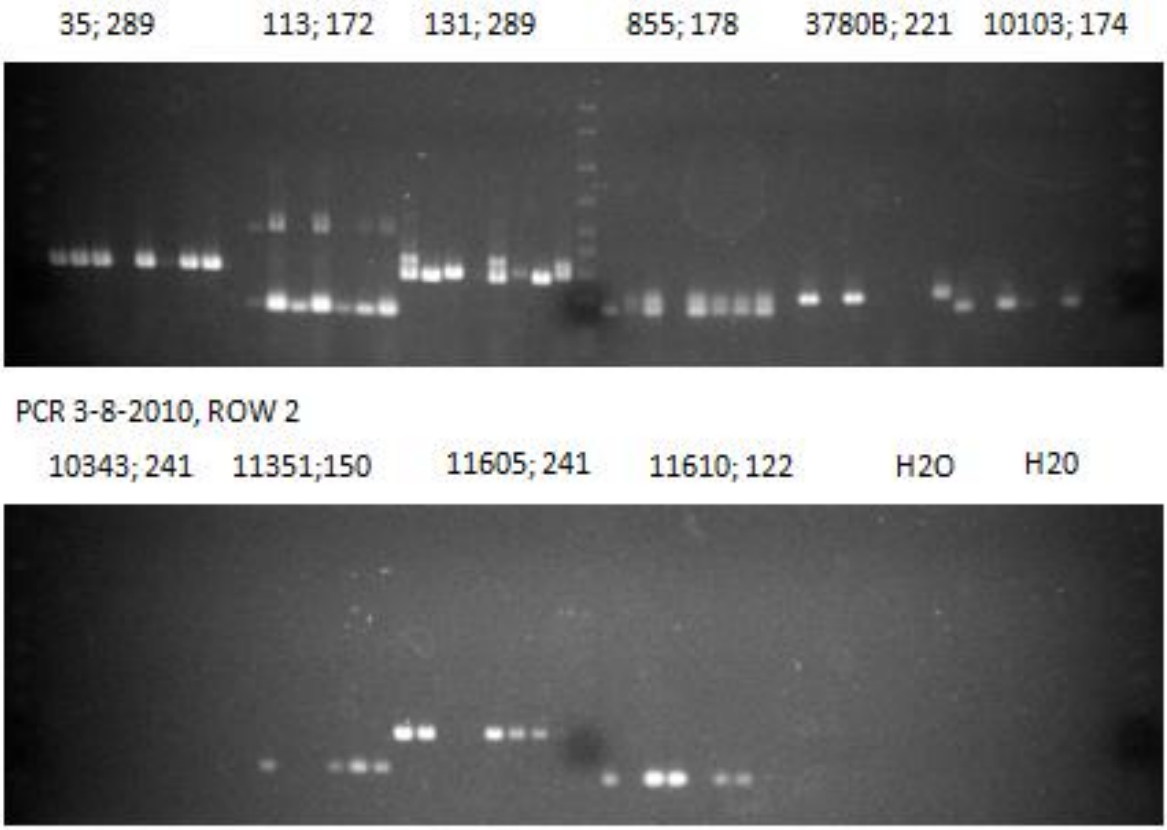

Figure 8: Agarose gel with 8 markers tested. Marker 241 showed no amplification.

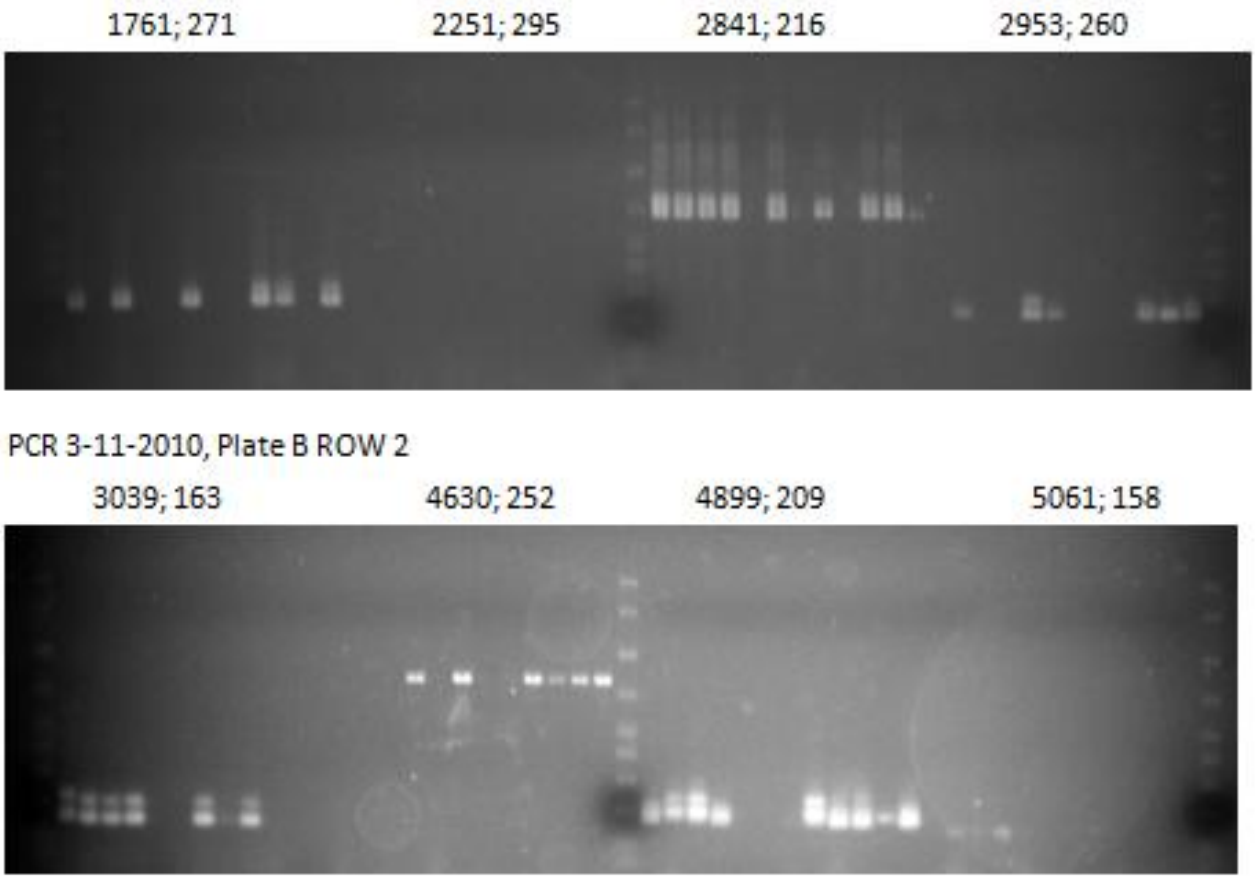

Figure 9: Agarose gel with 8 markers tested. Markers 295 and 158 showed no amplification.

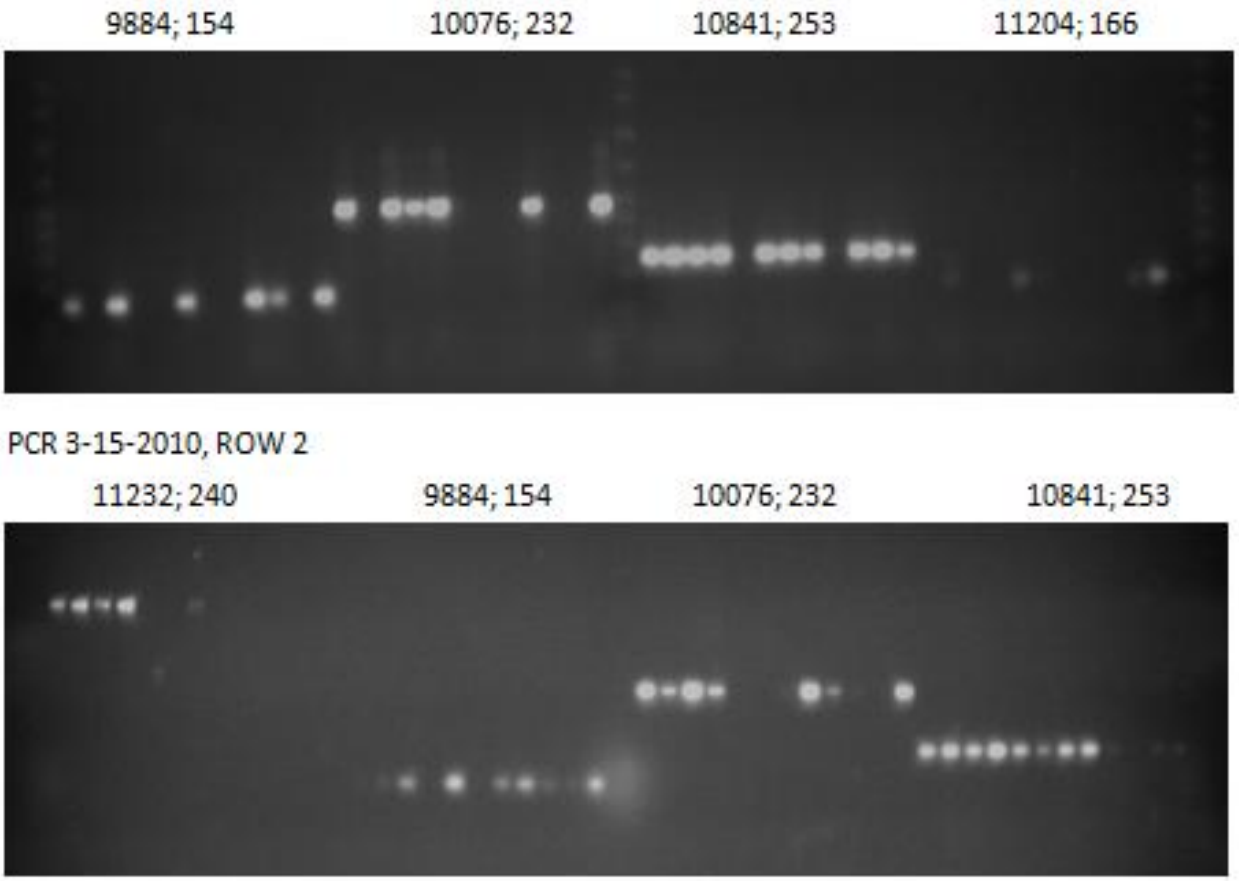

Figure 10: Agarose gel with 8 markers tested. All markers showed amplification.

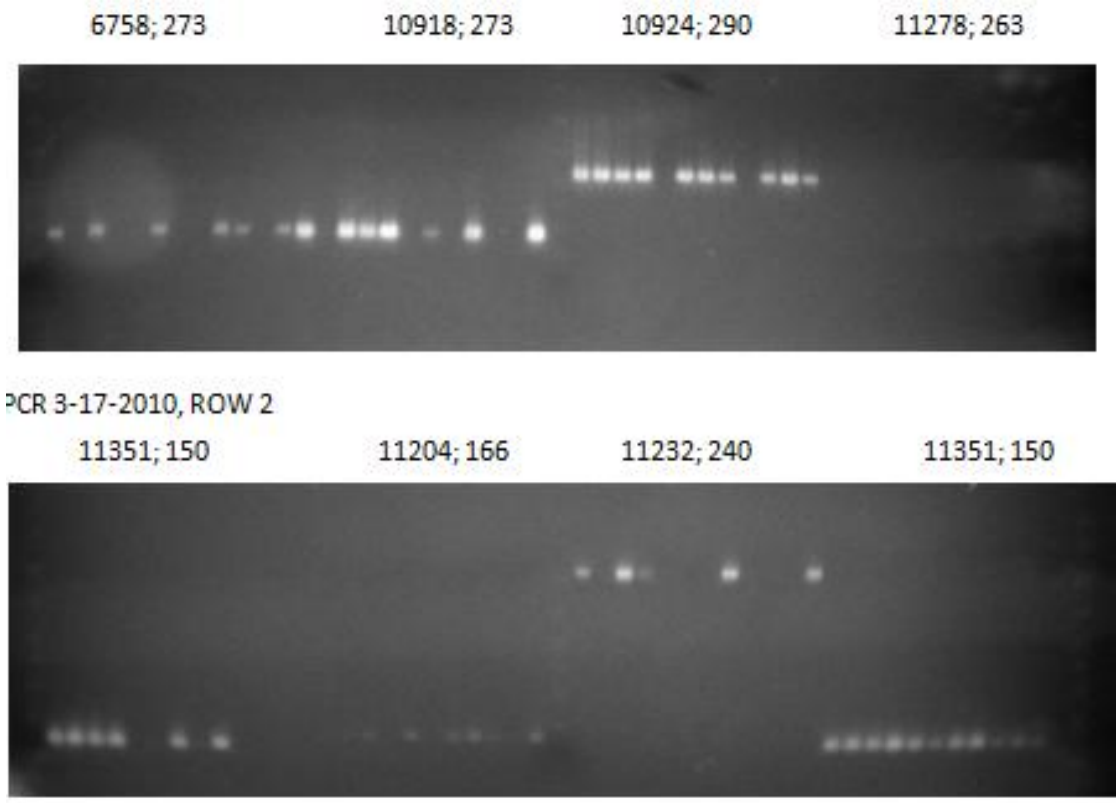

Figure 11: Agarose gel with 8 markers tested. Marker 263 showed no amplification.

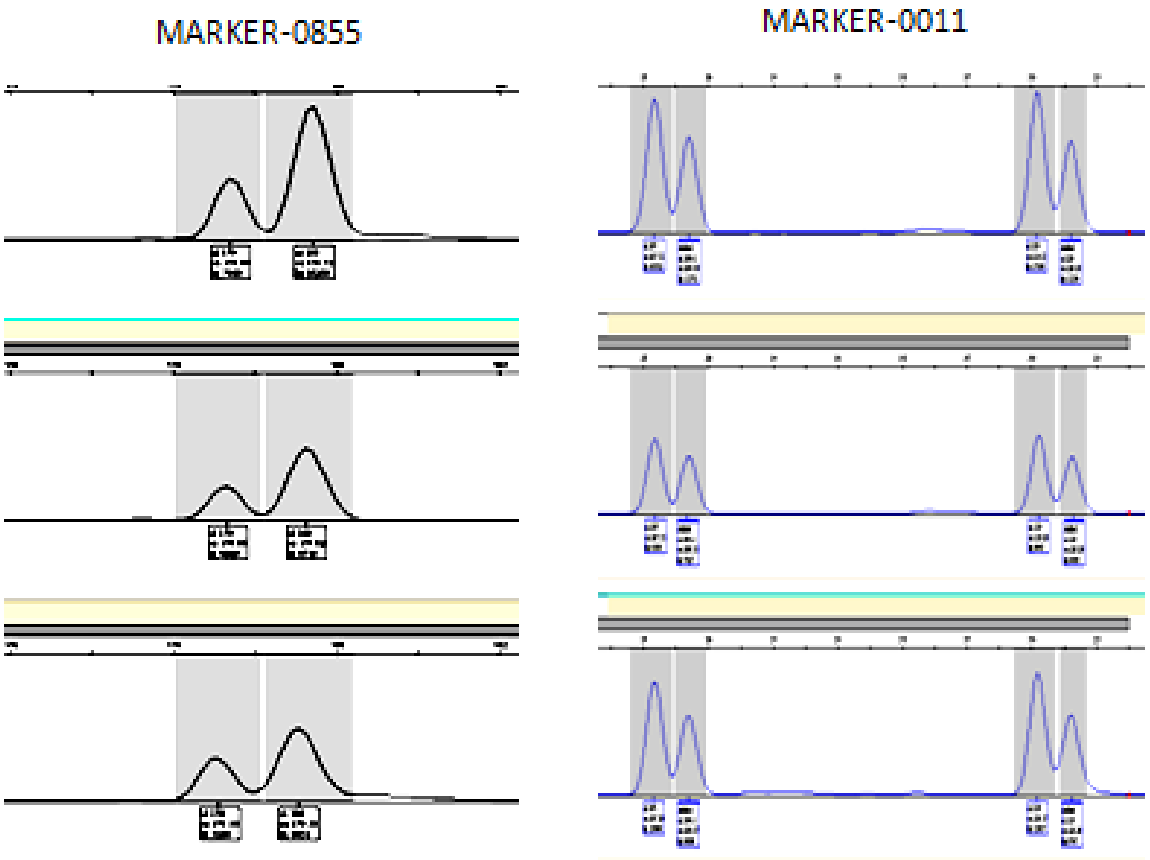

Figure 12: Graphs showing lack of heterozygosity between 3 samples for both marker 855 and marker 11.

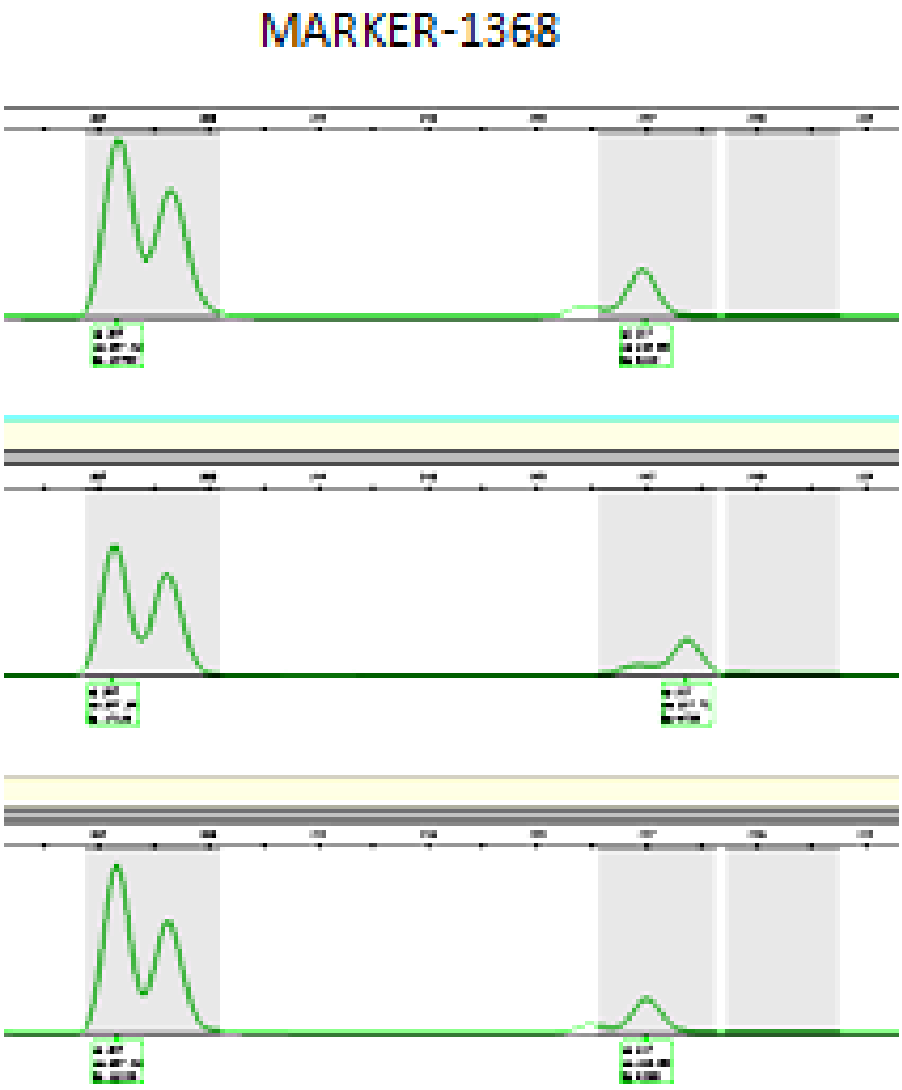

Figure 13: Graph showing lack of heterozygosity between 3 samples for marker 1368.

ACADEMIC VITA of Daniel J. Devine

Daniel J. Devine
4002 Evergreen Drive
Bethlehem, Pennsylvania, 18020
djd5073@@gmail.com

Education:

Bachelor of Science Degree in Biology, Penn State University, Spring 2010
Honors in Biology
Thesis Title: DNA Fingerprinting and Genotyping in *Prunus serotina*
Thesis Supervisor: John E. Carlson

Awards:

2007 Penn State Lehigh Valley Outstanding Biology Student
2007 Penn State Lehigh Valley Freshman with highest GPA: Science
2008 Penn State Berks Biology Academic Award
2008 Penn State Berks Organic Chemistry Academic Award
2008 Boscov's Scholarship
Dean's List
Phi Kappa Phi National Honor Society
Phi Beta Kappa

Presentations/Activities:

Volunteer Researcher at Caribbean Conservation Corporation; Costa Rica
Volunteer Researcher at *Asociación* ANAI; Costa Rica
Volunteer Researcher at La Selva Biological Station; Costa Rica
Volunteer Judge for Pennsylvania Junior Academy of Science
2010 Penn State Undergraduate Exhibition