Proceedings of The National Conference
On Undergraduate Research (NCUR) 2016
University of North Carolina Asheville
Asheville, North Carolina
April 7-9, 2016

Genetic diversity and population structure in the clonal plant *Trillium recurvatum*

Catherine K. Major
Department of Biology
University of Memphis
3774 Walker Ave
Memphis, TN 38152

Faculty Advisors: Dr. Jennifer Mandel and Dr. James Moore

Abstract

Trillium recurvatum is a long-living herbaceous perennial plant found in the central and eastern United States. It is currently threatened in Michigan and rare in Wisconsin. Threats to this plant include: forest management practices, land-use conversion, and habitat fragmentation. Pollen limitation also threatens this plant as the species is self-incompatible so one genetic individual must be pollinated by a different genetic individual. *T. recurvatum* is clonal, so pollination could be difficult if there are few different genetic individuals in a specific region. My research project aims to explore the genetic diversity of *T. recurvatum* in the University of Memphis Meeman Biological Field Station. This will provide the first population genetics for Shelby County populations *of T. recurvatum*. In March 2014, 220 *T. recurvatum* leaf samples were collected from a population at the University of Memphis Meeman Biological Field Station. The DNA was extracted from these leaves using a commercially available kit, Omega Biotek E.Z.D.N.A Plant DNA Kit, and stored at -80 degrees Celsius. Following that, PCR was performed on all 220 samples for the 10 microsatellite loci. Locus 2,3,4,7,9, and 10 were pooled and run on a capillary electrophoresis at the Molecular Research Core at The University of Tennessee Health Science Center where the instrument can detect one base pair differences. The alleles were recorded using the software package GeneMarker, and the data was analyzed for genetic and clonal diversity using the GenAlEx software. These results were correlated with the demographic data collected by Dr. James Moore and colleagues[8]. Based on the data analyzed, the genetic diversity of *T. recurvatum* was higher than expected. 108 unique genotypes were found from 182 plants sampled, which means there is 60% diversity.

Keywords: Demography, microsatellite, population genetics

1. Introduction

Trillium recurvatum Beck (common name: Prairie Trillium), Melanthiaceae family, is a long living perennial clonal plant found in the United States. It extends as far north as Michigan, while also extending down into Texas. The species is often found in shady woodland areas with rich, well-drained soil. Flowering typically begins anywhere from late February to March in partial to full shade. Currently, *T. recurvatum* is listed by the U.S. federal government as a threatened species in Michigan (Fig. 1). It is also considered rare in Wisconsin, Alabama, North Carolina, Ohio, Texas, Louisiana, and Oklahoma. (Fig. 1). In these states, *T. recurvatum*, as well as other forest understory species, is threatened by climate change, forest management practices, land-use conversion, and habitat fragmentation[4]. This species is an understory herb, which means that it is present on the bottom layer of the forest, under the main canopy. Understory herbs have some of the largest species diversity and thus are important in

supporting the structure of forests[3]. Additionally, it has been shown that higher genetic diversity among plant populations supports a high genetic diversity among arthropod communities[11].

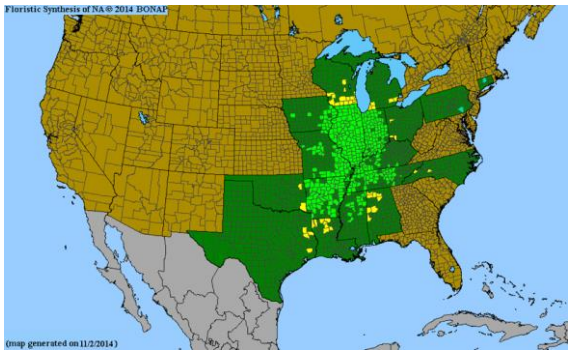

Figure 1: BONAP map of *Trillium recurvatum*.

Brown indicates absence, dark green indicates the state in which it is found, light green indicates the specific county, yellow demonstrates rarity, and the light blue indicates that *T. recurvatum* is adventive in that specific county[5].
Kartesz, J.T., The Biota of North America Program (BONAP). 2015. North American Plant Atlas. (http://bonap.net/napa). Chapel Hill, N.C. [maps generated from Kartesz, J.T. 2015. Floristic Synthesis of North America, Version 1.0. Biota of North America Program (BONAP). (in press)].

Prairie Trillium is characterized by maroon flowers and mottled leaves. The flowers are pollinated by insects and produce a capsule-like berry. Ants feed on the eliasomes, which are fleshy structures attached to the seed. Once the seeds are discarded, it takes approximately two springs for a juvenile to appear and a minimum of four years for flowers to develop[9]. The stems originate from underground rhizomes that slowly spread through the soil. The underground rhizomes then give rise to multiple *T. recurvatum* individuals making it a clonal plant.
This species has three life history stages: juvenile, flowering, and non-flowering. The juvenile stage is characterized by a single leaf, the flowering has three mottled leaves with a maroon flower in the center, and the non-flowering has three mottled leaves (Fig. 2). Flowering plants can revert to non-flowering and vice versa; however, the mechanism by which this occurs is unknown.

Figure 2: Life stages of *Trillium recurvatum*. Red circle represents juvenile, blue circle represents non-flowering adult, and white circle represents flowering adult.

An ongoing population demography study has been carried out on a population of *T. recurvatum* at the University of Memphis Meeman Biological Field Station (referred to hereafter as the Meeman Trillium Population) since 1996. From this study, it was discovered that non-flowering ramets had the greatest population growth rate and flowering plants comprised 18% of the population on average[8]. Although *T. recurvatum* is not rare in Tennessee, understanding patterns of clonal genetic diversity in the Meeman Trillium Population coupled with the long-term

demographic data offer the possibility of providing unique insight into population genetic dynamics in a clonal species. Moreover, populations of *T. recurvatum* may display pollen limitation; because the species is self-incompatible, one genetic individual must be pollinated by a different genetic individual[10]. If *T. recurvatum* demonstrates a high degree of clonality, insect pollinators flying from plant to plant might actually transfer pollen to the same genetic individual where fertilization would be unsuccessful. An understanding of pollen limitation and clonal genetic structuring is also important for extrapolating to populations/sites where the species is rare.

Clonal plants are abundant across the planet, often dominating in forest understory, aquatic environments, and extreme conditions[2]. Without the need for sexual reproduction, which can be a risky strategy in some environments, clonal plants can propagate by rhizomes, tillers, surface runners, etc.[2]. Each group of genetically identical individuals is called a genet; whereas, an individual stem is known as a ramet. Traditionally, "stem counts" were used to measure population numbers; however, those can overestimate the level of diversity of clonal plants[6]. Genetic markers, which are non-invasive, can be used to better estimate the number of genetic individuals and calculate levels of genetic diversity.

Clonality can, sometimes, be associated with a low level of genetic diversity -- an entire plot could be one genetic individual, connected underground by rhizomes. The ramets possess identical phenotypic and genotypic characteristics; this could be problematic if the specific genotype were to be adversely affected by a predator or by the environment. The focus of this research is on the genetic diversity of *T. recurvatum* at the University of Memphis Meeman Biological Field Station. Population genetic markers were used to differentiate genetic individuals, determine the number of genotypes in populations, and measure population levels of genetic diversity. We hypothesized that there would be low levels of genetic diversity in the *T. recurvatum* populations and that the population would consist of just a few distinct genetic individuals i.e., high clonality.

In this research, simple sequence repeats (SSR) were used as genetic markers to identify genetic diversity and clonal structure[6]. The SSRs, also known as microsatellites, are repeated segments of DNA where each repeat unit is usually 2-5 base pairs long (for example AT, GCC, or CAGT). The microsatellite markers used in this study are the first microsatellites to be developed for this species. Because of the high mutation rate of microsatellites, individuals will vary by the number of repeat units they contain[1]. One can isolate and amplify these repeated regions of the genome via the polymerase chain reaction (PCR) and subject the resulting product to gel electrophoresis. When different individuals differ by one or more repeat units, the PCR products of different sizes will migrate at different rates and be distinguishable from one another.

The microsatellite markers that were used in the study were developed using a microsatellite enrichment method and subsequent next-generation sequencing provided by the microsatellite discovery services at the Evolutionary Genetics Core Facility (EGCF) at Cornell University. The Core digested two pooled samples of whole genomic *T. recurvatum* DNA with several restriction enzymes and ligated appropriate adaptors for next-generation sequencing on the Illumina platform. Following this, they performed microsatellite enrichment by hybridization to 3'-biotinylated oligonucleotide repeat probes (for example, ATATATAT). These hybridized fragments were captured by streptavidin-coated magnetic beads and a magnet was used to wash away the unwanted portion of the *T. recurvatum* genome. After this, the capture fragments were sequenced, and the EGCF used a bioinformatic pipeline to search the sequences for repeats. The EGCF then designed primers for PCR amplification of the microsatellites, and ten of these regions, or loci, were chosen for this project.

2. Materials and Methods

2.1 Sampling method

Leaf samples of *T. recurvatum* were collected from a population at the University of Memphis Meeman Biological Field Station in March 2014 (Fig. 3). Individuals were sampled in a five-by-five array of 2x2 meter quadrats. Plants were sampled to ensure the greatest spatial variation, i.e., from as many spatially distinct clusters as possible within the 2x2m quadrat. The leaves were stored in a refrigerator prior to DNA extraction. To extract the DNA, the leaf was placed in a 2.0 ml tube along with three metal beads. The tube was placed in the SPEX Geno grinder 2000 at 500 rpm until the leaf was ground. An OMEGA bio-tek E.N.Z.A. SQ Plant DNA kit was used to extract the DNA. The DNA was then stored at -80 degrees Celsius until ready for use.

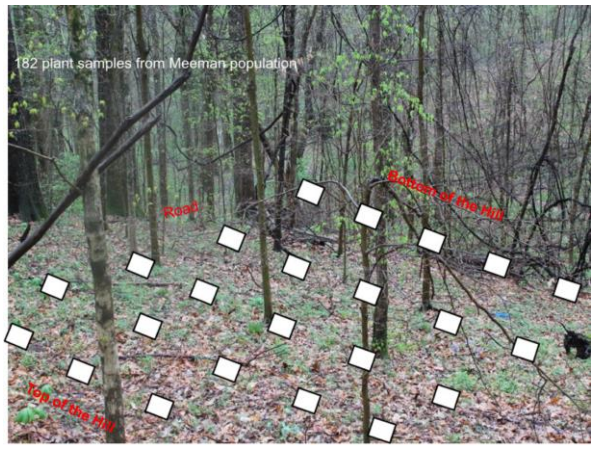

Figure 3: Collection site for *T. recurvatum* at University of Memphis Meeman Biological Field Station. White box indicates one 2x2 meter quadrat.

2.2 Sample Assessment

Each individual was assessed for purity and measured for nucleic acid concentration by testing it on a NanoDrop Spectrophotometer. One µl of DNA was used in this process. Following the results of the Spectrophotometer, the DNA was diluted by a factor of 10 based on the concentration of nucleic acids. The individuals were split up into 3 master plates. Master Plates 1 and 2 contain 95 individuals, each with a water control. Master Plate 3 contains 26 individuals and a water control.

2.3 Experimental Design

Polymerase Chain Reaction (PCR) was used to amplify the extracted DNA. Loci 1-10 were run on sample E11 and T4, from Master Plate 3. A master mix composed of 9.55 µl deionized water, 1.5 µl buffer, 0.35 µl $MgCl_2$, 0.2 µl dNTPs, 0.35 µl forward primer at 5 µM, 0.35 µl of reverse primer at 20 µM, 0.70 µl M13 (FAM, NED, or VIC), and 1.0 µl *Taq* polymerase in a micro-centrifuge tube. Primers with and without extra base pairs were used in this experiment (Table 1). The reverse primer had extra base pairs (GTTTCTT) so that it would further stabilize the primer annealing; however, the primer annealing was unsuccessful so the primers without extra base pairs were used. Each value was multiplied by 11 in order to provide enough solution for 10 individuals. 1.0 µl of DNA from each individual was mixed with 14 µl of the master mixer and run on the thermocycler using the program Td_50_1min. The temperature cycles for the program were as follows: 95º C for 3 min, [94ºC for 30 sec, 60ºC for 30 sec, 72ºC for 1 min] x 9, [94ºC for 30 sec, 50ºC for 30 sec, 72ºC for 1 min] x 29, 72ºC for 10 min, and 4ºC for 15 min. This process was repeated with samples A9 and T3, from Master Plates 1 and 3 respectively.

Following the results of that trial, Loci 2, 3, 4, 7, and 9 were selected to be amplified on all of the *T. recurvatum* samples. Loci 2 and 3 were amplified using the same protocol and thermocycler program as used when testing Loci 1-10. Locus 9 was slightly modified by using .27 µl $MgCl_2$ in order to decrease double banding. Additionally, .20 µl of forward primer, reverse primer, and M13 were used instead of the .35 µl used previously. The primers used in Locus 9 also contained extra base pairs. The decrease in solution yielded a 13 µl reaction. Program Td_55_1min was used to amplify Locus 9 on the thermocycler. Individuals from Master Plates 2 and 3 were amplified at Loci 2, 3, and 9 using the same protocol and thermocycler program used for Master Plate 1.

Locus 4 was amplified using a master mix of 9.80 µl of deionized water, 1.5 µl of buffer, .30 µl of $MgCl_2$, .20 µl dNTPs, .20 µl of forward primer at 5 µM, .20 µl of reverse primers with extra base pairs at 20 µM, .20 M13 (NED) at 10 µM, and .60 µl of *Taq* polymerase. 1 µl of DNA was added to 13 µl of the master mix. Individuals were run on the program Td_57_45sec. Instead of lasting 1 minute, the 94 º C, 67-57 º C, and 72 º C cycle lasted 45 seconds. This protocol and program was used for all individuals from Master Plates 1, 2, and 3. Locus 7 was amplified using a master mix of 9.35 µl of deionized water, 1.5 µl of buffer, .35 µl of $MgCl_2$, .20 µl of dNTPs, .20 µl of forward primer at 5 µM, .20 µl of reverse primer with extra bases at 20 µM, .20 µl of M13 (VIC) at 10 µM, and 1.0 µl of

Taq polymerase. One µl of DNA was added to 13 µl of the master mix. Individuals were on the thermocycler program Td_50_1min. This protocol and program was used for all individuals from Master Plates 1, 2, and 3.

Locus 6 was amplified using a master mix of 9.85 µl of deionized water, 1.5 µl of buffer, .35 µl of $MgCl_2$, .20 µl dNTPs, .20 µl of forward primer at 5 µM, .20 µl of reverse primers at 20 µM, .20 µl M13 (NED), and .50 µl of *Taq* polymerase. 1.0 µl of DNA was added to 13 µl of the master mix. Individuals were run on the program Td_57_45sec. This protocol was used for Master Plates 1, 2, and 3. Table 2 indicates the fluorophore (M13) used with each locus.

Table 1: Forward and Reverse Primer Sequences for each amplified locus.

	Forward Primer Sequence	**Reverse Primer Sequence**
Locus 2	CACGACGTTGTAAAACGACTAA ACAAACATCGCCAGGCC	TCATTCTGACTGAGGGCGAG
Locus 3	CACGACGTTGTAAAACGACAAG GGTTGTGAGAGGGAAGG	GTCGCCCGCATCAAATAGAG
Locus 4	CACGACGTTGTAAAACGACAAC TGCACAAAGGTTTCCGC	AGCGATCAAATTCCAGGAACC
Locus 6	CACGACGTTGTAAAACGACTTC GCTGCTTCCCTCGTATC	AAGGACATGAGCGTTTGTGG
Locus 7	CACGACGTTGTAAAACGACCCT TACCTCGAGAAGCTTTGG	ATCGCTCGAACCCATACTCC
Locus 9	CACGACGTTGTAAAACGACTAA TGCCAAGCCCTCAGAAG	TCGCCGTTGTTTCGAGTTC

Table 2: Fluorophore used with each locus.

Locus	**Fluorophore**
Locus 2	VIC
Locus 3	FAM
Locus 4	NED
Locus 6	NED
Locus 7	VIC
Locus 9	NED

Gel electrophoresis was used to ensure that the PCR was successful and to estimate fragment length. The PCR products were mixed with loading dye and loaded into a 1% agarose gel along with TAEx buffer. The first well of each gel contained a ladder, which is a standard by which fragment lengths are measured. The last well of the gel contained a water control, unless otherwise specified. The gel electrophoresis was run at 80 volts and the results were visualized by placing the gel in a UV light box (Fig. 4). A dilution plate was created by adding 3 µl from each plate (Loci 2,3, and 9) and then 21 µl of water. One µl of the dilution plate and 12 µl of formamide ladder were combined to make a run plate. It was heated at 95ºC for 5 minutes on the thermocycler in order to denature the DNA so that they could be the sequenced. The purpose of the formamide was to keep the strands of DNA separated after being heated. The run plate was then taken to the University of Tennessee Molecular Research Core for fragment analysis using an ABI 3130XL Capillary Sequencer. GeneMarker 2.6.3 was used to score the fragments on the computer.

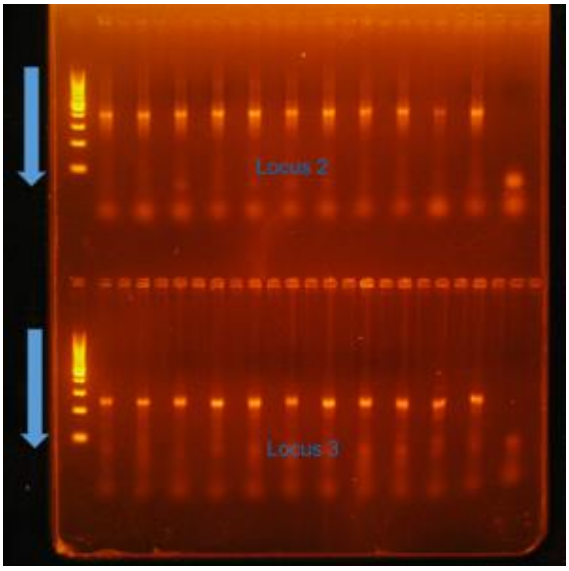

Figure 4: Gel electrophoresis for Locus 2 and 3. Individuals H1-H12 from Master Plate 1. Blue arrows indicate downward movement of ladder. Fragment size decreases with downward movement. Gel was run to ensure successful completion of Polymerase Chain Reaction.

A dilution plate was made for Locus 6 by mixing 3 µl of PCR product with 27 µl of water. One µl of the dilution plate was mixed with 9 µl of formamide ladder, heated at 95 ºC for five minutes and then sent to the University of Tennessee Molecular Research Core for fragment analysis using an ABI 3130XL Capillary Sequencer. Due to the faintness of the bands, 6 µl of Locus 4 and 9 µl of Locus 7 were added to the dilution plate. One µl of this dilution plate was mixed with 12 µl of formamide ladder before being sent off. GeneMarker 2.6.3 was used to score the fragments on the computer. Each allele was manually recorded based on fragment length (Figs. 5-7). Once the fragments were scored, patterns were identified and labeled for each individual. Labels were recorded as A-Z, and then AA, BB, etc. These patterns were recorded in an Excel spreadsheet and used to name the genotype of each individual (loci were concatenated to record a multi-locus genotype). For example, individuals A1-6NF and A1-8J both have the genotype FCAT (Table 3).

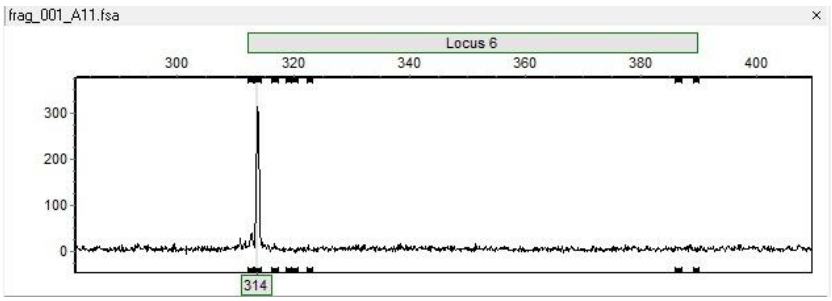

Figure 5: Homozygous 314 for Locus 6. 314 indicates number of base pairs.

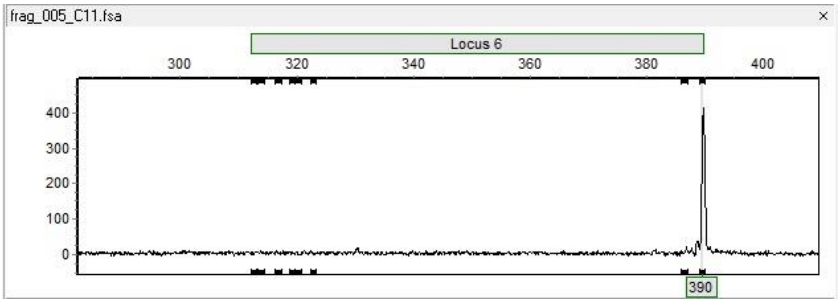

Figure 6: Homozygous 390 for Locus 6. 390 indicates number of base pairs.

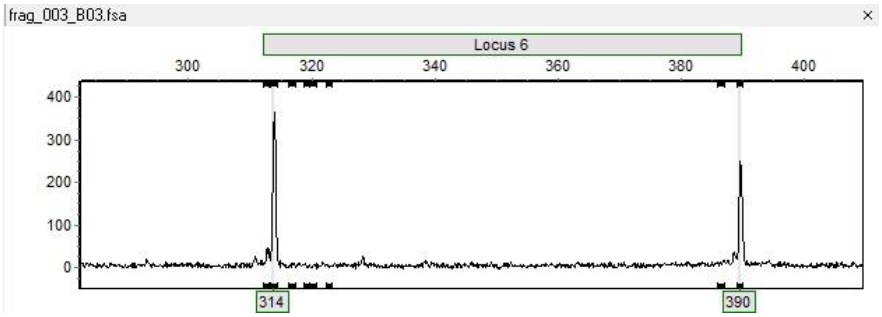

Figure 7: Heterozygous 314, 390 for Locus 6.

3. Results

GS+, a comprehensive geostatics program manufactured by gammadesign, was used to create the heat map of genetic diversity (Fig. 8). The data were placed into X, Y, Z form with X being the transect position running up the hill parallel to the road and Y being the transect position moving away from the road and Z being the number of unique genotypes counted in each quadrat. The road is labeled in Figure 3 for reference. A semivariogram was used to determine spatial autocorrelation of the samples and then point krigging was used to interpolate the map. Power Point was used to created the figures featuring Plot A2, D4, and D5 (Figs. 9-11). Each life stage was assigned a shape: square for non-flowering, circle for flowering, and triangle for juvenile. Clusters were created to mimic the position of the samples within the plots. Each genotype was assigned a different color in order to showcase the diversity found in the plot.

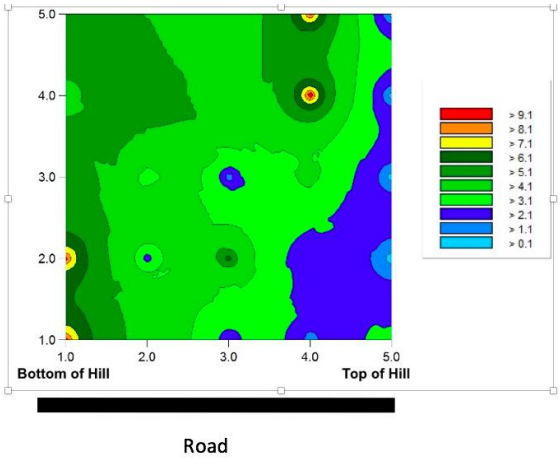

Figure 8: Heat map of genetic diversity for all sampled individuals.

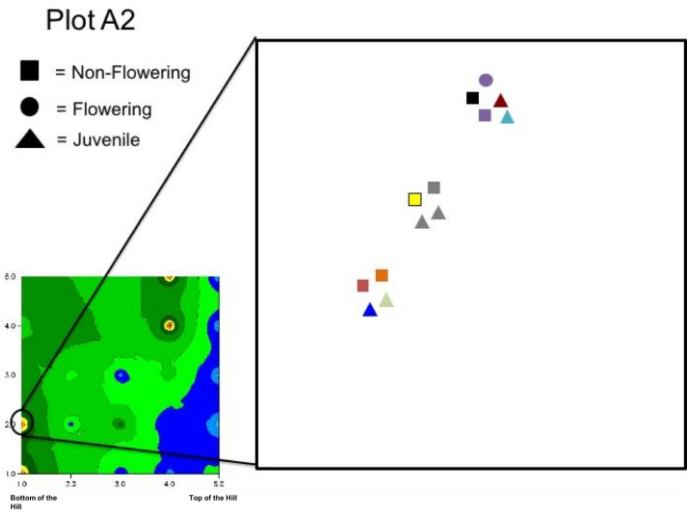

Figure 9: Genetic diversity of Plot A2. Colors indicate different genotypes.

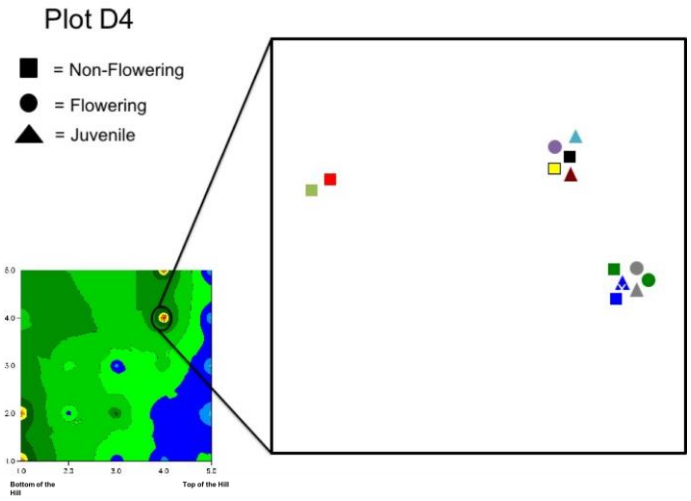

Figure 10: Genetic diversity of Plot D4. Colors indicate different genotypes.

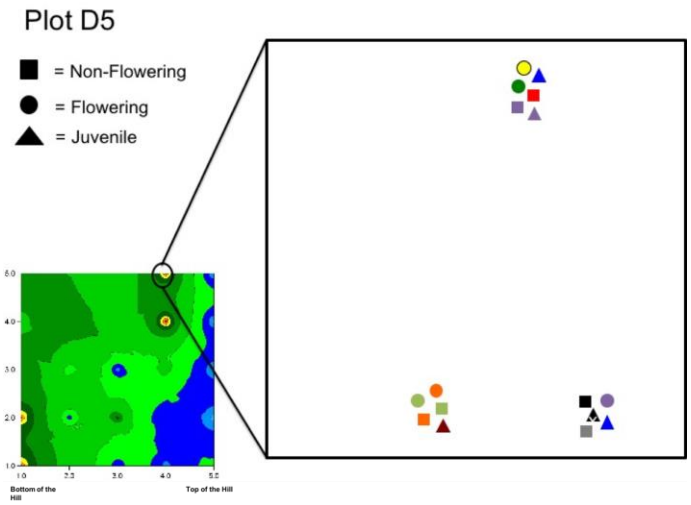

Figure 11: Genetic diversity of Plot D5. Colors indicate different genotypes.

4. Discussion

Locus 2 was determined to be a monomorphic locus following analysis by GeneMarker 2.6.3. A monomorphic locus is common to all individuals in the population and thus, does not provide pertinent information as to levels of genetic diversity. Locus 3 amplified well in the majority of individuals and showed considerable diversity. In order to determine the number of genotypes, allele patterns were recorded and assigned an arbitrary letter code. This was done for each locus, with the exception of Loci 2 and 9. Locus 2 was excluded due to its monomorphic properties and Locus 9 was excluded due to indiscernible banding patterns. From this, 108 unique genotypes were found in 182 individuals, which corresponds to roughly 59% genotypic diversity in the population. Most clusters of individuals were found to have multiple genotypes. These results were somewhat surprising and suggest a substantial amount of sexual reproduction as compared to vegetative propagation occurring in the Meeman *Trillium* Population. The findings here also suggest that pollination within clusters has a high likelihood of generating a successful fertilization since many individuals within clusters were actually different genets.

With regard to the spatial structure of genetic diversity in the Meeman *Trillium* Population, this study provided evidence of a more random genetic spatial structuring as opposed to a clumped distribution which is often observed in clonal plants[6]. The bottom of the hill appears to have more genotypic diversity than the top, which could be due to soil erosion pushing individuals towards the bottom of the hill. The soils in this area are a loess silt and demonstrate high erodability[7].

Future plans include amplifying more loci for the existing samples and then collecting more samples in the spring from the University of Memphis Meeman Biological Field Station and the Botanic Garden. Additionally, obtaining permits to collect from states in which *T. recurvatum* is rare could provide vital insight on genetic differences between the rare populations and the Meeman population. Greenhouse studies could also be done to see how *T. recurvatum* responds to changes in temperature, humidity, light, etc. The results from these future studies can be extrapolated to rare *Trillium* species as well as other rare, clonal species. As climate change, deforestation, and habitat fragmentation progress, it will be important to know how this species will respond to these significant changes.

5. Acknowledgements

The author wishes to thank the Helen Hardin Honors Program for funding this research project. She also expresses appreciation to her faculty research advisors Dr. Jennifer Mandel and Dr. James Moore along with the Mandel Lab for their guidance and support throughout this project.

6. References

1. Brooker, R.J. "Genetics: Analysis and Principles." 2d. ed. Boston.: Mcgraw Hill, (1999). Print.

2. Cook, R.E. "Clonal Plant Populations: A Knowledge of Clonal Structure Can Affect the Interpretation of Data in a Broad Range of Ecological and Evolutionary Studies." *American Scientist* 71.3 (1983). 244–253.

3. Gilliam, F.S. "The Ecological Significance of the Herbaceous Layer in Temperate Forest Ecosystems." *Bioscience* 57.10 (2007). 845-858.

4. Jules, E.S. "Habitat Fragmentation and Demographic Change for a Common Plant: Trillium in Old-growth Forest." *Ecology* 79.5 (1998). 1645–1656.

5. Kartesz, J.T. The Biota of North America Program (BONAP). 2015. North American Plant Atlas. (http://bonap.net/napa). Chapel Hill, N.C. [maps generated from Kartesz, J.T. 2015. Floristic Synthesis of North America, Version 1.0. Biota of North America Program (BONAP). (in press)].

6. Mandel, J.R. "Clonal diversity, spatial dynamics, and small genetic population size in the rare sunflower, Helianthus verticillatus." *Conservation Genetics* 11.5 (2010). 2055-2059.

7. McCarthy, K.P. "An analysis of gully development in Meeman-Shelby Forest State Park, Tennessee." Master's Thesis, The University of Memphis, Memphis, TN. 1990.

8. Moore, J.E., Franklin S.B., Wein, G., Collins, B.S. "Long-term population demography of Trillium recurvatum on loess bluffs in western Tennessee, USA." *AoB PLANTS* 2012: pls015; doi:10.1093/aobpla/pls015 2012.

9. O'Connor, R.P. Special Plant Abstract for Trillium recurvatum (prairie trillium). Lansing, MI: Michigan Natural Features Inventory. 2007. 3 pp.

10. Sawyer, N.W. "Reproductive Ecology of Trillium Recurvatum (Trilliaceae) in Wisconsin." *The American Midland Naturalist* 175.1 (2009). 146-160.

11. Williams, R.S., Avakian M.A. "Colonization of *Solidago altissima* by the Specialist Aphid *Uroleucon nigrotuberculatum*: Effects of Genetic Identity and Leaf Chemistry." *Journal of Chemical Ecology* 2015: 41:129-139; doi:10.1007/s10886-015-0546-1