To the Graduate Council:

I am submitting herewith a dissertation written by Denita Hadziabdic entitled "Evaluation of genetic diversity of flowering dogwood (*Cornus florida* L.) in the eastern United States using microsatellites." I have examined the final electronic copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Plants, Soils, and Insects.

Robert N. Trigiano, Major Professor

We have read this dissertation
and recommend its acceptance:

Bonnie H. Ownley

Timothy A. Rinehart

Benjamin M. Fitzpatrick

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

Evaluation of genetic diversity of flowering dogwood (*Cornus florida* L.) in the eastern United States using microsatellites.

A Dissertation
Presented for the
Doctor of Philosophy Degree
The University of Tennessee, Knoxville

Denita Hadziabdic
May 2010

Copyright © 2010 by Denita Hadziabdic
All rights reserved.

DEDICATION

I would like to dedicate this work to my family for their unconditional love, support, and never-ending faith in me. In particular, I would like to express my sincere gratitude to my mom, Sefika Hadziabdic, for unwavering moral support during the times I needed her the most.

ACKOWLEDGMENTS

I would like to express my sincere gratitude to my major professor, Dr. Robert N. Trigiano, for his guidance, encouragement, and above all, his friendship during my graduate studies here at the University of Tennessee. Thank you for believing in me.

I would like to thank my committee members, Dr. Bonnie H. Ownley, Dr. Timothy A. Rinehart and Dr. Benjamin M. Fitzpatrick for their help and valuable suggestions. It was my true pleasure to get to know and work with you.

My deepest thanks to Dr. Xinwang Wang for his mentorship and invaluable guidance.

Special thanks to my labmates, Phillip Wadl and Deborah Robertson for their help, guidance, friendship, and great sense of humor, which have made the lab a great working environment and wonderful place to be.

I am also very thankful to numerous people who collected dogwood samples throughout the eastern United States and made this project possible. Special thanks to Sarah Boggess, Nicholas Fedor, Ashley Gann, Chris Rosson, and Ben Lindner for their laboratory and technical assistance.

I thank Matthew Johnson for his invaluable help with ArcView.

Thanks to everyone in the department of Entomology and Plant Pathology, my fellow graduate students, and everyone who made my graduate life interesting. It has been a true pleasure working with all of you.

Most importantly, I would like to thank my family: my mom, Sefika, for her selfless love and support throughout all these years; my sisters, Sanela Sahinovic and Enida Islamagic, as well as their families.

This work was supported by USDA Agreement no. 58-6404-2-0057. Mention of trade names or commercial products in this dissertation is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the University of Tennessee or the U.S. Department of Agriculture.

ABSTRACT

Flowering dogwood (*Cornus florida* L.) populations have experienced severe declines caused by dogwood anthracnose in the past three decades. Mortality has ranged from 48 to 98%, raising the concern that genetic diversity of this native tree has been reduced significantly. Microsatellite data were used to evaluate the level and distribution of genetic variation throughout much of the native range of the tree. In the first conducted study, we found that genetic variation in areas affected by anthracnose was as high as or higher than areas without die-offs. We found evidence of four widespread, spatially contiguous genetic clusters. However, there was little relationship between geographic distance and genetic difference. These observations suggest that high dispersal rates and large effective population sizes have so far prevented rapid loss of genetic diversity. The effects of anthracnose on demography and community structure are likely to be far more consequential than short-term genetic effects. The second study examined levels and distribution of genetic variation of *C. florida* throughout Great Smoky Mountains National Park (GSMNP). Significant genetic structure at both landscape and local levels were found. We infer that two genetic clusters exist within the park, mostly separated by the main dividing ridge of the Great Smoky Mountains. The differentiation is statistically significant, but subtle, with gene flow evident through low-elevation corridors. It seems unlikely that recent demographic dynamics have resulted in a depletion of genetic variation in flowering dogwoods.

Table of Contents

Chapter 1. Introduction 1
Microsatellites or SSRs (Simple Sequence Repeats) 2
References 6
Chapter 2. Genetic diversity of flowering dogwood maintained despite massive mortality caused by dogwood anthracnose. 9
Abstract 10
Introduction 10
Materials and Methods 14
Study system 14
DNA Collection, Isolation and PCR Amplification 16
Data Analyses 18
Results 21
Discussion 22
References 27
Appendix 1: Tables 36
Appendix 2: Figures 43
Chapter 3. Fine scale genetic structure of flowering dogwood in the Great Smoky Mountains National Park. 51
Abstract 52
Introduction 52
Materials and Methods 59
Study system 59
Cornus florida L. (Flowering dogwood) 59
Discula destructiva Redlin 61
Field Sampling 64
DNA Collection, Isolation and PCR Amplification 65
Data Analyses 67
Results 69
Discussion 71
References 78
Appendix 3: Tables 89
Appendix 4: Figures 101
Vita 108

Lists of Tables

Table 2-1. Cornus florida sampling locations throughout eastern United States with GPS coordinates and county location……36

Table 2-2. Primer information for 18 microsatellite loci used to analyze 234 *Cornus florida* samples collected from 46 sampling localities in the eastern United States……39

Table 2-3 (A-B). Estimates of Analysis of Molecular Variance (AMOVA) from dogwood microsatellite data using Arlequin (ver 3.1). Two analyses were conducted: the first included all sites as one hierarchical group (A); the second analysis accounted for sampling sites grouped according to regions identified by the program STRUCTURE (TX was excluded from this analysis since it did not contain any subpopulations) (B)……..41

Table 3-1. *Cornus florida* sampling locations throughout Great Smoky Mountains National Park, United States with GPS coordinates and trail/road location within the park……..…89

Table 3-2. Primer information for 19 microsatellite loci used to analyze 194 *Cornus florida* samples collected in Great Smoky Mountains National Park, United States……....97

Table 3-3. Nei's estimation of heterozygosity across all loci in *Cornus florida* populations was computed using FSTAT version 2.9.3. HP-Rare version 1.0 was used for allelic richness calculations. Allelic richness (A_R) for both populations, observed heterozygosity (Ho), expected heterozygosity (gene diversity) under Hardy Weinberg equilibrium (He), total expected heterozygosity (Ht), inbreeding coeficient (F_{IS}) ……99

Table 3-4. Estimates of Analysis of Molecular Variance (AMOVA) from *Cornus florida* microsatellite data using Arlequin (version 3.1). This analysis included all sampling sites as one hierarchical group……100

List of Figures

Figure 2-1**.** Sampling localities and distribution of *Cornus florida* throughout eastern United States (total of 234 individuals). Counties shaded in beige had confirmed cases of dogwood anthracnose as of 2006. Pie charts at each locality represent average inferred ancestry of individuals for each cluster estimated by STRUCTURE. Dogwood anthracnose data courtesy of United States Department of Agriculture (USDA) - Forest Service-Forest Health Protection; distribution map courtesy of United States Geological Survey (USGS)……43

Figure 2-2. Distribution map of *Cornus florida* in the eastern United States. Image courtesy of United States Geological Survey (USGS)……45

Figure 2-3. Tukey boxplots representing no association between dogwood anthracnose outbreak and allelic richness (A) and gene diversity (heterozygosity) (B) within localities. The box represents 50% of ordered data stretching between the lower and the upper quartiles. The bar in the box indicates the median, and the whiskers extending from the boxes include all data within 1.5 $\times$ the interquartile range. Extreme outliers are shown as dots. The notches illustrate approximate 95% confidence intervals for the medians. Since these notches overlap broadly, there is no evidence that the two medians differ significantly…46

Figure 2-4. Delta K (Δk) calculations for unbinned (a and c) and binned (b and d) data. Determination of the number of clusters (k) vs. the second order rate of change in k (Δk) (upper 2 graphs). The clear maximum for Δk at k=4 indicates that four clusters fit the microsatellite data best for *Cornus florida* populations. The number of clusters (k) vs. the negative natural log of the likelihood [-ln (likelihood)] of the microsatellite data (lower 2 graphs)……47

Figure 2-5. Bar plots showing Bayesian assignment probabilities using software STRUCTURE 2.2 for four clusters (K=4) (TN-Tennessee, AL-Alabama, MS-Mississippi, NC-North Carolina, LA-Louisiana, FL-Florida, GA-Georgia, KY-Kentucky, OH-Ohio, PA-Pennsylvania, SC-South Carolina, WV-West Virginia, TX-Texas) divided according to geographical location. The proportion of each bar that is yellow, red, green and blue indicates the assignment probability of individuals to each of these clusters, respectively……49

Figure 2-6 (A-D). Scatterplots of pairwise genetic distances (a_r) vs. geographical distances (km) of all sampled *Cornus florida* individuals (A), and partitioned into sites with predominant membership in the northern cluster (B), middle cluster (C), and southern cluster (D). *P*-values are from Mantel tests with 10,000 randomizations. Values on the x-axis are in log scale. Lines indicate the best fit of least-squares regression of a_r on log distance……50

Figure 3-1. Confirmed cases of dogwood anthracnose for the past three decades. Dogwood anthracnose data courtesy of United States Department of Agriculture (USDA) - Forest Service-Forest Health Protection; distribution map courtesy of United States Geological Survey (USGS)……...101

Figure 3-2. *Cornus florida* sampling locations in the Great Smoky Mountains National Park……103

Figure 3-3. Delta K (Δk) calculations for unbinned (a and c) and binned (b and d) data. Determination of the number of clusters (k) vs. the second order rate of change in k (Δk) (upper 2 graphs). The clear maximum for Δk at k=2 indicates that two clusters fit the microsatellite data best for C*ornus florida* populations. The number of clusters (k) vs. the negative natural log of the likelihood [-ln (likelihood)] of the microsatellite data (lower 2 graphs)…….....105

Figure 3-4. Bar plots showing Bayesian assignment probabilities using software STRUCTURE 2.2 for two clusters (K=2) (Cluster 1–north of Appalachian trail and Cluster 2 –south of Appalachian trail). The proportion of each bar that is red or green indicates the assignment probability of *Cornus florida* individuals to each of the two clusters, respectively……..……..107

Chapter 1. Introduction

Microsatellites or SSRs (Simple Sequence Repeats)

Microsatellites are mutation-prone DNA tracts composed of tandem repetitions of relatively short motifs, which contain several (one to six nucleotides) units (Fujimori et al.; 2003, Kashi and King 2006) in tracts up to 10^2 bp (Rakoczy-Trojanowska and Bolibok 2004). They can be classified into the following three groups: pure, compound and interrupted repeats (Weber 1990), and are distributed through the whole genome. Microsatellites distributions vary by species and chromosome and are flanked by highly conserved sequences (Chambers and MacAvoy 2000). In plants, the most frequent microsatellite repeats are composed of dinucleotide motifs, usually $(AT)_n$ and $(GT)_{n,}$ whereas $(AC)_n$ repeats are common in animals (Rakoczy-Trojanowska and Bolibok 2004). Microsatellites are less abundant in plants when compared to animals (Lagercrantz et al. 1993), however, they exist in both nuclear and the organellar genomes (Pleines et al. 2009).

Microsatellites are used as a genetic marker due to their ubiquitous distribution in the genome and have a high level of allelic variation (Rakoczy-Trojanowska and Bolibok 2004). Variation in the number of tandemly repeated units is primarily due to strand slippage during DNA replication (Schlotterer and Tautz 1992; Agarwal et al. 2008). Microsatellites represent a significant resource for creating genetic and physical genome maps, distinguishing individuals, investigating genetic relatedness, and studying genome organization (Thomas et al. 1993; Debener and Mattiesch 1998). These markers have shown to be useful for integrating the genetic, physical and sequence-based physical maps in many

plant species, simultaneously providing breeders and geneticists with an efficient tool to link phenotypic and genotypic variation (Varshney et al. 2005). Their popularity is also due to the fact they are assayed using PCR, thus requiring minute amounts of tissue for the analysis. Since molecular markers are not affected by environmental factors or age of the organism, they are an excellent tool to provide an unbiased insight into species under examination. Microsatellites can be used to detect DNA polymorphisms (genetic similarity or differences between individuals or populations) and like RFLPs, these markers are codominant (both alleles of a diploid organism are detected); therefore, homozygote and heterozygote individuals can be distinguished (Weising et al. 1995). Microsatellites are preferred markers in plant breeding programs due to their uniform genome coverage, high levels of polymorphism, co-dominance and reproducibility (Pejic et al. 1998). In addition, they can be multiplexed to allow the screening of multiple markers simultaneously as well as to detect duplications within breeding populations (Wadl 2009).

Most microsatellite loci are identified from small (less then 1000 bp) insert genomic libraries using oligonucleotide probes (Maguire 2001). Commonly used methods for isolation involve several step processes. Library construction starts with digestion of genomic DNA, ligation into a plasmid vector, transformation of the vector into *Escherichia coli*, followed by hybridization with a labeled microsatellite oligonucleotide probe, sequencing of positive clones, designing of primers, locus specific PCR amplification, and finally, identification of polymorphic loci (Maguire 2001).

Microsatellites are present in both non-coding and coding regions (which can cause differences in their composition). Loci located in coding regions may activate or inactivate genes or truncate proteins (Ince et al. 2010). However, their frequency is higher in transcribed regions, especially in mRNA untranslated regions (UTRs) (Hongtrakul et al.1998; Morgante et al. 2002; Panaud et al. 1995). Microsatellites found in 3'-UTRs are involved in gene silencing and transcription slippage; microsatellites within 5'-UTRs have effects on gene transcription and/or regulation and microsatellites in introns have effects on gene transcription and regulation, messenger RNA spicing, and gene silencing (Ince et al. 2010).

Microsatellites have proved to be useful as molecular markers in many aspects of research, particularly for population studies. However, they are not without limitations. Microsatellites developed for particular species can often be applied to both closely related and unrelated species, but the percentage of amplified loci typically decreases with increasing genetic distance (Jarne and Lagoda 1996). Additionally, non-amplifying alleles or null alleles are often observed since mutations in one or both primer binding sites prevent PCR amplification (Weising et al. 1995). These can cause problems for population genetic studies because of their recessive behavior since they are recognized in the homozygous state only (de Sousa et al. 2005). Additionally, high variability of these markers is also the major disadvantage, as microsatellite loci often show high levels of homoplasy, especially when distant populations or higher taxonomic levels are studied (Jakob et al. 2007; Pleines et al. 2009).

Even though microsatellites have been considered to be selectively neutral and randomly distributed throughout the genome, recent studies have found evidence of evolutionary role of microsatellites as important sources of adaptive genetic variation within and between species (Kashi and King 2006).

Two studies were conducted in this doctoral research project. In both studies microsatellite loci were used to study genetic differentiation and population structure of flowering dogwoods in the eastern United States and Great Smoky Mountains National Park (GSMNP), respectively. By assessing genetic diversity and population dynamics of these native trees, the following issues were addressed in the first study: (1) to quantify genetic diversity of existing flowering dogwood populations to provide a baseline for understanding future population changes; (2) to compare levels of diversity in regions that have experienced dogwood anthracnose-associated die-offs with those in regions with no history of die-offs; and (3) to describe population structure and geographic variation of flowering dogwood for comparison with other North American species in the southeast. In the second study conducted in the GSMNP, the following questions were addressed: (1) How much genetic diversity exists within and among populations of flowering dogwoods? (2) Is there an evidence of population structure? (3) Is there an effect of seed vs. pollen dispersal on genetic structure? (4) Can a baseline be determined for understanding future population changes in newly colonized areas?

References

Agarwal M, Shrivastava N, Padh H (2008) Advances in molecular marker techniques and their applications in plant sciences. Plant Cell Rep 27:617–631.

Chambers GK, MacAvoy ES (2000) Microsatellites: consensus and controversy. Comp Biochem Physiol 126:455-476.

Debener T, Mattiesch L (1998) Effective pairwise combination of long primers for RAPD analyses in roses. Plant Breeding 117:147–151.

De Sousa SNE, Finkeldey R, Gailing O (2005) Experimental Verification of microsatellite null alleles in Norway spruce *(Picea abies* [L.] Karst.): Implications for Population Genetic studies. Plant Mol Biol Rep 23:113-119.

Fujimori S, Washio T, Higo K, Ohtomo Y, Murakami K, Matsubara K, Kawai J, Carninci P, Hayashizaki Y, Kikuchi S, Tomita M (2003) A novel feature of microsatellites in plants: a distribution gradient along the direction of transcription. FEBS Lett 554:17-22.

Hongtrakul V, Slabauch MB, Knapp J (1998) DFLP, SSCP, and SSR markers for 9-stearoyl-acyl carrier protein desaturases strongly expressed in developing seeds of sunflower: intron lengths are polymorphic among elite inbred lines. Mol Breeding 4:195-203.

Ince AG, Karaca M, Onus AN (2010) Differential expression patterns of genes containing microsatellites in *Capsicum annuum* L. Mol Breeding 25:645–665.

Jakob SS, Ihlow A, Blattner FR (2007) Combined ecological niche modeling and molecular phylogeography revealed the evolutionary history of *Hordeum marinum* (Poaceae)—niche differentiation, loss of genetic diversity, and speciation in Mediterranean Quaternary refugia. Molec Ecol 16:1713–1727.

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

Kashi Y, King DG (2006) Simple sequence repeats as advantageous mutators in evolution. Trends Genet 22: 253-259.

Lagercrantz U, Ellegren H, Andersson L (1993) The abundance of various polymorphic microsatellite motifs differs between plants and vertebrates. Nucleic Acids Res 21:1111–1115.

Maguire TL (2001) Producing and exploiting enriched microsatellite libraries. In: Henry RJ (ed) Plant Genotyping: The DNA Fingerprinting of Plants, CAB International, Wallingford, UK, pp.193-209.

Morgante, M, Hanafer M, Powell W (2002) Microsatellites are preferentially associated with nonrepetitive DNA in plant genomes. Nat Genet 30:194-200.

Panaud O, Chen X, McCouch SR (1995) Frequency of microsatellite sequences in rice (*Oryza sativa* L.). Genome 38:1170-1176.

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castaglioni P, Taramino G, Motto M (1998) Comparative analysis of genetic similarity among

maize inbred lines detected by RFLPs, RAPDs, SSRs and AFLPs. Theor Appl Genet 97:1248-1255.

Pleines T, Jakob SS, Blattner FR (2009) Application of non-coding DNA regions in intraspecific analyses. Plant Syst Evol 282:281–294.

Rakoczy-Trojanowska M, Bolibok H (2004) Characteristics and comparison of three classes of microsatellite-based markers and their application in plants. Cell Mol Biol Lett 9:221–238.

Schlotterer C, Tautz D (1992) Slippage synthesis of simple sequence DNA. Nucleic Acids Res 20:2211–2215.

Thomas MR, Satsumoto S, Cain P, Scott NS (1993) Repetitive DNA of grapevine: classes present and sequences suitable for cultivar identification. Theor Appl Genet 86:173–180.

Varshney RK, Graner A, Sorrells ME (2005) Genic microsatellite markers in plants: features and applications. Trends Biotechnol 23:48-55.

Wadl PA. 2009. Simple sequence repeats (SSRs) and their application to breeding *Cornus* species. Diss. University of Tennessee.

Weber JL (1990) Informativeness of human (dC-dA)n.(dC-dT)n polymorphisms. Genomics 7:524-530.

Weising K, Nybom H, Wolff K, Meyer W (1995) DNA Fingerprinting in Plants and Fungi. CRC Press, Boca Raton, Florida.

Chapter 2. Genetic diversity of flowering dogwood maintained despite massive mortality caused by dogwood anthracnose.

This chapter is a revised version of paper by the same name submitted in the journal *Genetica* in 2009 by Denita Hadziabdic, Benjamin Fitzpatrick, Xinwang Wang, Phillip Wadl, Timothy Rinehart, Bonnie Ownley, Mark Windham and Robert Trigiano.

Hadziabdic D, Fitzpatrick BM, Wang X, Wadl PA, Rinehart TA, Ownley BH, Windham MT, and Trigiano RN. Genetic diversity of flowering dogwood maintained despite massive mortality caused by dogwood anthracnose. *Genetica* (accepted for publication 2010).

My primary contributions to this paper include (1) topic selection and project development into a work relevant to population genetics of flowering dogwood; (2) some sampling of plant material, lab work related to the project and data analysis; (3) gathering and interpretation of literature; (4) summarizing and editing contributions and suggestions from other co-authors into a single paper, and (5) majority of the writing related to this paper.

Abstract

Flowering dogwood (*Cornus florida*) populations have recently experienced severe declines caused by dogwood anthracnose. Mortality has ranged from 48 to 98%, raising the concern that genetic diversity has been reduced significantly. Microsatellite data were used to evaluate the level and distribution of genetic variation throughout much of the native range of the tree. Genetic variation in areas affected by anthracnose was as high as or higher than areas without die-offs. We found evidence of four widespread, spatially contiguous genetic clusters. However, there was little relationship between geographic distance and genetic difference. These observations suggest that high dispersal rates and large effective population sizes have so far prevented rapid loss of genetic diversity. The effects of anthracnose on demography and community structure are likely to be far more consequential than short-term genetic effects.

Introduction

Recently introduced diseases and pests have caused tremendous mortality of various tree species over wide geographic areas (Merkle et al. 2007; Crowl et al. 2008). These die-offs often result in restructuring of forest communities and disruption of ecosystem processes (Hall et al. 2002; Becker et al. 2008; Elliott and Swank 2008; Eschtruth et al. 2008). Well-known examples in North America include the near elimination of American chestnuts [*Castanea dentata* (Marsh.) Borkh] by chestnut blight [*Cryphonectria parastica*, (Murrill)

Barr] (in addition to extensive logging) (Elliott and Swank 2008); American elms (*Ulmus americana* L.) due to Dutch elm disease [*Ophiostoma ulmi* (Buisman) Nannf] (Brasier 2000) and Eastern hemlocks [*Tsuga canadensis* (L.) Carrière] by hemlock wooly adelgids [*Adelges tsugae* Annand] (Havill et al. 2006). In the United States alone, the projected economic loss due to invasion of non-native species is over $120 billion annually and about one-fourth of this loss is caused by exotic plant pathogens (Pimentel 1997; USBC 2001; Pimentel et al. 2005).

In the past three decades, flowering dogwoods (*Cornus florida* L.) have been severely affected by dogwood anthracnose caused by *Discula destructiva* Redlin (Redlin 1991) (Figure 2-1, in appendix). Poor seed production due to decreased flowering (Rossell et al. 2001) and greater susceptibility of smaller, younger trees (Hiers and Evans 1997; Holzmueller et al. 2006; Jenkins and White 2002) resulted in reduced dogwood regeneration and dramatic decline of native dogwood populations. Mortality of flowering dogwood caused by dogwood anthracnose has ranged from 48 to 98% in the northeast and Appalachian highlands (Sherald et al. 1996; Hiers and Evans 1997; Williams and Moriarity 1999; McEwan et al. 2000; Jenkins and White 2002). In 1992, a survey of trees in Great Smoky Mountain National Park (GSMNP) revealed that 25% of flowering dogwoods had been killed by dogwood anthracnose (Windham et al. 1995). Two years later, an estimated 75% of the remaining trees had been eliminated (Windham et al. 1995). In a study by Wilds (1997), dogwood anthracnose infection was as high as 98% in experimental, biodiversity plots located in the western part of GSMNP. In the periods from 1977-1979 and 1995-2000, Jenkins

and White (2002) collected flowering dogwood data from 86 permanent plots located in the western portion of GSMNP. In their study, flowering dogwood density within different forest types–(typic cove, acid cove, alluvial, oak-hickory, and oak-pine stands) decreased between two sampling intervals, ranging from 69-94%. The largest decline, over 94% was among smaller, younger trees (1 to 4.9 cm diameter) in all forest types (Jenkins and White 2002).

With large-scale mortality, genetic variation of species can decline rapidly (Hawley et al. 2006), with potentially deleterious effects on the ability of tree populations to regenerate and respond to future changes in selection pressures. As a result, the relationships between plant population size, fitness and genetic diversity are of fundamental importance in plant ecology, evolution and conservation (Leimu et al. 2006). In smaller, fragmented populations, the loss of genetic variation through genetic drift may cause reduced mean population fitness, elevating the risk of extinction and reducing prospects for adaptive change (Fischer and Matthies, 1998; Reed and Frankham, 2003; Severns 2003).

However, no broad-scale survey of genetic variation for *C. florida* trees has been published, and there may be reason to expect tree populations to be somewhat resistant to apparent bottleneck effects, at least in the short term. The full genetic effect of a population bottleneck may take many generations to manifest (Chakraborty and Nei 1977; Nei 1975). In a simple model of drift with no mutation, the time it takes heterozygosity to fall to one-half of its initial value is on the order of *N* generations, where *N* is the effective population size (Wright 1931). Therefore, the effect of a bottleneck depends on the actual population

size, rather than on the percent reduction in abundance, because a high level of heterozygosity may be maintained following a bottleneck (Nei 1975). For example, if 90% mortality takes a population from $N = 10^6$ to $N = 10^5$, expected heterozygosity in the reduced population will remain greater than 99% of its original value for 2,000 generations (Hartl and Clark 1997).

Flowering dogwoods are widespread in the eastern United States and do not appear to be clustered into spatially discrete demes (USDA, NRCS, 2008) (Figure 2-1; Figure 2-2, in appendix). Thus, it has been difficult to evaluate the geographic scale of demographically or genetically meaningful "populations." If the genetic neighborhood (Wright 1946) of an individual tree is typically small, then large die-offs may cause rapid and dramatic declines in local genetic diversity and increase differences among localities. On the other hand, if geographically distant trees are connected by extensive gene flow, the effective population size may be so large that ecologically dramatic die-offs have little short-term effect on genetic diversity.

In this study, we used 18 microsatellite loci to study genetic differentiation and population structure of flowering dogwoods in the eastern United States. By assessing genetic diversity and population dynamics of these native trees, we aimed to address the following issues: (1) to quantify genetic diversity of existing flowering dogwood populations to provide a baseline for understanding future population changes; (2) to compare levels of diversity in regions that have experienced dogwood anthracnose-associated die-offs with those in regions with no history of die-offs; (3) to describe population structure and geographic

variation of flowering dogwood for comparison with other North American species in the southeast.

Materials and Methods

Study system

Flowering dogwoods are ubiquitous understory trees throughout most eastern United States forest types (Hannah 1993; Martin 1987; Quarterman et al. 1972). It has been argued that flowering dogwoods have played an important role in structure and function of Appalachian forests (Hiers and Evans 1997). Individual dogwood trees are capable of producing up to 10 kg of high protein berries each fall, supplying a valuable food source for more than 50 bird and numerous small game species (Rossell et al. 2001; Holzmueller et al. 2006). Dogwoods are important in rapid nutrient cycling (64% of reduction in litter mass in two years) within forest communities, thus providing high concentrations of available calcium (2-3.5% dry weight) (Jenkins and White, 2002).

Flowering dogwoods are most common in second growth forests and human suppression of disturbance might contribute to local and regional declines (Pierce et al. 2008; Jenkins and White 2002; McEwan et al. 2000). However, extremely rapid and severe die-offs have been associated with the fungus *D. destructiva*, which causes dogwood anthracnose disease (Redlin 1991; Daughtrey et al. 1996; Holzmueller et al. 2006). Purple-rimmed lesions initially appear on the leaf margins and extend through the veins into the petiole in trees infected by dogwood anthracnose (Jenkins and White 2002). The disease usually

starts in the lower canopy and moves upwards in the tree. During cool, wet weather, the first signs of dogwood anthracnose are gray or drooping foliage hanging on the twigs. With disease progression, the fungus moves from the leaf petiole into stems resulting in twig and branch die-back due to the formation of annual cankers. Eventually the disease may cause death of the tree. Even though spread of dogwood anthracnose southward appeared to be relatively uninhibited, several factors, including light availability, moisture and evaporative potential of the leaf surface can determine the probability and severity of the infection (Hibben and Daughtrey 1988; Hiers and Evans 1997). Disease symptoms are usually more apparent on slopes facing northeast (understory canopies with low evaporative potential) than in southwest facing slopes (exposed canopies) (Chellemi and Briton 1992; Chellemi et al. 1992; Hiers and Evans 1997, Jenkins and White 2002).

Currently, there are no management options for controlling dogwood anthracnose in large forested areas (Holzmueller et al. 2006). Fungicides are impractical and conventional selection and breeding approaches typically used for development of pathogen and/or pest resistant crop species are very difficult to apply to forest trees (Merkle et al. 2007). This is partially due to long juvenile periods of many forest trees including flowering dogwood (about seven years) before they can be selected for breeding qualities or resistance purposes (Merkle et al. 2007). Breeding for resistance is also hindered by the lack of resistant germplasm; only one tree, the ‘Appalachian Spring’ has confirmed resistance to

the disease (Windham et al. 1998), self-incompatibility (Wadl et al. 2009), and therefore not subject to inbreeding depression (Trigiano, unpublished data).

DNA Collection, Isolation and PCR Amplification

Genomic DNA was isolated from young leaves or unopened flower buds of flowering dogwood collected from 234 individuals at 46 sampling localities throughout the eastern United States (Figures 2-1 and 2-2, in appendix). Samples used in this population study were collected during 2006-2008 from the following states: Alabama, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Ohio, Pennsylvania, South Carolina, Tennessee, Texas, and West Virginia. For each sampling location, GPS coordinates were recorded (Table 2-1, in appendix). Identification of individuals collected from each location was based on morphological characteristics of the species. We made use of extensive sampling from southern portions of the range, where no anthracnose has been detected (Figure 2-1; Table 2-1, in appendix). Flowering dogwood samples were taken 30 years, or approximately four generations, after the first appearance of anthracnose in New York (Daughtrey and Hibben 1994).

Plant materials were placed into labeled sample bags on ice and stored at -80℃ until used. Genomic DNA was extracted from sa mples using Qiagen DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA) according to the manufacturer's protocol. All dogwood samples were ground in autoclaved mortars and pestles using liquid nitrogen. The concentration and purity of DNA samples were measured with a NanoDrop ND-1000 spectrophotometer

(NanoDrop Technologies, Inc., Wilmington, DE, USA) before analysis. PCR amplifications were performed in 10 µl reaction mixtures containing 1 µl DNA template (4 ng/µl), 1 µl of GeneAmp 10×PCR Buffer II (Applied Biosystems, Branchburg, NJ, USA), 1 µl of 20 mM $MgCl_2$, 1 µl of each 2 mM of dNTPs, 1 µl of 2.5 µM specific microsatellite primers, 0.08 µl of 5 U AmpliTaq Gold® DNA polymerase, and 4.92 µl of sterile nanopure water. Eighteen polymorphic microsatellite primer pairs, developed from a previous study (Wang et al. 2008) were selected and used for analysis (Table 2-2, in appendix). Amplification reactions were temperature cycled in 96-well plates using an Eppendorf Autorisierter Thermocycler (Eppendorf AG, Hamburg, Germany) with the following program: 94℃ for 3 min, followed by 35 cycles of 94℃ for 40 seconds, 58℃ for 40 seconds, 72℃ for 30 seconds, and a final extension at 72℃ for 4 min.

PCR products were analyzed with the QIAxcel Capillary Electrophoresis System (Qiagen, Valencia, CA, USA) using an internal 25-bp DNA size marker. Data were automatically recorded and exported using BioCalculator™ software, which provides both a gel view and an electropherogram of the separated PCR products (alleles) (Wang et al. 2009).

Program FLEXIBIN (Amos et al. 2007) was used for automated binning of allelic data. In order to achieve the best possible fit and find the most suitable binning parameter, the program uses a simple algorithm to conduct searches in two different phases and steps through all possible parameter combinations. When the best fit values are established, all alleles are replaced with their repeat

unit equivalents and graphical output file with summary statistics is generated. Binned and unbinned data gave identical results and we present only analyses of binned data (Figure 2-4, in appendix).

Data Analyses

To address the question of whether genetic variation is lower in regions affected by dogwood anthracnose, we estimated Nei's gene diversity (Nei 1987) and allelic richness for each site and then compared each of those summary statistics to a null distribution generated by randomizing sites between affected and unaffected regions. Nei's gene diversity is the average probability of nonidentity for two randomly chosen alleles from a randomly chosen locus (also the average expected heterozygosity). Allelic richness was estimated using rarefaction to three trees per site following Kalinowski's method (Kalinowski 2005). Both allelic richness and gene diversity calculations were performed using HP-Rare 1.0 software package (Kalinowski 2005). We used randomization tests (one tailed test) to evaluate whether anthracnose sites and no-anthracnose sites were more different than expected by chance. Sampling sites were randomly relabeled (anthracnose or no-anthracnose) 10,000 times and we recorded the number of randomizations with a greater mean difference between groups than the observed data.

The population structure across the range was evaluated using Bayesian cluster analysis (Pritchard et al. 2000), analysis of molecular variance (Excoffier et al. 1992), and Mantel tests for isolation by distance (Bohonak 2002). Bayesian

cluster analysis was performed using STRUCTURE version 2.2 (Pritchard and Donnelly 2001). This program groups individuals on the basis of their multilocus genotypes. The method is based on the assumption that Hardy–Weinberg or linkage disequilibria are caused by population structure and attempts to find population groupings that are not in disequilibrium (Pritchard and Donnelly 2001). This is an individual-based analysis that makes no assumptions about the relationship between sample sites and population structure. Evanno's method (Evanno et al. 2005) was used to identify the appropriate number of clusters using the *ad hoc* statistic Δ*k,* which is based on the second order rate of change in the log probability of the data between successive values of *k*. We used an admixture model with correlated allele frequencies (assuming no prior information of population origin). For assessing alternative values of *k*, we ran ten independent runs for each *k* value between one and eight for 100,000 generations and a burn-in period of 100,000 iterations. Additional 10 runs for four clusters (the best *k*, see results) were run for 1,000,000 generations and a burn-in period of 1,000,000 iterations to verify that results were consistent across independent Markov Chains.

Differentiation among sample locations was quantified using a hierarchical analysis of variance (AMOVA - Analysis of Molecular Variance) using Arlequin (version 3.1). Two different analyses were conducted, the first one included all sites as a single hierarchical group and the second one grouped sites according to majority representation in the clusters identified by STRUCTURE. AMOVA complements STRUCTURE by providing estimates of the amount of variation

explained by grouping sites as implied by STRUCTURE, but should not be taken as independent confirmation of that specific hierarchical structure.

Both STRUCTURE and AMOVA assume hierarchical clustering of individuals into genetic sub-populations. Alternatively, individuals might be continuously distributed throughout a geographic range with genetic similarity falling off with distance (because procreation is more likely to involve spatially close parents), but no inherent clustering of individuals into demes (Wright 1943). Under this model, each individual has a unique "genetic neighborhood" of likely parents (or mates). In a homogeneous habitat with "gaussian" dispersal, Rousset (2000) predicted a linear relationship between a pairwise genetic distance between individuals [a_r, analogous to $F_{ST}/(1 - F_{ST})$] and the logarithm of geographic distance. The slope of the relationship is an inverse function of population density and dispersal distance (providing an estimate of genetic neighborhood "size" in numbers of individuals). While the theoretical model is unlikely to apply to dogwoods (because habitats are not homogenous and dispersal can occur at both the pollen and seed stages), we can use the statistical analysis to evaluate whether geographic structure is best explained by hierarchical clustering, a continuous distribution, or a mixture of the two. We estimated a_r following Roussett (2000) and compared it to pairwise geographic distance (km) using R 2.7.1 (Hornik 2008). Geographic distances were computed as great circle distance (accounting for the curvature of the Earth). Statistical significance of the correlation between geographic and genetic distance was evaluated with a modified Mantel test in which sample sites (rather than

individuals) were randomized in order to account for potential non-independence of individuals from the same locality (within a 1 km radius).

Results

All 18 microsatellite loci were highly polymorphic across the 46 sampling localities and binned and unbinned data gave identical results; only analyses of binned data are presented. The average allelic richness was 3.08 (calculated by rarefaction to three trees per site) and gene diversity (heterozygosity) averaged 0.65. There was no evidence that sites within the region affected by anthracnose were genetically depauperate relative to sites from unaffected regions (P=0.8803 and P=0.6267 one-tailed p-values for difference in means for allelic richness and gene diversity, respectively) (Figure 2-3, in appendix). In fact, the levels of genetic variation reported here are on par with the most diverse animals and plants reported in the literature (El Mousadik and Petit 1996; Innan et al. 1997; Bachtrog et al. 2000; Neff and Gross 2001).

STRUCTURE analysis resulted in a clear maximum for Δk at k =4 [four different clusters (Figures 2-4 and 2-5, in appendix)]. These clusters were more or less spatially contiguous, including “northern” (Kentucky, Ohio, Pennsylvania, South Carolina, and West Virginia), “middle” (Alabama, Mississippi and Tennessee), “southern” (Florida, Georgia, Louisiana, and North Carolina), and Texas clusters (Figure 2-5, in appendix). The multiple runs of the same k, used in this analysis, produced highly consistent individual assignment probabilities. Furthermore, an additional 10 runs for four clusters were run for 1,000,000

generations and a burn-in period of 1,000,000 iterations to verify that results were consistent across independent Markov Chains.

AMOVA analysis (Excoffier et al. 2005) indicated that most of the genetic variance can be explained by individual variation rather than divergence among sample localities 84.24 and 15.76%, respectively (Table 2-3A, in appendix). When genetic variances within individuals, among individuals within and among sites within three clusters (northern, middle, and southern) were compared, they partitioned similarly to the previous finding, 51.57, 37.31 and 8.96% respectively (Table 2-3B, in appendix). The Bayesian clusters explained only 2.03% of variation. Due to small sample size (13 individuals from a single locality), the Texas cluster was not included in this analysis.

Individual-based analysis of isolation by distance across all samples showed a modest positive relationship between geographic distance (km) and genetic differentiation (a_r) (Figure 2-6A). However, when sites were grouped by inferred cluster (from STRUCTURE) and analyzed separately, there was no evidence of isolation by distance within clusters (Figure 2-6B-D). In terms of Rousset's (2000) model, we cannot statistically reject infinite genetic neighborhoods (numbers of individuals) within the broadscale geographic clusters suggested by STRUCTURE.

Discussion

Despite severe die-offs of flowering dogwood over the last 30 years, there is no measurable difference in genetic diversity between parts of the geographic

range affected and unaffected by anthracnose. Although it is possible that other factors have depressed genetic variation in the southern part of the range, making the comparison an invalid test of the effect of anthracnose, the fact that all sites had extremely high levels of variation indicates that the genetic effects of recent demographic fluctuations have been minimal. More likely, the population structure of flowering dogwoods is such that even a 10-fold reduction in local abundance has little short-term effect on loss of genetic variation.

In particular, the typical genetic neighborhood (Wright 1943; Wright 1946) likely still contains hundreds or thousands of trees after severe die-off events. The proposition of large genetic neighborhoods was supported by our inference that flowering dogwoods in eastern North America fall into a few widespread genetic clusters with no evidence of isolation by distance within clusters. This analysis supports the contention that the geographically contiguous genetic clusters identified by STRUCTURE resemble extended genetic neighborhoods where gene flow between the most distant sites (800 – 1,000 km apart) is just as likely as gene flow between sites 10 km apart.

The geographic distributions of the Bayesian clusters were congruent with phylogeographic patterns exhibited by many other plants and animals in eastern North America (Soltis et al. 2006; Fontanella et al. 2007). We infer that this spatial genetic structure reflects the historical population structure of flowering dogwoods resulting from the distribution of suitable habitats over the Quaternary and from the tendency of dogwoods to be dispersed by birds that feed on dogwood fruits (Rossell et al. 2001).

These analyses suggest that any given stand of dogwoods shares a gene pool with thousands of other trees spread over thousands of square kilometers. This inference is consistent with the known breeding system and dispersal ecology of flowering dogwoods (Sork et al. 2005). Seeds in fruits eaten by migratory birds are likely to be deposited many kilometers from their sources. In addition, flowering dogwoods are obligate outcrossers with gametophytic self-incompatibility (Reed 2004). This breeding system enforces a degree of disassortative mating and maintenance of local genetic variation. As self-incompatible species, genetic variation of flowering dogwood is expected to be greater within populations then between populations when compared with self-compatible species (Hamrick and Godt 1989; Leimu et al. 2006). In a continuous population structure, the relationship between genetic and geographic distance might be expected to be much weaker in self-incompatible species. These predictions are supported in this study.

Flowering dogwood trees are native to the United States (Witte et al. 2000) and are important components of the forests they inhabit not only for their ornamental values during all seasons, but also as important source of food for wildlife. As a result, the loss of this species would cause a serious ecological impact throughout the southeastern region of the United States. The flowers, leaves, bark and fruits provide nourishment for a variety of invertebrates and vertebrates, including American beavers (*Castor canadensis* Kuhl), which feed largely on the bark and outer layers of deciduous trees such as dogwood (Linzey and Brecht 2003a), and Eastern gray squirrels (*Sciurus carolinensis* Gmelin),

which prefer dogwood berries (Linzey and Brecht 2003b). Fruits of flowering dogwood contain one of the highest fat content (approximately 18%) of any food available in the forest, and also contain high levels of available calcium (Blair et al. 1983; Halls and Epps 1969), which is used as a fall food source for neo-tropical migratory birds (Stiles 1980). A recent study reported that individuals of flowering dogwood have primarily disappeared in the sub-canopy layer allowing other species such as *Lindera benzoin* Blume (Northern spicebush), *Viburnum acerifolium* L. (Mapleleaf viburnum), *Nyssa sylvatica* Marsh (Blackgum) and *Tsuga canadensis* (Eastern hemlock) (Hiers and Evans 1997; Jenkins and White 2002) to become more dominant and prevalent components of forests they inhabit (Myers et al. 2004).

Despite high mortality and reduced fecundity caused by dogwood anthracnose, this study confirmed that considerable genetic diversity still exists among native populations of flowering dogwood. We suggest that large-scale dispersal and continuous population structure has so far prevented major genetic consequences from the clearly dramatic local demographic effects of dogwood anthracnose over the last three decades. Unfortunately, long-distance dispersal by migratory birds may also explain the rapid spread of dogwood anthracnose from the north to the south; *D. destructiva*, the causal agent of the disease, has been isolated from both pulp and seeds of dogwood fruit (Rossell et al. 2001). While the impact of anthracnose on genetic diversity may be small, consequences for flowering dogwood population viability and the community structure of eastern North American forests are still of great concern (Daughtrey

and Hibben 1994; Lande 1988). However, more research is needed on the ecology of post-die-off recovery as well as disease ecology and population dynamics in widespread populations of flowering dogwood.

References

Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill VS (2007) Automated binning of microsatellite alleles: problems and solutions. Mol Ecol Notes 7:10-14.

Becker DA, Brittingham MC, Goguen CB (2008) Effects of hemlock woolly adelgid on breeding birds at Fort Indiantown Gap, Pennsylvania. Northeast Nat 15:227–240.

Bohonak AJ (2002) IBD (Isolation by distance): A program for analyses of isolation by distance. J Hered 93:153-154.

Brasier CM (2000) Intercontinental spread and continuing evolution of the Dutch elm disease pathogens. In: Dunn CP (ed) The Elms: Breeding, Conservation and Disease Management. Kluwer Academic Publishers, Boston, pp. 61–72.

Chellemi DO, Britton KO (1992) Influence of canopy microclimate on incidence and severity of dogwood anthracnose. Can J Bot 70:1093-1096.

Chellemi DO, Britton KO, Swank WT (1992) Influence of site factors on dogwood anthracnose in the Nantahala mountain range of western North Carolina. Plant Dis 76:915-918.

Chakraborty R, Nei M (1977) Bottleneck effects on average heterozygosity and genetic distance with the stepwise mutation model. Evolution 31:347–356.

Crowl TA, Crist TO, Parmenter RR, Belovsky G, Lugo AE (2008) The spread of invasive species and infectious disease as drivers of ecosystem change. Front Ecol Environ 6:238-246.

Daughtrey ML, Hibben CR (1994) Dogwood anthracnose: a new disease threatens two native *Cornus* species. Annu Rev Phytopathol 32:61-73.

Daughtrey ML, Hibben CR, Britton KO, Windham MT, Redlin SC (1996) Dogwood anthracnose: Understanding a disease new to North America. Plant Dis 80:349 –357.

Elliott KJ, Swank WT (2008) Long-term changes in forest composition and diversity following early logging (1919–1923) and the decline of American chestnut (*Castanea dentata*). Plant Ecol 197:155–172.

El Mousadik A, Petit RJ (1996) High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. Theor Appl Genet 92:832-839.

Eschtruth AK, Battles JJ (2008). Deer herbivory alerts forest response to canopy decline caused by an exotic insect pest. Ecol Appl 18:360-376.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform Online 1:47-50.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Fischer M, Matthies D (1998) Effects of population size on performance in the

rare plant *Gentianella germanica*. J Ecol 86:195–204.

Fontanella FM, Feldman CR, Siddall ME, Burbrink FT (2007) Phylogeography of Diadophis punctatus: Extensive lineage diversity and repeated patterns of historical demography in a trans-continental snake. Mol Phylogent Evol 46**:**1049-1070.

Hall B, Motzkin G, Foster DR, Syfert M, Burk J (2002) Three hundred years of forest and land-use change in Massachusetts, USA. J Biogeogr 29:1319-1335.

Halls LK, Epps EA (1969) Browse quality influenced by tree overstory in the South. J Wildl Manage 33:1028-1031.

Hamrick JL, Godt MJW (1989) Allozyme diversity in plant species. In: Brown HD, Clegg MT, Kahler AL, Weir BS (eds) Plant Population Genetics, Breeding, and Genetic Resources, Sinauer Associates, Sunderland, Massachusetts, pp. 43–63.

Hannah PR (1993) Composition and development of two Appalachian hardwood stands in North Carolina. J Elisha Mitch Sci 109:87-98.

Hartl DL, Clark AG (1997) Principles of population genetics, 3rd edn. Sinauer Associates, Sunderland, MA.

Havill NP, Montgomery ME, Yu G, Shiyake S, Caccone A (2006) Mitochondrial DNA from hemlock woolly adelgid (Hemiptera: Adelgidae) suggests cryptic speciation and pinpoints the source of the introduction to eastern North America. Ann Entomol Soc Am 99:195-203.

Hawley DM, Hanley D, Dhondt AA, Lovette IJ (2006) Molecular evidence for a founder effect in invasive house finch (*Carpodacus mexicanus*) populations experiencing an emergent disease epidemic. Mol Ecol 15:263-275.

Hiers JK, Evans JP (1997) Effects on anthracnose on dogwood mortality and forest composition of the Cumberland Plateau (U.S.A.). Conserv Biol 11:1430-1435.

Hibben CR, Daughtrey ML (1988) Dogwood anthracnose in the northeastern United States. Plant Dis 72:199-203.

Hornik K (2008), The R FAQ available at: http://CRAN.R-project.org/doc/FAQ/R-FAQ.html

Holzmueller E, Jose S, Jenkins M, Camp A, Long A (2006) Dogwood anthracnose in eastern hardwood forests: what is known and what can be done? J Forest 104:21-26.

Innan H, Terauchi R, Miyashita NT (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana.* Genetics 146:1441-1452.

Jenkins MA, White PS (2002) *Cornus florida* L. mortality and understory composition changes in western Great Smoky Mountains National Park. J Torrey Bot Soc 129:194-206.

Kalinowski ST (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Mol Ecolo Notes 5:187-189.

Lande R (1998) Genetics and demography in biological conservation. Science 241:1455-1460.

Leimu R, Mutikainen P, Koricheva J, Fischer M (2006) How general are positive relationships between plant population size, fitness and genetic variation? J Ecol 94:942–952.

Linzey DW, Brecht C (2003)a American beavers (*Castor canadensis*) - Biodiversity of Great Smoky Mountains National Park, Discover Life in America, Gatlinburg, Tennessee 37738.

Linzey DW, Brecht C (2003)b Eastern gray squirrels (*Sciurus carolinensis*) - Biodiversity of Great Smoky Mountains National Park, Discover Life in America, Gatlinburg, Tennessee 37738.

Martin WG (1987) Savage gulf ecosystem project. Tennessee Department of Conservation, Nashville.

McEwan RW, Mueller RN, Arthur MA, Housman HH (2000) Temporal and ecological patterns of flowering dogwood mortality in the mixed mesophytic forest of eastern Kentucky. J Torrey Bot Soc 127:221-229.

Merkle SA, Andrade GM, Nairn CJ, Powell WA, Maynard CA (2007) Restoration of threatened species: a noble cause for transgenic trees. Tree Genet Genomes 3:111-118.

Myers BR, Walck JL, Blum KE (2004) Vegetation change in a former chestnut stand on the Cumberland Plateau of Tennessee during an 80-year Period (1921–2000). Castanea 69:81–91.

Neff BD, Gross MR (2001) Microsatellite evolution in vertebrates: Inference from AC dinucleotide repeats. Evolution 55:1717–1733.

Nei M (1975) *Molecular population genetics and evolution*. North-Holland Publishing, Amsterdam, The Netherlands.

Nei M (1987) *Molecular evolutionary genetics*. Columbia University Press, New York.

Pierce AR, Bromer WR, Rabenold KN (2008) Decline of *Cornus florida* and forest succession in a *Quercus–Carya* forest. Plant Ecol 195**:**45–53.

Pimentel D (1997) Techniques for Reducing Pesticides: Environmental and Economic Benefits. John Wiley and Sons, Chichester, UK.

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecol Econ 52:273-288.

Pritchard JK, Donnelly P (2001) Case–control studies of association in structured or admixed populations. Theor Popul Biol 60:227–237.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945-959.

Quarterman E, Turner BH, Hemmerly TE (1972) Analysis of virgin mixed mesophytic forests in Savage Gulf, Tennessee. Bull Torrey Bot Club 99:228-232.

Redlin SC (1991) *Discula destructiva* sp. nov., cause of dogwood anthracnose. Mycologia 83:633-642.

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. Conserv Biol 17:230–237.

Reed S (2004) Self-incompatibility in *Cornus florida* L. HortScience 39:335-338.

Rossell IM, Rossell CR, Hining KJ, Anderson RL (2001) Impacts of dogwood anthracnose (*Discula destructiva* Redlin) on the fruits of flowering dogwood (*Cornus florida* L.): implications for wildlife. Am Midl Nat 146:379-387.

Rousset F (2000) Genetic differentiation between individuals. J Evol Biol 13:58-62.

Severns P (2003) Inbreeding and small population size reduce seed set in a threatened and fragmented plant species, *Lupinus sulphureus* ssp. kincaidii (Fabaceae). Biol Conserv 110:221–229.

Sherald JL, Stidham TM, Hadidian JM, Hoeldtke JE (1996) Progression of the dogwood anthracnose epidemic and the status of flowering dogwood in Catoctin Mountain Park. Plant Dis 80:310-312.

Soltis DE, Morris AB, McLachlan JS, Manos PS, Soltis PS (2006) Comparative phylogeography of unglaciated eastern North America. Mol Ecol 15:4261-4293.

Sork VL, Smouse PE, Apsit VJ, Dyer RJ, Westfall RD (2005) A two-generation analysis of pollen pool genetic structure in flowering dogwood, *Cornus florida* (Cornaceae), in the Missouri Ozarks. Am J Bot 92:262-271.

Stiles EW (1980) Patterns of fruit presentation and seed dispersal in bird-disseminated woody plants in the eastern deciduous forest. Am Natur 116:670-688.

USBC (2001) Statistical Abstract of the United States 2001. Washington, DC: U.S. Bureau of the Census, U.S. Government Printing Office.

USDA, NRCS (2008) The PLANTS Database (http://plants.usda.gov) National Plant Data Center, Baton Rouge, LA 70874-4490 USA.

Wadl PA, Skinner JA, Dunlap JR, Reed SM, Rinehart TA, Pantalone VR, Trigiano RN (2009) Honeybee-mediated controlled pollinations in *Cornus florida* and *C. kousa* intra- and interspecific crosses. HortScience 44:1527-1533.

Wang X, Rinehart TA, Wadl PA, Spiers JM, Hadziabdic D, Windham MT, Trigiano RN (2009) A new electrophoresis technique to separate microsatellite alleles. Afr J Biotechnol 8:2432-2436.

Wang XW, Trigiano RN, Windham MT, Scheffler BE, Rinehart TA, Spiers JA (2008) Development and characterization of simple sequence repeats for flowering dogwood (*Cornus florida* L.). Tree Genet Genomes 4:461-468.

Wilds SP (1997) Gradient analysis of the distribution of a fungal disease of *Cornus florida* in the southern Appalachian Mountains, Tennessee. J Veg Sci 8:811–818.

Williams CE, Moriarity WJ (1999) Occurrence of flowering dogwood (*Cornus florida* L.), and mortality by dogwood anthracnose (*Discula destructiva*

Redlin), on the Northern Allegheny Plateau. J Torrey Bot Soc 126:313-319.

Windham MT, Graham ET, Witte WT, Knighten JL, Trigiano RN (1998) *Cornus florida* 'Appalachian Spring': A white flowering dogwood resistant to dogwood anthracnose. HortScience 33:1265-1267.

Windham WT, Montgomery-Dee M, Parham JM (1995) Factors associated with dogwood anthracnose incidence and severity. Tennessee AgriScience 175:17-20.

Witte WT, Windham MT, Windham AS, Hale FA, Fare DC, Clatterbuck WK (2000) Dogwoods for American gardens. The University of Tennessee Agriculture Extension Service, Knoxville. PB1670.

Wright S (1931) Evolution in Mendelian populations. Genetics 16:97-159.

Wright S (1943) Isolation by distance. Genetics 28:114-138.

Wright S (1946) Isolation by distance under diverse systems of mating. Genetics 31:39-59.

Appendix 1: Tables

Table 2-1. *Cornus florida* sampling locations throughout eastern United States with GPS coordinates and county location.

Sample	Subsample	Latitude	Longitude	County	State
TN 1	(1-5)	35.00466667	-85.308283	Hamilton	Tennessee
TN 2	(1-3)	36.01073333	-87.379250	Dickson	Tennessee
TN 3	(1-5)	36.01456667	-85.133200	Cumberland	Tennessee
TN 4	(1-4)	35.88146667	-84.810217	Cumberland	Tennessee
TN10	(1-5)	35.79283333	-85.637317	Warren	Tennessee
TN 20	(1-6)	36.57225	-82.388350	Sullivan	Tennessee
TN 30	(1-3)	35.62645	-88.942850	Madison	Tennessee
TN 41	(1-5)	35.94411667	-83.202433	Cocke	Tennessee
AL 1	(1-5)	34.7049	-85.555733	DeKalb	Alabama
AL 2	(2-5)	33.86745	-86.269583	St. Clair	Alabama
AL 3	(1-5)	33.18536667	-87.321983	Tuscaloosa	Alabama
AL 4	(1-3)	33.43668333	-88.201617	Pickens	Alabama
AL 5	(1-5)	30.68951667	-88.171800	Mobile	Alabama
MS 1	(1-5)	33.48821667	-88.907850	Oktibbeha	Mississippi
MS 2	(1-5)	33.37121667	-89.139700	Choctaw	Mississippi
MS 3	(1-5)	33.47568333	-89.687233	Montgomery	Mississippi
MS 4	(1-5)	32.30078333	-90.801500	Warren	Mississippi
MS 5	(1-5)	31.99165	-90.356533	Copiah	Mississippi
MS 6	(1-5)	30.78618333	-89.501300	Pearl River	Mississippi
NC 1	(1-6)	35.70246667	-83.041250	Haywood	N. Carolina
NC 2	(1-5)	35.42521667	-82.631083	Henderson	N. Carolina
NC 3	(1-5)	35.00263333	-84.083733	Cherokee	N. Carolina
LA 1	(1-5)	30.78916667	-89.839017	Washington	Louisiana
LA 2	1	30.40905	-89.785017	St.Tammany	Louisiana
FL 1	(1-4)	30.66846667	-86.818350	Santa Rosa	Florida

Table 2-1. Continued.

Sample	Subsample	Latitude	Longitude	County	State
FL 2	(1-5)	30.74181667	-86.272833	Walton	Florida
FL 3	(1-5)	30.7533	-85.317633	Jackson	Florida
GA 1	(1-5)	30.73655	-84.106050	Grady	Georgia
GA 2	(1-5)	31.74313333	-83.656667	Turner	Georgia
GA 3	(1-5)	32.85881667	-83.752517	Bibb	Georgia
GA 6	(1-5)	34.86956667	-83.518083	Rabun	Georgia
KY 1	(1-5)	36.65253333	-84.184100	Whitley	Kentucky
KY 2	(1-6)	36.995000	-86.445861	Warren	Kentucky
KY 2	(2-1)	37.160000	-86.198333	Edmonson	Kentucky
KY 2	(2-2)	37.160000	-86.198333	Edmonson	Kentucky
KY 2	(2-3)	36.858056	-86.036944	Barren	Kentucky
KY 2	(2-4)	36.629722	-85.682500	Monroe	Kentucky
KY 2	(2-5)	36.840833	-86.354167	Warren	Kentucky
KY 2	(2-6)	38.309444	-85.488333	Oldham	Kentucky
OH 1	(1-7)	39.22395	-84.826117	Hamilton	Ohio
OH 2	(1-5)	39.05558333	-84.105350	Clermont	Ohio
OH 3	(1-5)	39.04873333	-83.230433	Pike	Ohio
OH 4	(1-5)	39.35825	-81.983350	Athens	Ohio
PA 1	(1-4)	40.55972222	-76.364722	Schuylkill	Pennsylvania
PA 2	(1-3)	40.54416667	-76.326111	Schuylkill	Pennsylvania
SC 1	(1-5)	35.16558333	-82.132833	Spartanburg	S. Carolina
SC 2	(1-5)	34.60138333	-82.653633	Anderson	S. Carolina
SC 3	(1-5)	34.68166667	-83.145167	Oconee	S. Carolina
WV 1	(1-4)	38.25176667	-81.427017	Kanawha	W. Virginia
WV 2	(1-5)	38.1525	-81.127150	Fayette	W. Virginia
WV 3	(1-7)	37.49175	-81.102033	Mercer	W. Virginia
TX 1	1	30.559444	-96.629722	Burleson	Texas
TX 3	1	30.559472	-96.408056	Brazos	Texas

Table 2-1. Continued.

Sample	Subsample	Latitude	Longitude	County	State
TX 4	1	30.559639	-96.409389	Brazos	Texas
TX 5	1	30.560056	-96.409639	Brazos	Texas
TX 6	1	30.560306	-96.409667	Brazos	Texas
TX 7	1	30.559250	-96.409861	Brazos	Texas
TX 8	1	30.559167	-96.409778	Brazos	Texas
TX 9	1	30.558306	-96.410278	Brazos	Texas
TX10	1	30.558139	-96.410500	Brazos	Texas
TX 11	1	30.557333	-96.410833	Brazos	Texas
TX12	1	30.560444	-96.407472	Brazos	Texas
TX 13	1	30.560528	-96.407417	Brazos	Texas

Table 2-2. Primer information for 18 microsatellite loci used to analyze 234 *Cornus florida* samples collected from 46 sampling localities in the eastern United States.

Locus	Primer sequences (5'−3')	GenBank accesion number	Expected size (bp)
C020	F: TATGGCTTGCTTTGGCTAATTGTT R: CCAACTTATGCACACAGTGACACA	ED651708	144
C025	F: ATTATTTTTCGGGACCATTGCAT R: CGGCTAAGAAAATATGTCCCTTTGT	ED651711	158
C046	F: TGCAATTGACATTTGTTGCATTTT R: CAACTACGTGATCAGGTGGACAAC	ED651730	151
C072	F: AAACACCCCACTGCCATATAATGT R: CCATATGGGACCCAGTGATTTTA	ED651755	132
C091	F: GCACTAGTATAAAAACATACGCGCC R: TGCCTACATGGTAGCATTCCTTTT	ED651772	143
C100	F: CCAAAGGCTATGAATTTACGATGC R: GATGCAGTCTTAATCCTCAACGTG	ED651777	112
C110	F: AGGCTGCGTTGATAAATGATATGC R: TTGACTTGACTTGAGGCAAAAACA	ED651786	98
C121	F: CCCGAAAATCAAATATGGTAAATAAGTG R: CTCTCTTAAACCACCTTTTGATGTGA	ED651797	154
C203	F: TCCTACGCACCCACCCTTAT R: ACGGTGGTGCCATTCTTTCT	ED651865	135
C209	F: TACAAAAGTGCAATGCCAATACC R: ACTCCAAAACTTCATTCCTGAGC	ED651871	197
C214	F: TGCAAATGGTTATTGATTGCTCTC R: ATTTGTTTCCCATGACCTGAAAGA	ED651874	129
C261	F: GGACATGTGGGCAGTTTGAT R: GCAATGAGATCCCTCCCCTA	ED651912	143
C262	F: ATGACAAATGGCAAGTTGATGTTG R: CAACTGTGACTCAATTCATGTCTGC	ED651913	197
C267	F: TTGGGTAGGTGAGGGAATTTAGGT R: TGCTATTACACATTTACCCCTGCC	ED651915	95
C290	F: GGTGGATCAAAACCCACAAATAAG R: CCTATAGCCCAGCGAGAGTCTTTA	ED651936	153
C398	F: CCTCGCTTAATTCATATACTTCTT R: AATAAAAGTTTGTATGAACTGCAC	ER870420	156

Table 2-2 Continued.

Locus	Primer sequences (5'−3')	GenBank accesion number	Expected size (bp)
C488	F: CTCAAAATCACCCTTCTTATATCT R: GATAGTTCACAAGCAATGTTACTG	ER870510	114
C597	F: AAGTCAGATCATTTCAGATTAACA R: CGAATTGACGATAAATACAAAATA	ER870619	107

Table 2-3 (A-B). Estimates of Analysis of Molecular Variance (AMOVA) from dogwood microsatellite data using Arlequin (ver 3.1). Two analyses were conducted: the first included all sites as one hierarchical group (A); the second analysis accounted for sampling sites grouped according to regions identified by the program STRUCTURE (TX was excluded from this analysis since it did not contain any subpopulations) (B).

A.

Variance partition	d.f.	Sum of squares	Variance component	% of variation	P value
Among sites	45	334.285	0.49037 Va	15.76	<0.0001
Within sites	406	1064.544	2.62203 Vb	84.24	<0.0001
Total	451	1398.830	3.11240		
Fixation Index Fst:	0.15755				

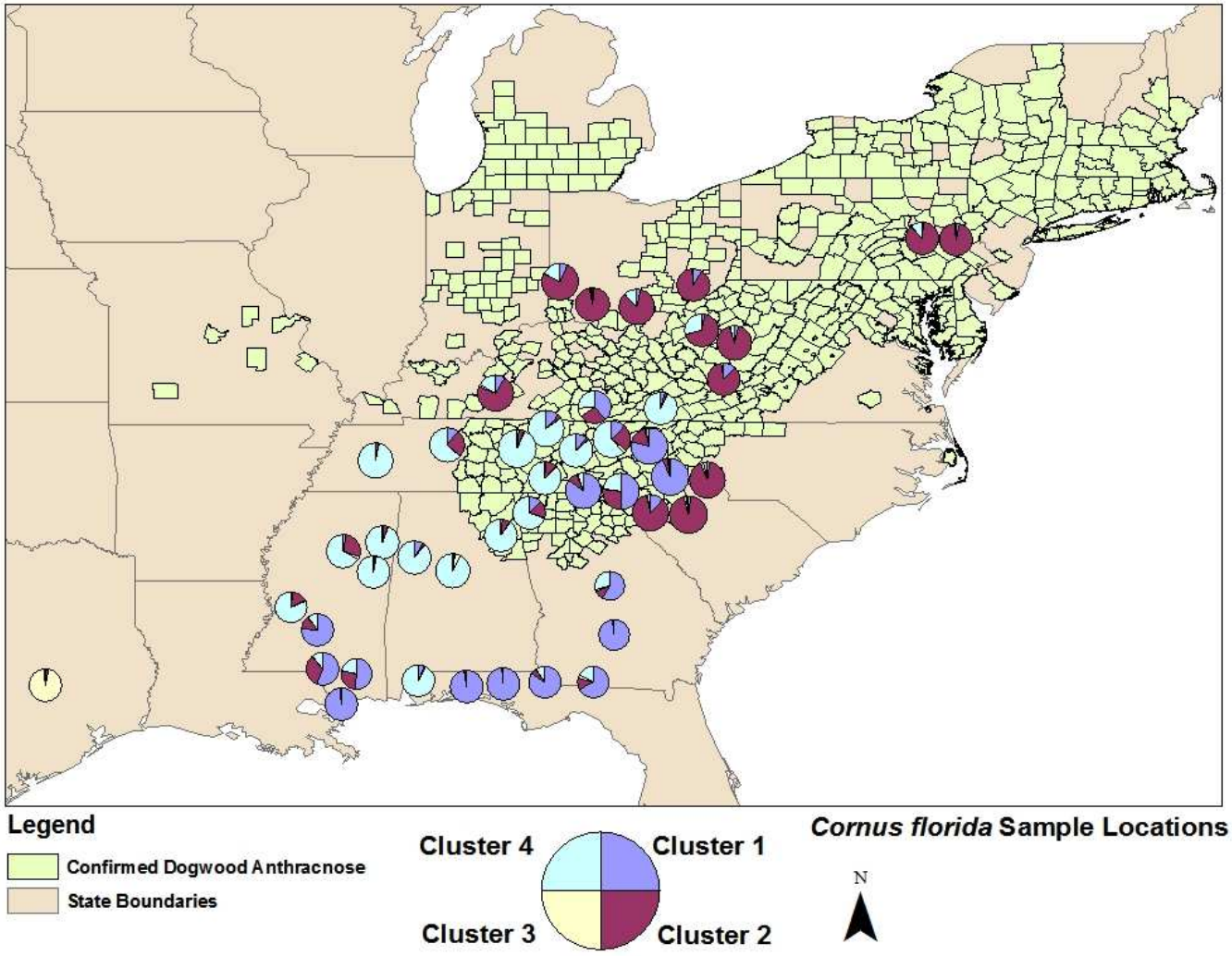
Table 2-3. Continued.

B.

Variance partition	d.f.	Sum of squares	Variance component	% of variation	P value
Among groups	2	31.417	0.6318 Va	2.03	<0.0001
Among sites within groups	42	276.136	0.27951 Vb	8.96	<0.0001
Among individuals within sites	168	661.890	1.16357 Vc	37.31	<0.0001
Within individuals	213	343.500	1.61268 Vd	51.71	<0.0001
Total	427	2911.923	6.91574		
Fixation Indices:	Fis: 0.41912	Fsc: 0.09147	Fct: 0.02026	Fit: 0.48294	

Appendix 2: Figures

Figure 2-1**.** Sampling localities and distribution of *Cornus florida* throughout eastern United States (total of 234 individuals). Counties shaded in beige had confirmed cases of dogwood anthracnose as of 2006. Pie charts at each locality represent average inferred ancestry of individuals for each cluster estimated by STRUCTURE. Dogwood anthracnose data courtesy of United States Department of Agriculture (USDA) - Forest Service-Forest Health Protection; distribution map courtesy of United States Geological Survey (USGS).


Legend
Confirmed Dogwood Anthracnose
State Boundaries
Cluster 4
Cluster 1
Cluster 3
Cluster 2
Cornus florida Sample Locations
N

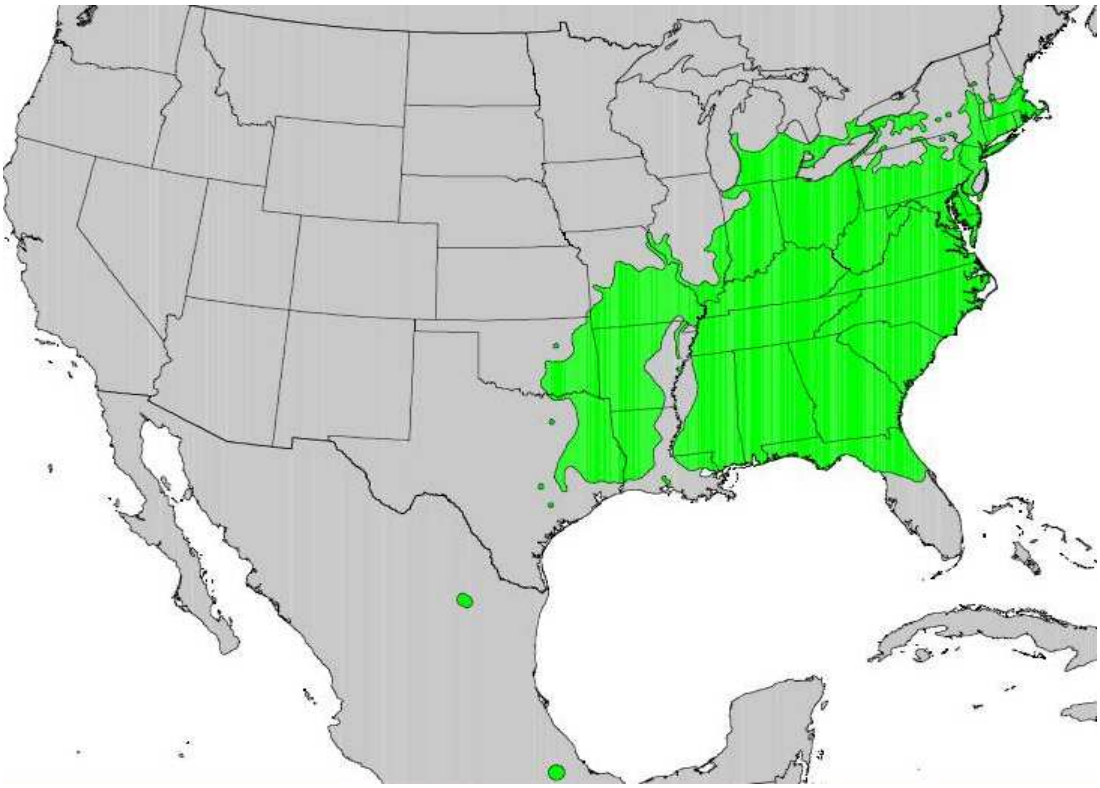


Figure 2-2. Distribution map of *Cornus florida* in the eastern United States. Image courtesy of United States Geological Survey (USGS).

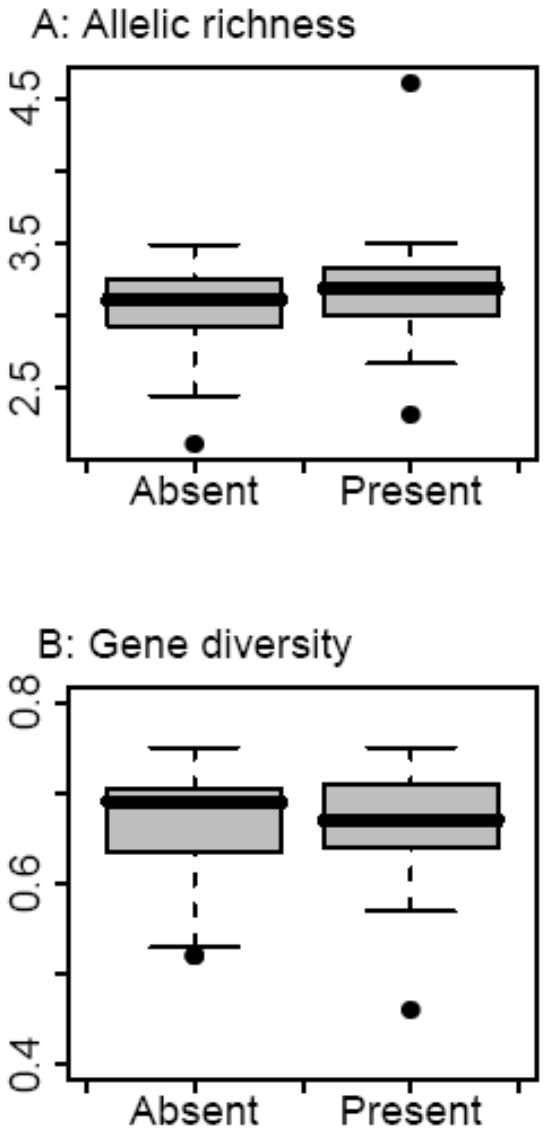
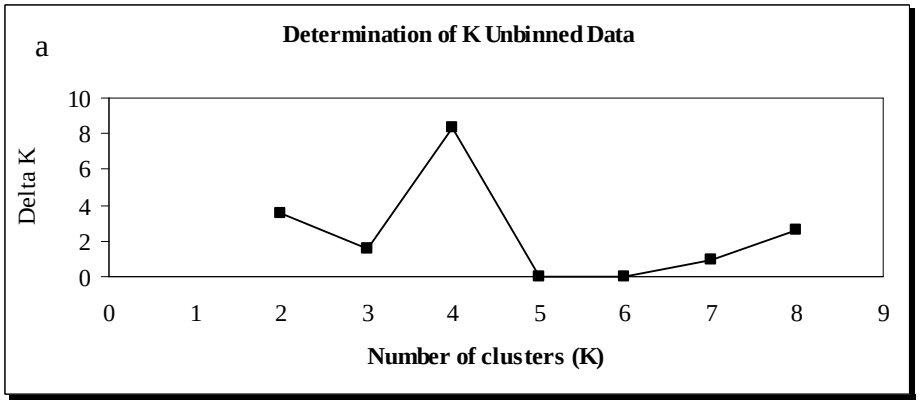
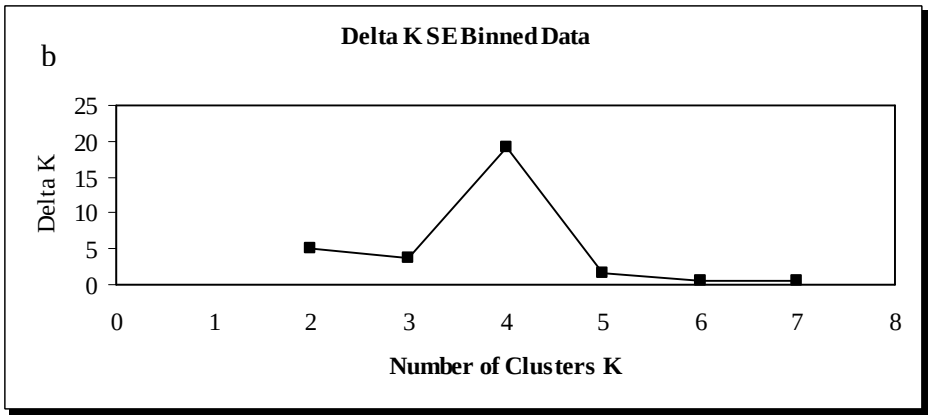
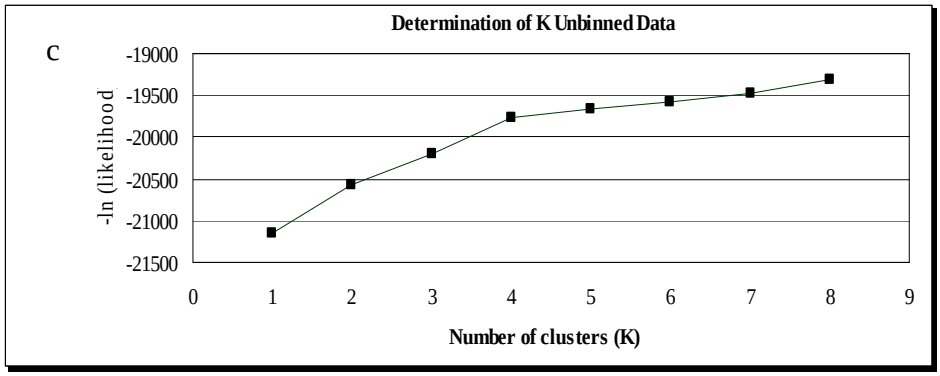


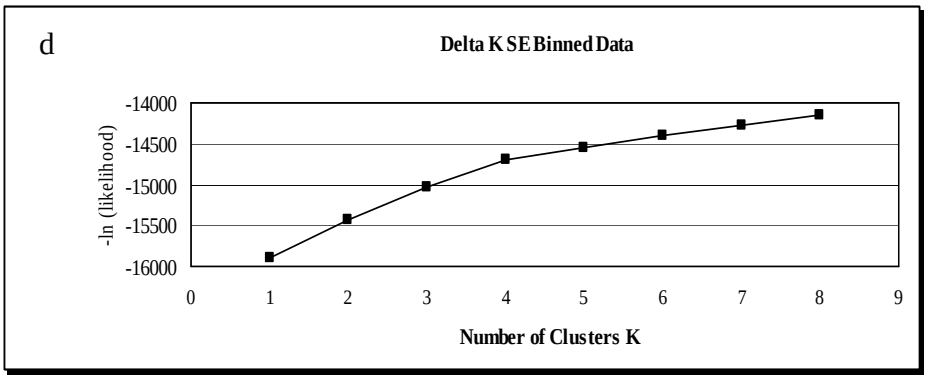
Figure 2-3. Tukey boxplots representing no association between dogwood anthracnose outbreak and allelic richness (A) and gene diversity (heterozygosity) (B) within localities. The box represents 50% of ordered data stretching between the lower and the upper quartiles. The bar in the box indicates the median, and the whiskers extending from the boxes include all data within 1.5 $\times$ the interquartile range. Extreme outliers are shown as dots. The notches illustrate approximate 95% confidence intervals for the medians. Since these notches overlap broadly, there is no evidence that the two medians differ significantly.

Figure 2-4. Delta K (Δk) calculations for unbinned (a and c) and binned (b and d) data. Determination of the number of clusters (k) vs. the second order rate of change in k (Δk) (upper 2 graphs). The clear maximum for Δk at k=4 indicates that four clusters fit the microsatellite data best for *Cornus florida* populations. The number of clusters (k) vs. the negative natural log of the likelihood [-ln (likelihood)] of the microsatellite data (lower 2 graphs).


a
Determination of K Unbinned Data
Delta K
10
8
6
4
2
0
0 1 2 3 4 5 6 7 8 9
Number of clusters (K)


b
Delta K SE Binned Data
Delta K
25
20
15
10
5
0
0 1 2 3 4 5 6 7 8
Number of Clusters K


c
Determination of K Unbinned Data
-ln (likelihood
-19000
-19500
-20000
-20500
-21000
-21500
0 1 2 3 4 5 6 7 8 9
Number of clusters (K)


d
Delta K SE Binned Data
-ln (likelihood)
-14000
-14500
-15000
-15500
-16000
0 1 2 3 4 5 6 7 8 9
Number of Clusters K

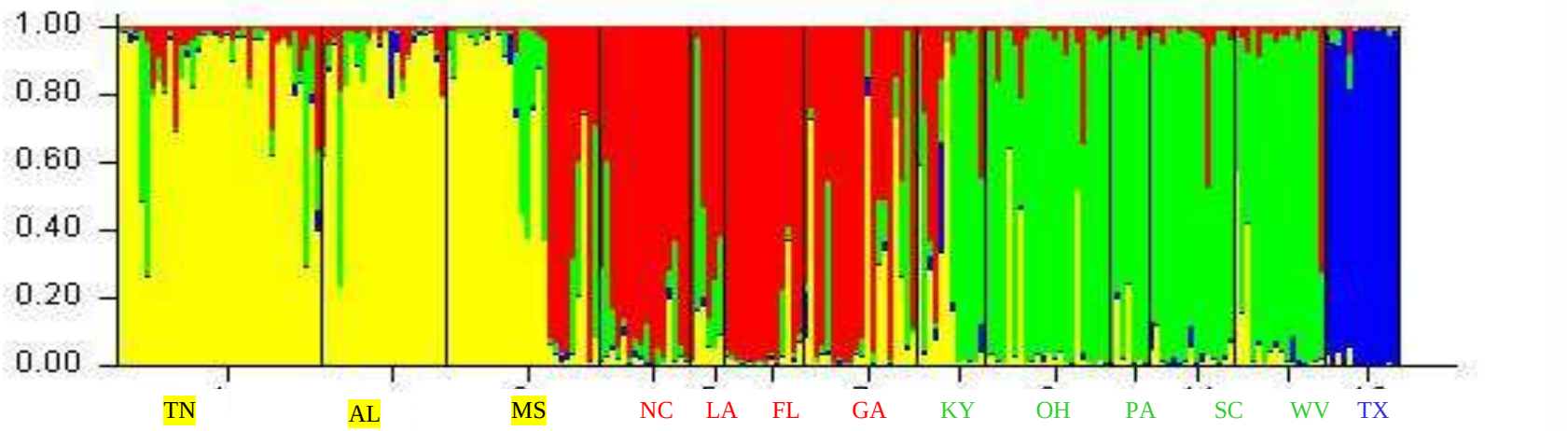


Figure 2-5. Bar plots showing Bayesian assignment probabilities using software STRUCTURE 2.2 for four clusters (K=4) (TN-Tennessee, AL-Alabama, MS-Mississippi, NC-North Carolina, LA-Louisiana, FL-Florida, GA-Georgia, KY-Kentucky, OH-Ohio, PA-Pennsylvania, SC-South Carolina, WV-West Virginia, TX-Texas) divided according to geographical location. The proportion of each bar that is yellow, red, green and blue indicates the assignment probability of individuals to each of these clusters, respectively.

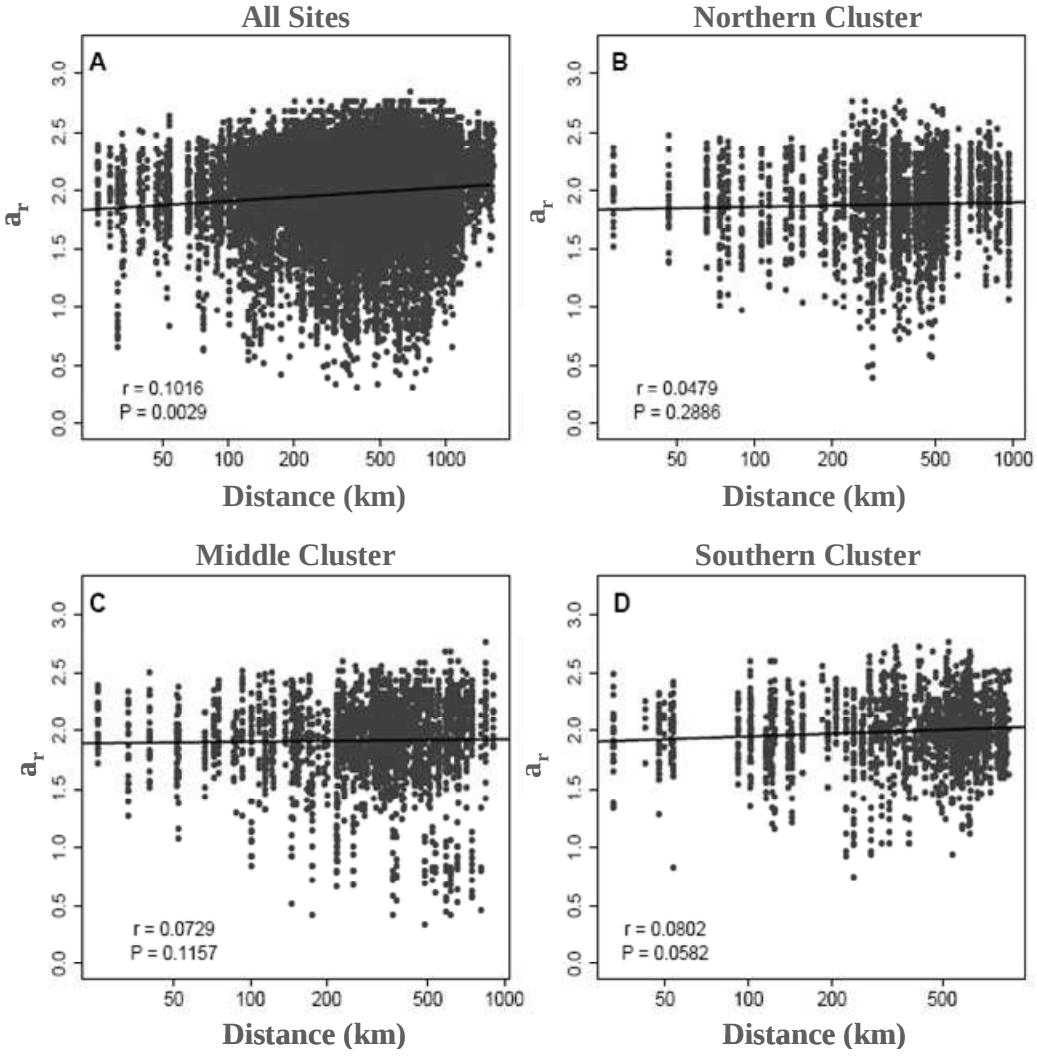


Figure 2-6 (A-D). Scatterplots of pairwise genetic distances (a_r) vs. geographical distances (km) of all sampled *Cornus florida* individuals (A), and partitioned into sites with predominant membership in the northern cluster (B), middle cluster (C), and southern cluster (D). *P*-values are from Mantel tests with 10,000 randomizations. Values on the x-axis are in log scale. Lines indicate the best fit of least-squares regression of a_r on log distance.

Chapter 3. Fine scale genetic structure of flowering dogwood in the Great Smoky Mountains National Park.

Abstract

In the past three decades, flowering dogwood (*Cornus florida*) populations have experienced severe declines caused by dogwood anthracnose. Mortality has ranged from 48 to 98%, raising the concern that the genetic diversity of this native tree has been reduced significantly. However, the response of each species to ecological disturbance may differ greatly depending on their biological attributes, particularly pollen and seed dispersal ability. Nineteen microsatellite loci were used to evaluate the level and distribution of genetic variation throughout the Great Smoky Mountains National Park (GSMNP). Significant genetic structure exists at both landscape and local levels. Two genetic clusters exist within the park separated by the main dividing ridges of the Great Smoky Mountains. The differentiation of the clusters is statistically significant, but subtle, with gene flow evident through low-elevation corridors. It seems unlikely that recent demographic dynamics have resulted in a depletion of genetic variation in flowering dogwoods.

Introduction

The levels of genetic diversity reflect genetic resources necessary for both short-term ecological adaptation and long-term evolutionary change. In the short term, genetic variability is a less critical factor in the survivability of the species and overall population persistence. However, long-term preservation requires understanding demography and genetics of small populations, as well as ecology and evolution of abundant taxa (Lande and Shannon 1996). Environmental

changes and factors influencing species abundance or rarity are often unpredictable. Because of those evolutionary oscillations, it is crucial to secure adequate genetic diversity that permit species to continuously evolve in response to those environmental fluctuations and changes in selection pressures. Preservation of biodiversity and conservation efforts has largely focused on endangered, threatened or relatively small populations in danger of extinction. However, decrease or complete loss of existing populations worldwide is mainly due to anthropogenic habitat destruction and the introduction of exotic species of parasites, predators and competitors (Lande and Shannon 1996). In the United States alone, the projected economic loss due to invasion of non-native species is over $120 billion annually and about one-fourth of this loss is caused by exotic plant pathogens (Pimentel 1997; USBC 2001; Pimentel et al. 2005).

Discula destructiva Redlin (Redlin 1991) is thought to be an introduced plant pathogen (Trigiano et al. 1995) and has caused severe decline and reduced reproduction of native flowering dogwood populations in the past three decades (Figure 3-1, in appendix). Infection by *D*. *destructiva,* causal agent of dogwood anthracnose, resulted in poor seed production due to decreased flowering (Rossell et al. 2001) and smaller, younger trees were more susceptible to this disease (Hiers and Evans 1997; Holzmueller et al. 2006; Jenkins and White 2002). Mortality of flowering dogwood has ranged from 48 to 98% in the northeast and Appalachian highlands (Sherald et al. 1996; Hiers and Evans 1997; Williams and Moriarity 1999; McEwan et al. 2000; Jenkins and White 2002) (Figure 3-1, in appendix) because of this disease. Previous research evaluated

native populations of flowering dogwoods for reduced genetic diversity and fragmentation, which was presumably due anthracnose disease (Hadziabdic et al. 2010, accepted for publication). Although four widespread, spatially contiguous clusters were detected in the native range of flowering dogwood distribution, this study found little relationship between geographic and genetic difference. These findings suggest that high dispersal rates and large effective population sizes have so far prevented loss of genetic diversity. Additionally, genetic variation in areas affected by dogwood anthracnose was as high as or higher than areas without large die-offs.

In smaller, fragmented populations, loss of genetic variation through genetic drift may cause reduced mean population fitness, elevating the risk of extinction and reducing prospects for adaptive change (Fischer and Matthies 1998; Reed and Frankham 2003; Severns 2003). Fragmentation can considerably reduce heterozygosity, cause limited gene flow among populations, and therefore reduce effective population size by separating larger plant populations that were historically linked via suitable habitat (Severns 2003). Genetic variation in plant species with continuous distributions located at different geographical regions may be affected by physical barriers such as a mountain range or historical events like glaciations that limits gene flow among populations (Persson et al. 2004; Mandek et al. 2005). In these small isolated populations, seed production, which affects species persistence, can be significantly decreased owing to inbreeding depression (Fischer and Matthies 1998).

Population fragmentation from disease pressure or physical barriers may play a significant role in flowering dogwood genetic diversity since seed dispersal is thought to occur mainly by neo-tropical migratory birds that use the fruits seeds as a fall food source (Stiles 1980). The birds ingest fruits that are either disgorged or voided after relatively short time ingestion (Stiles 1980). Fruits of flowering dogwoods contain one of the highest fat content (approximately 18%) of any food available in the forest and also contain high levels of available calcium (Halls and Epps 1969). Individual dogwood trees are capable of producing up to 10 kg of high protein berries each fall, supplying a valuable food source for more than 50 bird and numerous small game species (Rossell et al. 2001; Holzmueller et al. 2006). Understanding fine-scale genetic diversity of dogwood within the Great Smoky Mountains National Park will further our understanding of how seed dispersal affects population genetics.

The flowering dogwood mating system plays an important role in the genetic composition of populations affecting frequency of individual genotypes in subsequent generations and therefore influencing the degree of genetic variation within and among populations (Hamrick and Godt 1989; Karasawa et al. 2007). Mating system of the species is one of the fundamental parameters affecting colonizing success. Flowering dogwood has an obligate outcrossing mating system with self-incompatibility (Reed 2004). Consequently, they are generally not capable of breeding with related individuals (such as half-sibs or full-sibs), implying nonrandom mating and therefore deviation from Hardy-Weinberg equilibrium expectations.

Successful breeding systems involve several evolutionary factors such as size and population density, movement and selection of pollinators, number of flowers and presence of self-incompatibility systems for newly established colonizing species (Karasawa et al. 2007; Murawasky and Hamrick 1991; Franceschinelli and Bawa 2000). Greater population density increases the effective number of reproducing individuals within the genetic neighborhood, effectively increasing outcrossing rate (Franceschinelli and Bawa 2000). High plant densities can be found in newly colonizing species capable of fast establishment and quick dispersal of seeds. As long as numbers of seed parents are sufficiently large for reproduction purposes, landscape disturbances may not negatively affect the effective number of pollen parents and, hence, subsequent genetic diversity in flowering dogwood (Sork et al. 2005).

Colonization of new or expansion and domination of already existing populations is an evolutionary change that results in increased spatial heterogeneity and community complexity (Margalef 1962; Platt 1975) in these disturbed habitats. Gradual expansion and satellite population scenarios are two models explaining genetic variation in colonized areas (Mandek et al. 2005; Perrson 2004). Under the first model, migration is initiated from a gradually expanding continuous front with an assumption that genetic diversity will be maintained as the expansion continues. In contrast to gradual expansion, the satellite population model is focused on the number of spatially isolated and initially small and marginal populations, which eventually become a source of new founding events (Mandek et al. 2005), and represent only a portion of the

original population's genetic variation. In the case of plant invasions, historical studies reveal differing levels of genetic variation within colonizing species, due to founder effects, number of population bottlenecks, mating systems or ability of colonizing plants to hybridize (Barrett 1982; Mandek et al. 2005). In flowering plants, genetic diversity is strongly affected by interactions between reproductive systems and stochastic forces associated with immigration history and range expansion (Barrett et al. 2008). Hence, maintaining high levels of gene flow results in genetic cohesion among populations, whereas disrupting that flow produces genetic differentiation through genetic drift and local selection (Bittencourt and Sebbenn 2007).

Although both seeds and pollen contribute to gene flow between existing populations of forest trees (Godoy and Jordano 2001; Smouse and Sork 2004), only seed dispersal can result in colonization of new patches. Therefore selection pressures acting on seed dispersal may differ considerably from those affecting pollen dispersal. Pollen dispersal determines the pattern of mating, whereas seeds disperse after reproduction. Furthermore, male gametes are dispersed from the paternal to the maternal parent via pollen (Bittencourt and Sebbenn 2007) and embryos containing genetic material from both parents are dispersed in the form of seeds (Hamrick et al. 1993). As a result, both seeds and pollen are expected to respond differently to genetic population structure and inbreeding load (a decrease in fitness of inbred relative to outbred crosses) (Ravigne et al. 2006).

Kin-structured migration (via kin-structured seed dispersal) appears commonly in nature. In many plant species, multi-seeded fruits are a unit of dispersal in which larger proportions of the seeds may have closely related (or identical) fathers, as well as sharing the same mother plant (Torimaru et al. 2007). In angiosperms, movement of genes is in the form of full or half-sib seeds.

In disturbed habitats in which mortality rate of native populations result in near elimination of the existing flowering dogwoods, forest fragmentation may affect ecological determinants of reproductive parameters (seeds and pollen dispersal). The overall result is limited gene flow, inbreeding depression and/or genetic erosion at the reproductive stage. Extinction and recolonization after local disturbances could result in sampling from the available gene pool (founder effect) or in additional gene flow. As a result of dogwood anthracnose, flowering dogwood seedlings and saplings were reported to be absent in several studies (Sherald et al. 1996; Hiers and Evans 1997; Holzmueller et al. 2006; Jenkins and White 2002; Rossell et al. 2001) resulting in severe decline of native populations. With such large-scale mortality, genetic variation of species can decline rapidly (Hawley et al. 2006) having potentially deleterious effects on the ability of tree populations to regenerate and respond to future changes in selection pressures. Consequently, the relationships between plant population size, fitness, mating systems and genetic diversity are of fundamental importance in plant ecology, evolution and conservation (Leimu et al. 2006).

Both widespread and localized landscape disturbances because of anthropogenic expansion, invasion of non-native species or severe disease

outbreaks present an opportunity to understand effects of these disturbances on genetic structure and variability within and among populations. In this study, microsatellite loci were used to evaluate genetic differentiation and population structure of flowering dogwoods in the GSMNP. By assessing genetic diversity and population dynamics of this native tree, we aimed to answer the following questions: (1) How much genetic diversity exists within and among populations of flowering dogwoods? (2) Is there an evidence of population structure? (3) Is there an effect of seed vs. pollen dispersal on genetic structure? (4) Can a baseline be determined for understanding future population changes in newly colonized areas?

Materials and Methods

Study system

Cornus florida L. (Flowering dogwood)

Flowering dogwood is a small, shade tolerant tree indigenous to the eastern United States (Figure 2-2, in appendix). Flowering dogwood is an insect-pollinated, self-incompatible understory species found in temperate deciduous forests, such as oak (*Quercus*), yellow poplar (*Liriodendron*) and pine (*Pinus*). The floral display starts in the early spring, followed by attractive summer green foliage. In the late summer, bright red oblong drupes are accompanied by early development of dark, red brown foliage (Witte et al. 2000). These features make flowering dogwood trees important components of the forests they inhabit not

only for their ornamental value during all seasons, but also as a significant source of food for wildlife.

The flowers, leaves, bark, and fruits provide nourishment for a variety of vertebrates, including American beavers (*Castor canadensis*), which feed largely on the bark and outer layers (Linzey and Brecht 2003a); Eastern gray squirrels (*Sciurus carolinensis*), which prefer dogwood berries (Linzey and Brecht 2003b); and white-tailed deer (*Odocoileus virginianus*) that feed on dogwood twigs (Blair, 1982; Holzmueller et al. 2006). Flowering dogwoods are also important in rapid nutrient cycling (64% reduction in litter mass in two years) within forest communities, thus providing high concentrations (2-3.5% dry weight) of available calcium (Ca) (Jenkins and White 2002). Due to high calcium concentration in its foliage, quick decomposition, and abundance in the understory, flowering dogwoods influenced calcium availability in the mineral soil and forest floor by acting as a “Ca pump” in forests (Holzmueller 2006; Jenkins et al. 2006; Thomas 1969). Calcium, potassium and magnesium saturation in soils were positively correlated with the presence of flowering dogwood density in oak-hickory forests (Holzmueller 2006). Seedlings grown in soil with lower availability of calcium and potassium cations exhibited higher dogwood anthracnose severity earlier in the growing season when compared to seedlings grown in soils with greater inputs of calcium and potassium (Holzmueller 2006).

***Discula destructiva* Redlin**

The decline of native dogwood populations due to dogwood anthracnose (caused by *Discula destructiva* Redlin) (Figure 3-1, in appendix) was consistent with introduced diseases and pests that have caused tremendous mortality of various tree species over wide geographic areas (Merkle et al. 2007; Crowl et al. 2008). These die-offs often result in restructuring of forest communities and disruption of ecosystem processes (Hall et al. 2002; Becker et al. 2008; Elliott and Swank 2008; Eschtruth and Battles 2008) and some examples include near elimination of American chestnuts [*Castanea dentata* (Marsh.) Borkh] by chestnut blight [*Cryphonectria parastica*, (Murrill) Barr] (Elliott and Swank 2008); American elms (*Ulmus americana* L.) due to Dutch elm disease [*Ophiostoma ulmi* (Buisman) Nannf] (Brasier 2000); and Eastern hemlocks [*Tsuga canadensis* (L.) Carrière] by hemlock wooly adelgids [*Adelges tsugae* Annand] (Havill et al. 2006).

Dogwood anthracnose is manifested by purple-rimmed lesions that initially appear on the leaf margins and extend through the veins into the petiole in trees (Jenkins and White 2002). The disease usually starts in the lower canopy and moves upwards in the tree. During cool, wet weather, the first signs of dogwood anthracnose are gray or drooping foliage hanging on the twigs. With disease progression, the fungus moves from leaf petiole into stems resulting in twig and branch die-back due to the formation of annual cankers, which can eventually result in tree death. Although the spread of dogwood anthracnose southward in the United States appeared to be relatively uninhibited, several factors, including

temperature, light availability, moisture and evaporative potential of the leaf surface, can determine the probability and severity of the infection (Hibben and Daughtrey 1988; Hiers and Evans 1997). Disease symptoms are usually more apparent on slopes facing northeast (understory canopies with low evaporative potential) than in southwest facing slopes (exposed canopies) (Chellemi and Briton 1992; Chellemi et al. 1992; Hiers and Evans 1997; Jenkins and White 2002).

In the northeast and Appalachian highlands, dogwood mortality rates were as high as 98%, resulting in a severe decline of native dogwood populations (Sherald et al. 1996; Hiers and Evans 1997; Williams and Moriarity 1999; McEwan et al. 2000; Jenkins and White 2002) (Figure 3-1). In Catoctin Mountain Park in Maryland, Sherald et al. (1996) observed 77% reduction in flowering dogwood density between 1976 and 1992. Hiers and Evans (1997) compared responses of dogwood populations to the anthracnose in both mixed mesophytic forests of the Cumberland Plateau coves and the oak-hickory forests of the plateau uplands in Tennessee in the period between 1983 and 1995. Severity and incidence ratings in the study were followed by the methods of Chellemi et al. (1992) and symptoms included leaf necrosis, cankers, lower branch dieback and epicormic sprouting (Hiers and Evans 1997). They used two earlier studies for comparison to existing anthracnose outbreaks, changes in dogwood abundance and surveys of community composition (Hiers and Evans 1997; McGee 1986; Ramseur and Kelly 1981).

When compared to earlier surveys from 1976 and 1983 (McGee 1986; Ramseur and Kelly 1981), flowering dogwood populations showed precipitous decline in stem density within both examined sites, resulting in 87 and 98% mortality rates, respectively (Hiers and Evans 1997). The highest mortality occurred among smaller, younger trees. This outcome resulted in the absence of fruit and therefore seed production during the period of the study (Hiers and Evans 1997).

During the period between 1977 and 1985, 400 permanent, 0.1-hectare (20 m x 50 m) plots were established for monitoring long-term biodiversity in the GSMNP, during which time the distribution and severity of dogwood anthracnose increased. In 1992, 25% of flowering dogwoods had been destroyed by dogwood anthracnose (Windham et al. 1995). Two years later, an estimated 75% of the remaining trees had been eliminated (Windham et al. 1995). Jenkins and White (2002) collected flowering dogwood data from 86 permanent plots located in the western portion of GSMNP during the periods from 1977 to1979 and 1995 to 2000. In their study, flowering dogwood density within different forest types (typic cove, acid cove, alluvial, oak-hickory, and oak-pine stands) decreased between two sampling intervals, ranging from 69-94% (Jenkins and White 2002). The largest decline, over 94%, was among smaller, younger trees (1 to 4.9 cm diameter) in all forest types (Jenkins and White 2002).

Other factors including (but not limited to) canopy closure and environmental stress have contributed to the loss of 36% of flowering dogwoods in native habitats over the 10-year period (McEwan et al. 2000). Loss of flowering

dogwoods from eastern forests has reduced the rate of soil and forest floor calcium mineralization, which may have negative effects on many associated flora and fauna (Holzmueller 2006).

Currently, no management options exist for controlling dogwood anthracnose in large forested areas (Holzmueller et al. 2006). Fungicide applications are impractical and expensive, and breeding approaches commonly used for development of pathogen and/or pest resistant crop species are difficult to apply in large forest areas (Merkle et al. 2007). This is partially due to long juvenile periods of many forest trees including flowering dogwoods (about seven years) as well as the lack of disease resistant germplasm. To date, only one tree, 'Appalachian Spring' has confirmed resistance to the disease (Windham et al. 1998).

Field Sampling

In the spring and fall of 2007, 194 individual trees were selected for a population study in the GSMNP area (Figure 3-2, in appendix). Ninety-eight samples were randomly selected from different trail systems to the south and 96 samples north of the main dividing ridge of the Great Smoky Mountains, marked by the Tennessee-North Carolina border (Figure 3-2, in appendix). For each sampling location, Global Positioning System (GPS) coordinates were recorded (Table 3-1, in appendix) and the sampling locations were determined using the All Taxa Biodiversity Inventory (ATBI) Database Specimen Records (Ted

Simmons Bird Project) (http://www.dlia.org). Identification of individuals collected from each location was based on morphological characteristics of the species.

DNA Collection, Isolation and PCR Amplification

Genomic DNA was isolated from young leaves or unopened flower buds of flowering dogwoods collected from 194 individuals at different sampling localities throughout the GSMNP (Figure 3-2, in appendix). Collected samples were initially placed in 2 ml tubes containing 70% ethanol to preserve plant material during field sampling. Genomic DNA was extracted from samples using Qiagen DNeasy Plant Mini Kit (Qiagen, Valencia, CA, USA), according to the manufacturer's protocol. All samples were homogenized using Bio101 *FastPrep* Homogenization System FP120 (Thermo Savant, Waltham, MA, USA). The concentration and purity of DNA samples were measured with a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, Inc., Wilmington, DE, USA) before analyses. PCR amplifications were performed in 10 µl reaction mixtures containing 1 µl DNA template (4 ng/µl), 1 µl of GeneAmp 10×PCR Buffer II (Applied Biosystems, Branchburg, NJ, USA), 1 µl of 20 mM $MgCl_2$, 1 µl of each 2 mM of dNTPs, 1 µl of 2.5 µM specific microsatellite primers, 0.08 µl of 5 U AmpliTaq Gold® DNA polymerase, and 4.92 µl of sterile, nanopure water. Nineteen polymorphic microsatellite primer pairs for flowering dogwood, developed from a previous study (Wang et al. 2008) were selected and used for analyses (Table 3-2, in appendix). Amplification reactions were temperature-cycled in 96- well plates using an Eppendorf Autorisierter Thermocycler

(Eppendorf AG, Hamburg, Germany) with the following program: 94℃ for 3 min, followed by 35 cycles of 94℃ for 40 seconds, 58℃ for 40 seconds, 72℃ for 30 seconds, and a final extension at 72℃ for 4 min.

PCR products were analyzed with the QIAxcel Capillary Electrophoresis System (Qiagen, Valencia, CA, USA) using an internal 25-bp DNA size marker. Data were automatically recorded and exported using BioCalculator^{TM} software, which provided both a gel view and an electropherogram of the separated PCR products (alleles) (Wang et al. 2009). QIAxcel Capillary Electrophoresis System provides analyses with resolution of 2–4 bp compared to CEQ^{TM} 8000 (CEQ 8000: Genetic Analysis System, Beckman Coulter, USA) and ABI 3100xl DNA Sequencer (ABI 3100xl: Applied Biosystems, USA) systems that provide 1 bp resolution.

Program FLEXIBIN (Amos et al. 2007) was used for automated binning of allelic data. In order to achieve the best possible fit and to find the most suitable binning parameter, the program uses a simple algorithm to conduct searches in two different phases and steps through all possible parameter combinations. When the best fit values are established, all alleles are replaced with their repeat unit equivalents and a graphical output file with summary statistics is generated. Binned and unbinned data gave identical results and only analyses of binned data are presented (Figure 3-3, in appendix).

Data Analyses

To address the question of how much genetic diversity exists within populations of flowering dogwood affected by dogwood anthracnose, we estimated Nei's gene diversity (Nei 1987), allelic count and allelic richness (a measure of allele number independent of sample size) for all samples. Observed and expected heterozygosities within samples (either side of the main dividing ridge) were estimated using program FSTAT version 2.9.3 (Goudet 2001). Nei's gene diversity is the average probability of non-identity for two randomly chosen alleles from a randomly chosen locus (also, the average expected heterozygosity). Allelic richness was estimated following Kalinowski's rarefaction method (Kalinowski 2005) (minimum sample size of 81 trees (total of 162 genes)) using HP-Rare 1.0 software package (Kalinowski 2005). Gene diversity calculations were performed using both FSTAT version 2.9.3 and ML Relate (Kalinowski et al. 2006) (100,000 randomization performed using Monte Carlo randomization as described by Guo and Thompson (1992)).

The genetic structure of flowering dogwood populations was evaluated using two clustering methods based on Bayesian models and analysis of molecular variance (Excoffier et al. 1992). Bayesian cluster analysis was performed using STRUCTURE version 2.2 (Pritchard and Donnelly 2001). This program groups individuals on the basis of their multilocus genotypes. The method is based on the assumption that Hardy–Weinberg or linkage disequilibria are caused by population structure and attempts to find population groupings that are not in disequilibrium (Pritchard and Donnelly 2001). This is an individual-

based analysis that makes no assumptions about the relationship between sample sites and population structure. Evanno's method (Evanno et al. 2005) was used to identify the appropriate number of clusters using the *ad hoc* statistic Δk, which is based on the second order rate of change in the log probability of the data between successive values of k. An admixture model with correlated allele frequencies (assuming no prior information of population origin) was used. For assessing alternative values of k, we ran ten independent runs for each k value between one and eight for 100,000 generations and a burn-in period of 100,000 iterations.

Differentiation among groups of trees was quantified with an analysis of molecular variance (AMOVA) using Arlequin (version 3.1) (Excoffier et al. 2005) and the basic *F*-statistics to partition variation within vs. between individuals (F_{IS}) and populations (F_{ST}). Two different analyses were conducted, each with a single level of grouping. The first one partitioned individuals into groups north and south of the main dividing ridge of the Great Smoky Mountains (marked by the Appalachian Trail along the Tennessee-North Carolina border). The second one grouped sites according to majority representation in the clusters identified by STRUCTURE. AMOVA complements STRUCTURE by providing estimates of the amount of variation explained by grouping sites as implied by STRUCTURE, but should not be taken as independent confirmation of that specific hierarchical structure.

Continuous, fine-scale structure was evaluated by testing for a correlation between genetic and geographic distance. Roussett's (2000) a_R and the great-

circle distance (km) between each pair of trees in the sample were calculated and fitted linear regressions of a_R vs. the natural logarithm of distance for the entire sample, separately for the samples on either side of the dividing ridge, and separately for the main genetic clusters inferred from STRUCTURE. Computations were done using R 2.7.1 (Hornik 2008) and statistical significance of the correlation between geographic and genetic distance was evaluated using Mantel test. Given that Roussett's (2000) theoretical model does not include many features of the dispersal and colonization biology of dogwoods, and given that population genetic equilibrium is unlikely in GSMNP, estimation of demographic parameters (e.g., genetic neighborhood size, mean dispersal distance) were not attempted. Rather, this analysis was used to descriptively evaluate the kind of fine-scale, continuous genetic differentiation that is not captured by other analyses, which assume discrete demes or genetic clusters.

Results

Nineteen microsatellite loci were genotyped in each of 194 individuals of flowering dogwood, revealing a total of 233 unique alleles. Between nine and 18 alleles per locus with an average of 12 alleles per locus were detected. The observed frequency of heterozygotes (H_o) in the total sample was 0.384, which deviated from the expected calculated heterozygosity (H_e) of 0.829. Total gene diversity (expected heterozygosity) was 0.841 (Table 3-3, in appendix). The discrepancy between these numbers implies substantial population structure (or a systematic genotyping error, see below). All microsatellite loci were

polymorphic across sampling localities. The average allelic richness across all loci was 11.4 (calculated by rarefaction to 81 individuals or 162 gene copies per population) (Table 3-3, in appendix). The average private allelic richness was 0.78 (0.92 for northern and 0.64 for the southern area of the park).

STRUCTURE analysis resulted in a clear maximum for Δk at k =2 [two different clusters (Figures 3-3 and 3-4, in appendix)]. Multiple runs of the same k, used in this analysis, produced highly consistent individual assignment probabilities. These clusters were more or less spatially contiguous and corresponded well to the *a priori* grouping of trees into “northern” (above the Appalachian trail) and “southern” (below the Appalachian trail) groups (Figure 3-4, in appendix). The greatest evidence of gene flow between the two groupings was in the Abrams Creek area near the Little Tennessee River gorge at the west end of the mountain range, and in the Chattahoochee area near the French Broad River gorge at the east end of the mountain range.

AMOVA analysis indicated that most of the genetic variance can be explained by individual variation rather than divergence across the dividing ridge: 97.31 and 2.69% respectively, corresponding to F_{ST} = 0.027 (Table 3-4, in appendix). However, average deviation from Hardy-Weinberg expectations (F_{IS}) for both populations was 0.537, a significant deficit of heterozygotes (see discussion for additional explanations). The heterozygote deficiency was common to all markers (Table 3-3, in appendix), implying a population-level explanation (nonrandom mating) rather than a marker-specific explanation (selection or null alleles).

Analysis of isolation by distance across all samples showed significant correlation between genetic and geographic distance ($r=0.151$, $P < 0.0001$ and $r=0.143$, $P < 0.0001$, for populations south and north of the Appalachian trail, respectively), supporting the interpretation of spatial structure within the groups owing to restricted pollen flow.

Discussion

The patterns of genetic diversity and structure of native flowering dogwood populations showed evidence of high levels of genetic diversity within populations on both sides of the main ridge in the GSMNP, significant but low genetic differentiation between these populations, and evidence of substructure and isolation by distance within populations. This finding is consistent with other comparative surveys based on molecular markers suggesting that trees are more likely to have reduced among-population differentiation and increased within-population genetic diversity when compared to herbaceous plants and shrubs (Hamrick and Godt 1996; Nybom 2004; Petit and Hampe 2006). This is because of their obligate outcrossed mating system, extensive gene flow and diversifying selection, and large effective population sizes (Petit and Hampe 2006; Dubreuil et al. 2010). Flowering dogwoods are obligate outcrossers with gametophytic self-incompatibility (Reed 2004), which enforces a degree of disassortative mating and maintenance of local genetic variation. As a self-incompatible species, genetic variation of flowering dogwood is expected to be greater within populations than between populations when compared with self-compatible

species (Hamrick and Godt 1989; Leimu et al. 2006), which was confirmed in this study.

Substantial population structure for both northern (north of the Appalachian trail) and southern (south of the Appalachian trail) populations suggests strong limitations to gene flow across the main ridge of the Great Smoky Mountains. This pattern of variation implies a population-level explanation (nonrandom mating) that occurs between related individuals despite wide dispersal of seeds. In this study, significant heterozygote deficiencies were shown across all loci further suggesting that populations may not be panmictic. However, this result needs to be analyzed additionally because of constrains and limitations of QIAxcel system and allele scoring when compared to most commonly used capillary electrophoresis devices such as CEQ^{TM} 8000 (CEQ 8000:Genetic Analysis System, Beckman Coulter, USA) and ABI 3100xl DNA Sequencer (ABI 3100xl: Applied Biosystems, USA), the QIAxcel System (Qiagen, USA). Preprogrammed methods, combined with the corresponding disposable gel cartridges, allow separation and analysis of a variety of nucleic acids, including single or multiple PCR fragments, DNA digested with restriction endonucleases, synthesized oligonucleotides, total RNA, and cRNA (http://www.qiagen.com). When compared to two previously mentioned instruments, QIAxcel system is a relatively inexpensive machine since it utilizes disposable 12 micro-channel cartridges containing sieving-gel matrix with ethidium bromide dye to generate both gel image view and an electropherogram of the separated PCR products (Wang et al. 2009). Wang et al. (2009) compared

five different electrophoresis methods used in microsatellite studies and reported advantages and disadvantages of each including QIAxcel system, agarose and polyacrylamide gels, as well as CEQ^{TM} and ABI. Both CEQ^{TM} and ABI showed clear separation of products and were able to detect single base polymorphism. However, both high resolution systems are significantly more expensive when compared to other examined methods (Wang et al 2009). The QIAxcel system when compared to CEQ^{TM} and ABI showed lower resolution (2-4 bp difference) limiting possibility of showing heterozygotes that are 1-3 bp apart. Because of this, we believe that our calculations of heterozygote deficit may be overestimated and further analysis is necessary.

One of the biggest advantages of using QIAxcel system are the fact it is cost effective (20% of the price of CEQ^{TM} or ABI) and less labor intensive then other systems due to its minimal input volumes and automated loading and analysis. QIAxcel system offers higher detection sensitivity and rapid analysis of 96- well plate (approximately 2 h), resulting in automatic gel image view and possibility of automatic data export using Biocalculator software. Additionally, only small amounts of nucleic acid concentrations (as low as 0.1 ng/µl) are needed for system detection. Although small amounts of nucleic acid are needed to obtain results, resolution is between 3-5 bp.

The results obtained using classical population genetics were supported by Bayesian model-based clustering analysis. STRUCTURE identified two contiguous clusters that corresponded well to the *a priori* grouping of trees. This

result is unlikely to be an artifact of genotyping error, which would affect each region equally.

Flowering dogwood fruits (shiny, dark-red drupe, usually single seeded), are used as a food source by migratory birds and are likely to be deposited many kilometers from their sources (Trigiano, personal comm.). There are several ways in which long-distance seed dispersal can occur and they include vertebrate-mediated (by adhesion or ingestion), wind-mediated, and water-mediated seed dispersal (Yan et al. 2009). Since seeds are the only mean of natural propagation for flowering dogwood trees, frugivorous birds are important agents of seed distribution, carrying seeds in their bowels for up to 24 hours. This complies with the fact that broadscale, genetically autocorrelated dispersal of fruits in which large sets of seeds from a local area might all experience the same long-distance dispersal events, but possibly localized dispersal of pollen. Forest management or fragmentation can indeed influence pollen movement in a species pollinated by generalist pollinators (Sork et al. 2005). They estimated that pollen movement in both insect-pollinated *C. florida* and wind-pollinated *Pinus echinata* Mill. (Dyer 2002) is most restricted in the uncut sites and highest in the clear-cut sites (Sork et al. 2005).

In most models of population genetics, each offspring is assumed to have an independent dispersal path. However, seed dispersal in fruiting trees might not be independent for several reasons. Genetically autocorrelated dispersal might create a distinctive population structure, especially in species with highly dynamic spatial distributions. Genetic structure in species with animal-mediated

seed dispersal may depend on the type of fruit involved and the behavior of individual dispersers. Contrary to large, single seeded fruits in woody species where no clear genetic structure has been found, seedlings and/or adult plants in which birds may move seeds in groups of half- or full-siblings (kin-structured seed dispersal) through seed catching or multiseeded smaller fruits, showed clear fine-scale genetic structure (Torimaru et al. 2007). First, related individuals will be dispersed in groups (in the bowels of frugivorous birds due to repeated feeding off the same tree). Second, flocks of birds feeding on the same local set of trees might tend to move a large propagule *en masse* a long distance before depositing seeds all together in a new location. If these events tend to found new stands (e.g., after large disturbances due to fire, improper forest management or severe die-offs created by disease outbreaks), rather than join existing stands, a shifting mosaic of genotypic clusters with little large-scale isolation by distance, but significant small scale spatial-genetic autocorrelation could be established.

The migration of seed has the potential to introduce novel alleles into a population, thereby increasing genetic diversity and slowing the rate of divergence within the population. Subtle, but statistically significant difference in allele frequencies across the main dividing ridge of the Great Smoky Mountains was observed. Allele frequencies and genetic diversity in newly colonized populations could be negatively altered, too, considering that immigrant seeds may carry only a small portion of alleles from the source population. The founder event can be associated with a decline in genetic diversity, where most common alleles will more likely prevail and reduce frequency of rare alleles in the

population. Besides the number of founding propagules in comparison to migrants among existing populations, other factors such as kin structure and inbreeding within the colonizing populations, as well as the rate of population growth and immigration after colonization, can impact allele frequencies and genetic diversity in newly colonized populations. Since kin-structured seed dispersal can generate spatial aggregations of related individuals, it is likely to lead to evolutionary phenomena in which the spatial distribution of related individuals is important (biparental inbreeding depression and kin selection) (Torimaru et al. 2007).

Flowering dogwood trees are native to the United States (Witte et al. 2000) and therefore are important components of the forests they inhabit for both ornamental value during all seasons and as an important source of food for wildlife. Although high mortality rate and reduced fecundity caused by dogwood anthracnose severely affected native populations throughout the entire GSMNP, this study confirmed that considerable genetic diversity of flowering dogwoods exists at the population level. The study also implies that there is limited gene flow across the main ridge of the Great Smoky Mountains indicating nonrandom mating that occurs between related individuals despite wide dispersal of seeds. While the impact of dogwood anthracnose on genetic diversity may be small, consequences for flowering dogwood population viability and the community structure of eastern North American forests are still of great concern (Daughtre and Hibben 1994; Lande 1988). However, more research is needed on the

ecology of post-die-off recovery as well as disease and special ecology and population dynamics in widespread populations of flowering dogwood.

References

Amos W, Hoffman JI, Frodsham A, Zhang L, Best S, Hill VS (2007) Automated binning of microsatellite alleles: problems and solutions. Mol Ecol Notes 7:10-14.

ATBI Database Specimen Records (Ted Simmons Bird Project) http://www.dlia.org

Barrett SCH (1982) Genetic variation in weeds. In: Charudattan R, Walker H, eds. Biological control of weeds with plant pathogens. New York: John Wiley and Sons, pp. 73–98.

Barrett SCH, Colautti R, Eckert CG (2008) Plant reproductive systems and evolution during biological invasion. Mol Ecol 17:373–383.

Blair RM (1982) Growth and nonstructural carbohydrate content of southern browse species as influenced by light intensity. J Range Manage 35:756–760.

Bittencourt JVM, Sebbenn AM (2007) Patterns of pollen and seed dispersal in a small, fragmented population of the wind-pollinated tree *Araucaria angustifolia* in southern Brazil. Heredity 99:580–591.

Brasier CM (2000) Intercontinental spread and continuing evolution of the Dutch elm disease pathogens. In: Dunn CP (ed) The Elms: Breeding, Conservation and Disease Management. Kluwer Academic Publishers, Boston, pp. 61–72.

Chellemi DO, Britton KO (1992) Influence of canopy microclimate on incidence and severity of dogwood anthracnose. Can J Bot 70:1093-1096.

Chellemi DO, Britton KO, Swank WT (1992) Influence of site factors on dogwood anthracnose in the Nantahala mountain range of western North Carolina. Plant Dis 76:915-918.

Crowl TA, Crist TO, Parmenter RR, Belovsky G, Lugo AE (2008) The spread of invasive species and infectious disease as drivers of ecosystem change. Front Ecol Environ 6:238-246.

Dubreuil M, Riba M, Gonzlez-Martinez SC, Vendramin GG, Sebastiani F, Mayol M. (2010). Genetic effects on chronic habitat fragmentation revisited: Strong genetic structure in a temperate tree, *Taxus baccata* (Taxaceae), with great dispersal capability. Am J Bot 97:303–310.

Daughtrey ML, Hibben CR (1994) Dogwood anthracnose: a new disease threatens two native *Cornus* species. Annu Rev Phytopathol 32:61-73.

Dyer RJ (2002). Contemporary pollen movement in shortleaf pine, *Pinus echinata* Mill. Diss. University of Missouri, St. Louis, Missouri, USA.

Elliott KJ, Swank WT (2008) Long-term changes in forest composition and diversity following early logging (1919–1923) and the decline of American chestnut (*Castanea dentata*). Plant Ecol 197:155–172.

Eschtruth AK, Battles JJ (2008) Deer herbivory alerts forest response to canopy decline caused by an exotic insect pest. Ecol Appl 18:360-376.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. Mol Ecol 14:2611–2620.

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evol Bioinform Online 1:47-50.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. Genetics 131:479-491.

Franceschinelli EV, Bawa KS (2000) The effect of ecological factors on the mating system of a South American shrub species (*Helicteres brevispira*). Heredity 84:116–123.

Fischer M, Matthies D (1998) Effects of population size on performance in the rare plant *Gentianella germanica*. J Ecol 86:195–204.

Godoy JA, Jordano P (2001). Seed dispersal by animal: exact identification of source trees with endocarp DNA microsatellites. Mol Ecol 10:2275–2283.

Goudet J. 2001. Fstat 2.9.3. program available from: http://www2.unil.ch/popgen/softwares/fstat.htm

Hadziabdic D, Fitzpatrick BM, Wang W, Wadl PA, Rinehart TA, Ownley BH, Windham MT, Trigiano RN. Genetic diversity of flowering dogwood

maintained despite massive mortality caused by dogwood anthracnose. (Genetica, accepted for publication 2010)

Hall B, Motzkin G, Foster DR, Syfert M, Burk J (2002) Three hundred years of forest and land-use change in Massachusetts, USA. J Biogeogr 29:1319-1335.

Halls LK, Epps EA (1969) Browse quality influenced by tree overstory in the South. J Wildl Manage 33:1028-1031.

Hamrick JL, Murawski DA, Nason JD (1993). The influence of seed dispersal mechanisms on the genetic structure of tropical tree populations. Vegetatio 108:281–297.

Hamrick JL, Godt MJW (1989) Allozyme diversity in plant species. In: Brown HD, Clegg MT, Kahler AL, Weir BS (eds) Plant Population Genetics, Breeding, and Genetic Resources, Sinauer Associates, Sunderland, Massachusetts, pp. 43–63.

Hamrick JL, Godt MJW. 1996. Effects of life history traits on genetic diversity in plant species. Philos Trans R Soc London Ser B 351:1291–1298.

Havill NP, Montgomery ME, Yu G, Shiyake S, Caccone A (2006) Mitochondrial DNA from hemlock woolly adelgid (Hemiptera: Adelgidae) suggests cryptic speciation and pinpoints the source of the introduction to eastern North America. Ann Entomol SocAm 99:195-203.

Hawley DM, Hanley D, Dhondt AA, Lovette IJ (2006) Molecular evidence for a founder effect in invasive house finch (*Carpodacus mexicanus*)

populations experiencing an emergent disease epidemic. Mol Ecol 15:263-275.

Hiers JK, Evans JP (1997) Effects on anthracnose on dogwood mortality and forest composition of the Cumberland Plateau (U.S.A.). Conserv Biol 11:1430-1435.

Hibben CR, Daughtrey ML (1988) Dogwood anthracnose in the northeastern United States. Plant Dis 72:199-203.

Holzmueller E, Jose S, Jenkins M, Camp A, Long A (2006) Dogwood anthracnose in eastern hardwood forests: what is known and what can be done? J Forest 104:21-26.

Holzmueller E (2006) Ecology of flowering dogwood (*Cornus florida* L.) in response to anthracnose and fire in Great Smoky Mountains National Park, USA. Diss. University of Florida.

Innan H, Terauchi R, Miyashita NT (1997) Microsatellite polymorphism in natural populations of the wild plant *Arabidopsis thaliana.* Genetics 146:1441-1452.

Jenkins MA, White PS (2002) *Cornus florida* L. mortality and understory composition changes in western Great Smoky Mountains National Park. J Torrey Bot Soc 129:194-206.

Kalinowski ST (2005) HP-Rare: a computer program for performing rarefaction on measures of allelic diversity. Mol Ecol Notes 5:187-189.

Kalinowski ST, Wagner AP, Taper ML (2006) ML-Relate: a computer program for maximum likelihood estimation of relatedness and relationship. Mol Ecol Notes 6:576-579.

Karasawa MG, Vencovsky R, Silva CM, Zucchi MI, Oliveira GCX, Veasey EA (2007) Mating system of Brazilian *Oryza glumaepatula* populations studied with microsatellite markers. Ann Bot 99:245–253.

Lande R (1988) Genetics and demography in biological conservation. Science 241:1455-1460.

Lande RC, Shannon S (1996) The role of genetic variation in adaptation and population persistence in a changing environment. Evol 50:434–437.

Leimu R, Mutikainen P, Koricheva J, Fischer M (2006) How general are positive relationships between plant population size, fitness and genetic variation? J Ecol 94:942–952.

Linzey DW, Brecht C (2003)a American beavers (*Castor canadensis*) - Biodiversity of Great Smoky Mountains National Park, Discover Life in America, Gatlinburg, Tennessee 37738.

Linzey DW, Brecht C (2003)b Eastern gray squirrels (*Sciurus carolinensis*) - Biodiversity of Great Smoky Mountains National Park, Discover Life in America, Gatlinburg, Tennessee 37738.

Mandak,B, Bimova K, Plackova I, Mahelka V, Chrtek J (2005) Loss of genetic variation in geographically marginal populations of *Atriplex tatarica* (Chenopodiaceae). Ann Bot 96:901–912.

Margalef R (1962) Successions of populations. Adv Front Plant Sci. 2:137-188.

McEwan RW, Mueller RN, Arthur MA, Housman HH (2000) Temporal and ecological patterns of flowering dogwood mortality in the mixed mesophytic forest of eastern Kentucky. J Torrey Bot Soc 127:221-229.

McGee CE (1986). Loss of *Quercus* spp. dominance in an undisturbed old-growth forest. J Elisha Mitch Sci S 102:10-15.

Merkle SA, Andrade GM, Nairn CJ, Powell WA, Maynard CA (2007) Restoration of threatened species: a noble cause for transgenic trees. Tree Genet Genomes 3:111-118.

Murawasky DA, Hamrick JL (1991) The effect of the density of flowering individuals on the plant mating systems of nine tropical tree species. Heredity 67:167–174.

Nei M (1975) Molecular population genetics and evolution. North-Holland Publishing, Amsterdam, The Netherlands.

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, New York.

Nybom H (2004) Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. Mol Ecol 13:1143–55.

Persson H, Widen B, Andersson S, Svensson L (2004) Allozyme diversity and genetic structure of marginal and central populations of *Corylus avellana* (Betulaceae) in Europe. Plant Syst Evol 244: 157–179.

Petit RJ, Hampe A (2006) Some evolutionary consequences of being a tree. Annu Rev Ecol Evol Syst 37:187-214.

Pimentel D (1997) Techniques for Reducing Pesticides: Environmental and Economic Benefits. John Wiley and Sons, Chichester, UK.

Pimentel D, Zuniga R, Morrison D (2005) Update on the environmental and economic costs associated with alien-invasive species in the United States. Ecol Econ 52:273-288.

Plat WJ (1975) The colonization and formation of equilibrium plant species associations on badger disturbances in a tall-grass prairie. Ecol Monogr 45:285-305.

Pritchard JK, Donnelly P (2001) Case–control studies of association in structured or admixed populations. Theor Popul Biol 60:227–237.

QIAGEN: http://www1.qiagen.com/

Ramseur GS, Kelly JM (1981) Forest characterization and bio-mass estimates for two sites on the Cumberland Plateau. J Tennessee Acad Sci 56:99-104.

Ravigne V, Olivieri I, Gonzalez-Martinez SC, Rousset F (2006) Selective interactions between short-distance pollen and seed dispersal in self-compatible species. Evol 60:2257–2271.

Redlin SC (1991) *Discula destructiva* sp. nov., cause of dogwood anthracnose. Mycologia 83:633-642.

Reed DH, Frankham R (2003) Correlation between fitness and genetic diversity. Conserv Biol 17:230–237.

Reed S (2004) Self-incompatibility in *Cornus florida* L. HortScience 39:335-338.

Rossell IM, Rossell CR, Hining KJ, Anderson RL (2001) Impacts of dogwood anthracnose (*Discula destructiva* Redlin) on the fruits of flowering dogwood (*Cornus florida* L.): implications for wildlife. Am Midl Nat 146:379-387.

Rousset F (2000) Genetic differentiation between individuals. J Evol Biol 13:58-62.

Severns P (2003) Inbreeding and small population size reduce seed set in a threatened and fragmented plant species, *Lupinus sulphureus* ssp. kincaidii (Fabaceae). Biol Conserv 110:221–229.

Sherald JL, Stidham TM, Hadidian JM, Hoeldtke JE (1996) Progression of the dogwood anthracnose epidemic and the status of flowering dogwood in Catoctin Mountain Park. Plant Dis 80:310-312.

Smouse PE, Sork VL (2004). Measuring pollen flow in forest trees: an exposition of alternative approaches. Forest Ecol Manage 197:21–38.

Sork VL, Smouse PE, Apsit VJ, Dyer RJ, Westfall RD (2005) A two-generation analysis of pollen pool genetic structure in flowering dogwood, *Cornus florida* (Cornaceae), in the Missouri Ozarks. Am J Bot 92:262-271.

Stiles EW (1980) Patterns of fruit presentation and seed dispersal in bird-disseminated woody plants in the eastern deciduous forest. Am Natur 116:670-688.

Torimaru T, Tani N, Tsumura Y, Nishimura N, Tomaru N (2007) Effects of kin-structured seed dispersal on the genetic structure of the clonal dioecious shrub *Ilex leucoclada.* Evol 61:1289–1300.

Thomas WA (1969) Accumulation and cycling of calcium by dogwood trees. Ecol Monogr 39:101-120.

Trigiano RN, Caetano-Anollés G, Bassam, BJ, Windham MT (1995) DNA amplification fingerprinting provides evidence that *Discula destructiva*, the cause of dogwood anthracnose in North America, is an introduced pathogen. Mycologia 87:490-500.

Wadl PA, Skinner JA, Dunlap JR, Reed SM, Rinehart TA, Pantalone VR, Trigiano RN (2009) Honeybee-mediated controlled pollinations in *Cornus florida* and *C. kousa* intra- and interspecific crosses. HortScience 44:1527-1533.

Wang X, Rinehart TA, Wadl PA, Spiers JM, Hadziabdic D, Windham MT, Trigiano RN (2009) A new electrophoresis technique to separate microsatellite alleles. Afr J Biotechnol 8:2432-2436.

Wang XW, Trigiano RN, Windham MT, Scheffler BE, Rinehart TA, Spiers JA (2008) Development and characterization of simple sequence repeats for flowering dogwood (*Cornus florida* L.). Tree Genet Genomes 4:461-468.

Williams CE, Moriarity WJ (1999) Occurrence of flowering dogwood (*Cornus florida* L.), and mortality by dogwood anthracnose (*Discula destructiva* Redlin), on the Northern Allegheny Plateau. J Torrey Bot Soc 126:313-319.

Windham MT, Graham ET, Witte WT, Knighten JL, Trigiano RN (1998) *Cornus florida* 'Appalachian Spring': A white flowering dogwood resistant to dogwood anthracnose. HortScience 33:1265-1267.

Windham WT, Montgomery-Dee M, Parham JM (1995) Factors associated with dogwood anthracnose incidence and severity. Tennessee AgriScience 175:17-20.

Witte WT, Windham MT, Windham AS, Hale FA, Fare DC, Clatterbuck WK (2000) Dogwoods for American gardens. The University of Tennessee Agriculture Extension Service, Knoxville. PB1670.

Yan XB, Guo YX, Liu FY, Zhao C, Liu QL, Lu BR (2010) Population structure affected by excess gene flow in self-pollinating *Elymus nutans* and *E. burchan-buddae* (Triticeae: Poaceae) Popul Ecol 52:233–241.

Appendix 3: Tables

Table 3-1. *Cornus florida* sampling locations throughout Great Smoky Mountains National Park, United States with GPS coordinates and trail/road location within the park.

Sample	Location	Latitude	Longitude	Altitude (meters)
BC NC 1	Big Creek Parking Lot, NC	35.752883	-83.109417	564
BC NC 2	Baxter Creek Trail, NC	35.749383	-83.113267	557
BC NC 3	Bridge near Baxter Creek Trail	35.750900	-83.108750	550
BC NC 4	Big Creek Camping Site	35.749900	-83.112350	572
BC NC 5	Big Creek Trail	35.749783	-83.115533	560
BC NC 6	Big Creek Trail	35.750200	-83.115583	551
CBT 7	Chestnut Branch Trail, NC	35.759717	-83.120233	546
CBT 8	Chestnut Branch Trail, NC	35.759517	-83.120233	549
CBT 9	Chestnut Branch Trail, NC	35.765500	-83.132950	813
CBT 10	Chestnut Branch Trail, NC	35.764950	-83.131617	816
RD 11	Road to Mt. Sterling Gap	35.751650	-83.094500	686
RD 12	Road to Mt. Sterling Gap	35.749367	-83.094717	696
RD 13	Road to Mt. Sterling Gap	35.721317	-83.086333	978
RD 14	Road to Mt. Sterling Gap	35.698983	-83.097467	1266
CP 15	Cosby Picnic Area	35.756750	-83.208600	699
CP 16	Cosby Picnic Area	35.757767	-83.209500	671
CP 17	Cosby Picnic Area	35.757150	-83.207883	668
LGT 18	Low Gap Trail	35.754900	-83.207217	750
LGT 19	Low Gap Trail	35.754167	-83.206950	743
LMC 20	Lower Mountain Cammerer	35.755533	-83.198950	774
LMC 21	Lower Mountain Cammerer	35.757883	-83.195500	588
CCG 22	Cosby Campground	35.752133	-83.206967	824

Table 3-1 Continued.

Sample	Location	Latitude	Longitude	Altitude (meters)
CCG 23	Cosby Campground	35.752750	-83.210067	769
SDT 24	Snake Dan Trail	35.752150	-83.210467	799
CR 25	Cosby Road	35.759583	-83.209450	712
CR 26	Cosby Road	35.783533	-83.218433	556
HWY 32 #27	Highway 32	35.783783	-83.203167	543
HWY 32 #28	Highway 32	35.787233	-83.147650	686
HWY 32 #29	Highway 32	35.774883	-83.112500	628
HWY 32 #30	Highway 32	35.766300	-83.111067	602
G 31	Greenbriar	35.737867	-83.416250	468
G 32	Greenbriar	35.733033	-83.410633	464
G 33	Greenbriar	35.729700	-83.406533	495
G 34	Greenbriar	35.726633	-83.401683	511
G 35	Greenbriar	35.727283	-83.400917	513
GPA 36	Greenbriar Picnic Area	35.712367	-83.384267	511
GRT 37	Grapeyard Ridge	na	na	na
RCT 38	Greenbriar Rd towards Ramsey Cascades	na	na	na
RCT 39	Greenbriar Rd towards Ramsey Cascades	na	na	na
RCT 40	Ramsey Cascades Trail	35.705350	-83.356300	1007
RCT 41	Ramsey Cascades Trail	35.705950	-83.353033	992
RCT 42	Ramsey Cascades Trail	35.702233	-83.347667	812
OST 43	Old Settlers Trail	35.708200	-83.380367	548
OST 44	Old Settlers Trail	35.710533	-83.381583	549
OST 45	Old Settlers Trail	35.714883	-83.382100	550
OST 46	Old Settlers Trail	35.717417	-83.377500	551
PCT 47	Road towards Porter Creek Trail	na	na	na
PCT 48	Porter Creek Trail	35.696367	-83.388550	644

Table 3-1 Continued.

Sample	Location	Latitude	Longitude	Altitude (meters)
PCT 49	Porter Creek Trail	35.690917	-83.391917	638
PCT 50	Porter Creek Trail	35.684383	-83.396400	756
BMT 51	Brushy Mountain Trail	35.686400	-83.398100	747
BMT 52	Brushy Mountain Trail	35.686517	-83.400567	743
RMNT 53	Low Gap Road	35.717983	-83.486483	465
RCC 54	Road to Cades Cove	35.659017	-83.709283	411
RCC 55	Road to Cades Cove	35.643050	-83.713733	418
SHG 56	Schoolhouse Gap	35.630100	-83.732033	415
SHG 58	Schoolhouse Gap	35.634933	-83.732100	558
SHG 59	Schoolhouse Gap	35.638467	-83.738200	556
SHG 60	Schoolhouse Gap	35.643883	-83.747983	645
CTT 61	Chestnut Top Trail	35.646550	-83.746550	646
CTT 62	Chestnut Top Trail	35.648333	-83.743017	715
CTT 63	Chestnut Top Trail	35.651000	-83.735117	717
RTR 64	Rd. to Treemont	35.653117	-83.701200	561
RTR 65	Rd. to Treemont	35.650583	-83.696267	434
RTR 66	Rd. to Treemont	35.642117	-83.692083	439
WPT 67	West Prong Trail	35.641050	-83.695067	454
GTR 68	Gravel Rd. to Treemont	35.628717	-83.685050	456
GTR 69	Gravel Rd. to Treemont	35.623300	-83.682067	516
GTR 70	Gravel Rd. to Treemont	35.622200	-83.674383	554
MPT 71	Middle Prong Trail	35.609500	-83.671183	646
MPT 72	Middle Prong Trail	35.607317	-83.667750	644
LR 73	Little River Rd.	35.662717	-83.701200	645
LR 74	Little River Rd.	35.668533	-83.691633	640
LR 75	Little River Rd.	35.668533	-83.683567	632
MCT 76	Meig's Creek	35.670517	-83.660933	719

Table 3-1 Continued.

Sample	Location	Latitude	Longitude	Altitude (meters)
MBP 78	Metcalf Bottoms Picnic Area	35.679467	-83.648383	552
LBI 79	Little Briar Trail	35.686367	-83.638983	591
AWC 80	Alice Walker Cabin	35.695083	-83.629567	648
LR 81	Little River Rd.	35.672817	-83.640933	548
LR 82	Little River Rd.	35.670400	-83.626583	658
CC 83	Cades Cove	35.606167	-83.776300	634
CC 84	Cades Cove	35.606967	-83.787367	560
CC 85	Cades Cove	35.605517	-83.801767	550
CC 86	Cades Cove	35.607700	-83.825400	607
CC 87	Cades Cove	35.597533	-83.842867	558
RCT 88	Rabbit Creek Trail	35.585900	-83.865533	629
RCT 89	Rabbit Creek Trail	35.585950	-83.869183	608
RCT 90	Rabbit Creek Trail	35.587283	-83.876717	813
AF 91	Abram Falls	35.600517	-83.865967	637
AF 92	Abram Falls	35.607600	-83.873133	534
AF 93	Abram Falls	35.608333	-83.879400	531
AF 94	Abram Falls	35.609233	-83.879400	525
AF 95	Abram Falls	35.591333	-83.853433	523
CCRA 96	Cades Cove	35.584917	-83.842933	564
CC 97	Cades Cove	35.588783	-83.835433	565
CC 98	Cades Cove	35.589267	-83.817533	574
CC 99	Cades Cove	35.588300	-83.802617	574
CC 100	Cades Cove	35.597750	-83.787683	600
LBT 101	Little Bottom Trail	35.614083	-83.929883	404
CRT 102	Cooper Road Trail	35.624967	-83.916433	403
CRT 103	Cooper Road Trail	35.628417	-83.911967	402
GMT 104	Goldmine Trail	35.635383	-83.905450	489

Table 3-1 Continued.

Sample	Location	Latitude	Longitude	Altitude (meters)
CRT 105	Cooper Road Trail	35.637783	-83.898817	446
CRT 106	Cooper Road Trail	35.634200	-83.894333	568
CRT 107	Cooper Road Trail	35.626300	-83.883350	502
HMT 108	Hatcher Mountain Trail	35.619433	-83.890350	644
LBT 110	Little Bottom Trail	35.609817	-83.900133	492
LBT 111	Little Bottom Trail	35.611400	-83.905317	427
LBT 112	Little Bottom Trail	35.616617	-83.912450	422
LBT 113	Little Bottom Trail	35.619517	-83.919033	413
WRT 114	Walf Ridge Trail	na	na	na
TML 115	Twenty Mile Loop	35.484783	-83.865567	643
TML 116	Twenty Mile Loop	35.483533	-83.861100	826
TML 117	Twenty Mile Loop	35.484117	-83.849700	821
McKee 118	McKee Branch Trail	35.597933	-83.103400	265
Caldwell 119	Caldwell Fork Trail	35.597933	-83.103400	265
McKee 120	McKee Branch Trail	35.595383	-83.099050	278
McKee 121	McKee Branch Trail	35.585950	-83.088533	352
Hemphill 122	Hemphill Bald	35.582033	-83.114600	323
Caldwell 123	Caldwell Fork Trail	35.585750	-83.120400	299
Caldwell 124	Caldwell Fork Trail	35.585750	-83.120400	299
Caldwell 125	Caldwell Fork Trail	35.585750	-83.120400	299
Caldwell 126	Caldwell Fork Trail	35.585750	-83.120400	299
Rough 127	Rough Fork Trail	35.593650	-83.139617	332
Rough 128	Rough Fork Trail	35.607450	-83.128217	264
Rough 129	Rough Fork Trail	35.607450	-83.128217	264
Rough 130	Rough Fork Trail	35.616733	-83.121567	251
Rough 131	Rough Fork Trail	35.616733	-83.121567	251
Cat 132	Cataloochee Road	35.622317	-83.117567	246

Table 3-1 Continued.

Sample	Location	Latitude	Longitude	Altitude (meters)
Cat 133	Cataloochee Road	35.627183	-83.110017	244
Cat 134	Cataloochee Road	35.628700	-83.099900	239
Cat 135	Cataloochee Road	35.634067	-83.083983	236
BF 136	Bradley Fork	35.569267	-83.309883	818
BF 137	Bradley Fork	35.605083	-83.328467	808
CC 138	Chasteen Creek	35.607267	-83.292083	1345
CC 139	Chasteen Creek	35.586533	-83.310683	795
CC 140	Chasteen Creek	35.586533	-83.310683	795
CC 141	Chasteen Creek	35.578000	-83.312750	745
Scamp 142	Smokemont Campground	35.561767	-83.312500	690
Scamp 143	Smokemont Campground	35.561767	-83.312500	690
Entrance 144	Entrance to Smokemont	35.552633	-83.308983	680
MMP 145	Mingus Mill Parking	35.552583	-83.309100	680
MMP 146	Mingus Mill Parking	35.552583	-83.309100	680
Ocntr 147	Oconaluftee Center	35.520867	-83.309017	645
US 441 #148	US 441	35.532950	-83.303633	640
TSR 149	Tow String Road	35.542383	-83.297467	650
TSR 150	Tow String Road	35.542450	-83.297417	650
US 441 #151	US 441	35.554383	-83.314817	695
US 441 #152	US 441	35.563450	-83.329967	710
US 441 #153	US 441	35.563450	-83.329967	710
CCP 154	Collins Creek Picnic	35.563350	-83.330100	710
CCP 155	Collins Creek Picnic	35.568050	-83.338550	765
US 441 #156	US 441	35.588600	-83.364267	870
BMR 157	Balsam Mtn. Rd.	35.615283	-83.212417	940
BMR 158	Balsam Mtn. Rd.	35.615283	-83.212417	940
BMR 159	Balsam Mtn. Rd.	35.615283	-83.212417	940

Table 3-1 Continued.

Sample	Location	Latitude	Longitude	Altitude (meters)
BMR 160	Balsam Mtn. Rd.	35.615283	-83.212417	940
BMR 161	Balsam Mtn. Rd.	35.610033	-83.221667	905
BMR 162	Balsam Mtn. Rd.	35.609917	-83.221517	900
BMR 163	Balsam Mtn. Rd.	35.589183	-83.235583	825
BMR 164	Balsam Mtn. Rd.	35.586800	-83.236117	810
DCR 165	Deep Creek Rd.	35.460750	-83.436817	545
DCHT 166	Deep Creek Horse Trail	35.466150	-83.434133	600
DCHT 167	Deep Creek Horse Trail	35.471550	-83.435050	675
DCT 168	Deep Creek Trail	35.473600	-83.430283	595
DCT 169	Deep Creek Trail	35.472317	-83.428600	590
DCC 170	Deep Creek Campground	35.461433	-83.434983	560
DCR 171	Deep Creek Rd.	35.460750	-83.435050	560
FET 172	End Tunnel	35.461483	-83.544000	645
173	Lakeshore Trail	35.470033	-83.558367	605
174	Lakeshore Trail	35.466333	-83.558367	600
175	Lakeshore Trail	35.460200	-83.564050	600
176	Lakeshore Trail	35.456550	-83.576633	575
177	Lakeshore Trail	35.457867	-83.585733	580
179	Lakeshore Trail	35.453983	-83.564717	580
180	Lakeshore Trail	35.448017	-83.608017	555
181	Lakeshore Trail	35.445167	-83.622400	510
182	Lakeshore Trail	35.446483	-83.635750	515
183	Lakeshore Trail	35.445000	-83.648783	540
184	Lakeshore Trail	35.440683	-83.659450	580
185	Lakeshore Trail	35.446817	-83.665950	510
186	Lakeshore Trail	35.445833	-83.676967	635
187	Lakeshore Trail	35.446317	-83.689900	590

Table 3-1 Continued.

Sample	Location	Latitude	Longitude	Altitude (meters)
188	Lakeshore Trail	35.454267	-83.694600	555
189	Lakeshore Trail	35.451300	-83.703500	550
190	Lakeshore Trail	35.451217	-83.703483	550
191	Lakeshore Trail	35.472983	-83.722967	555
192	Lakeshore Trail	35.478950	-83.739833	585
193	Lakeshore Trail	35.483867	-83.759100	660
194	Lakeshore Trail	35.482250	-83.781267	590
195	Lakeshore Trail	35.482233	-83.781150	590
196	Lakeshore Trail	35.465550	-83.791267	540
197	Lakeshore Trail	35.465483	-83.791400	550
198	Lakeshore Trail	35.461050	-83.811283	595

Table 3-2. Primer information for 19 microsatellite loci used to analyze 194 *Cornus florida* samples collected in Great Smoky Mountains National Park, United States.

Locus	Primer sequences (5'−3')	GenBank accesion number	Expected size (bp)
CF 020	F: TATGGCTTGCTTTGGCTAATTGTT	ED651708	144
	R: CCAACTTATGCACACAGTGACACA		
CF 025	F: ATTATTTTTCGGGACCATTGCAT	ED651711	158
	R: CGGCTAAGAAAATATGTCCCTTTGT		
CF 046	F: TGCAATTGACATTTGTTGCATTTT	ED651730	151
	R: CAACTACGTGATCAGGTGGACAAC		
CF 072	F: AAACACCCCACTGCCATATAATGT	ED651755	132
	R: CCATATGGGACCCAGTGATTTTA		
CF 091	F: GCACTAGTATAAAAACATACGCGCC	ED651772	143
	R: TGCCTACATGGTAGCATTCCTTTT		
CF 100	F: CCAAAGGCTATGAATTTACGATGC	ED651777	112
	R: GATGCAGTCTTAATCCTCAACGTG		
CF 121	F: CCCGAAAATCAAATATGGTAAATAAGTG	ED651797	154
	R: CTCTCTTAAACCACCTTTTGATGTGA		
CF 203	F: TCCTACGCACCCACCCTTAT	ED651865	135
	R: ACGGTGGTGCCATTCTTTCT		
CF 209	F: TACAAAAGTGCAATGCCAATACC	ED651871	197
	R: ACTCCAAAACTTCATTCCTGAGC		
CF 214	F: TGCAAATGGTTATTGATTGCTCTC	ED651874	129
	R: ATTTGTTTCCCATGACCTGAAAGA		
CF 261	F: GGACATGTGGGCAGTTTGAT	ED651912	143
	R: GCAATGAGATCCCTCCCCTA		
CF 262	F: ATGACAAATGGCAAGTTGATGTTG	ED651913	197
	R: CAACTGTGACTCAATTCATGTCTGC		

Table 3-2 Continued.

Locus	Primer sequences (5'−3')	GenBank accesion number	Expected size (bp)
CF 267	F: TTGGGTAGGTGAGGGAATTTAGGT	ED651915	95
	R: TGCTATTACACATTTACCCCTGCC		
CF 290	F: GGTGGATCAAAACCCACAAATAAG	ED651936	153
	R: CCTATAGCCCAGCGAGAGTCTTTA		
CF 398	F: CCTCGCTTAATTCATATACTTCTT	ER870420	156
	R: AATAAAAGTTTGTATGAACTGCAC		
CF 488	F: CTCAAAATCACCCTTCTTATATCT	ER870510	114
	R: GATAGTTCACAAGCAATGTTACTG		
CF 597	F: AAGTCAGATCATTTCAGATTAACA	ER870619	107
	R: CGAATTGACGATAAATACAAAATA		
CF 599	F: TATGGTTTTCTTTTTCTTCTTTTT	ER870621	150
	R: ACCTAATAAAGGATGTTGGATAAG		
CF 862	F: AAGTGAGAGGAGAATGCTACAG	ED652042	150
	R: GCGTATTCACAAACTTGATTG		

Table 3-3. Nei's estimation of heterozygosity across all loci in *Cornus florida* populations was computed using FSTAT version 2.9.3. HP-Rare version 1.0 was used for allelic richness calculations. Allelic richness (A_R) for both populations, observed heterozygosity (Ho), expected heterozygosity (gene diversity) under Hardy Weinberg equilibrium (He), total expected heterozygosity (Ht), inbreeding coeficient (F_{IS}).

Locus	**A_R**		**H_o**	**H_e**	**H_t**	**F_{IS}**
	Pop1	**Pop2**				
CF20	13.684	11.997	0.618	0.901	0.902	0.314
CF25	9.893	9.820	0.124	0.689	0.724	0.821
CF46	11.934	10.832	0.232	0.851	0.854	0.728
CF72	17.957	12.915	0.465	0.911	0.911	0.49
CF91	7.843	7.840	0.047	0.725	0.755	0.935
CF100	9.000	9.850	0.45	0.86	0.868	0.476
CF121	9.931	8.975	0	0.574	0.61	1
CF203	11.861	12.508	0.559	0.859	0.875	0.35
CF209	8.979	8.931	0.641	0.81	0.813	0.209
CF214	16.812	15.984	0.408	0.917	0.916	0.555
CF261	14.000	12.982	0.46	0.919	0.918	0.499
CF262	13.861	13.998	0.551	0.915	0.915	0.398
CF267	7.998	10.567	0.346	0.714	0.762	0.516
CF290	13.858	14.841	0.601	0.862	0.862	0.303
CF398	9.000	8.000	0.105	0.762	0.776	0.862
CF488	11.820	11.973	0.615	0.899	0.899	0.316
CF597	11.844	12.976	0.573	0.896	0.895	0.36
CF599	9.000	10.000	0.137	0.848	0.87	0.838
CF862	9.931	8.910	0.358	0.839	0.844	0.573
Overall	11.537	11.258	0.384	0.829	0.841	0.537

Table 3-4. Estimates of Analysis of Molecular Variance (AMOVA) from *Cornus florida* microsatellite data using Arlequin (version 3.1). This analysis included all sampling sites as one hierarchical group.

Variance partition	d.f.	Sum of squares	Variance component	% of variation	P value
Among populations	1	32. 958	0.14323 Va	2.69	<0.0001
Within populations	386	1997.122	5.17289 Vb	97.31	<0.0001
Total	387	2030.080	5.31712		
Fixation Index	Fst:	0.027			

Appendix 4: Figures

Figure 3-1. Confirmed cases of dogwood anthracnose for the past three decades. Dogwood anthracnose data courtesy of United States Department of Agriculture (USDA) - Forest Service-Forest Health Protection; distribution map courtesy of United States Geological Survey (USGS).

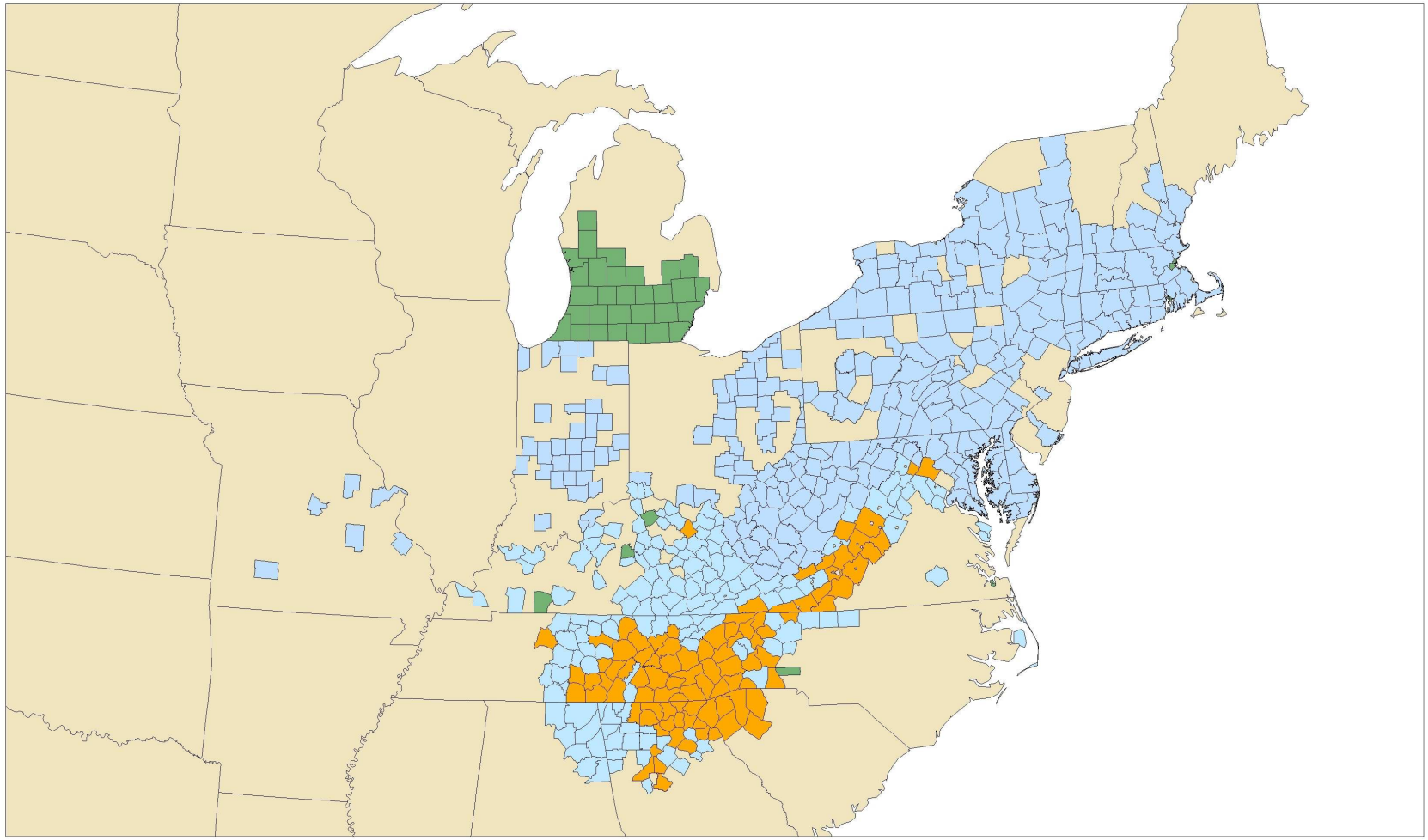

Legend
2000 to 2006
1986 to 1989
1990 to 1999
State Boundaries
N
Confirmed Cases of Dogwood Anthracnose
Data courtesy of USDA - Forest Service - Forest Health Protection

Figure 3-2. *Cornus florida* sampling locations in the Great Smoky Mountains National Park.

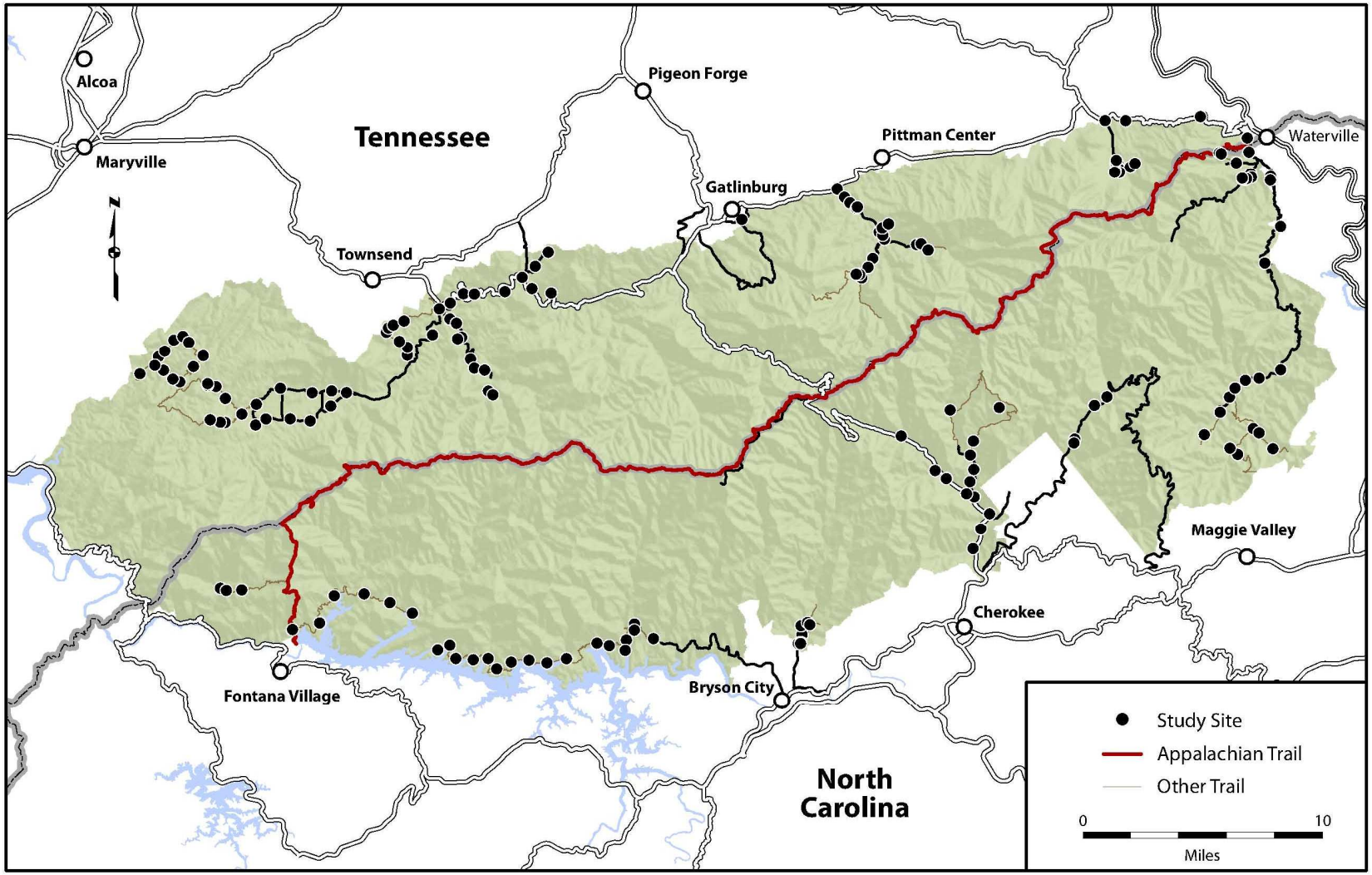
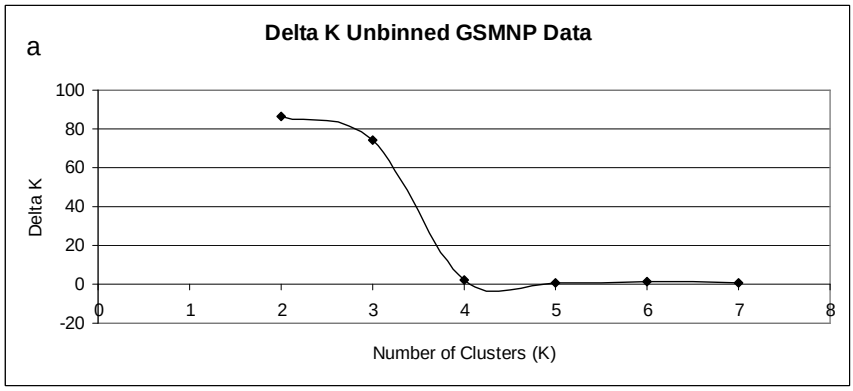
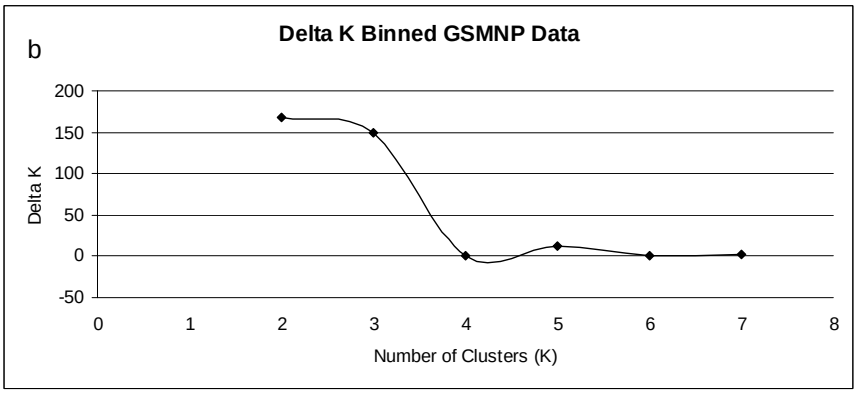
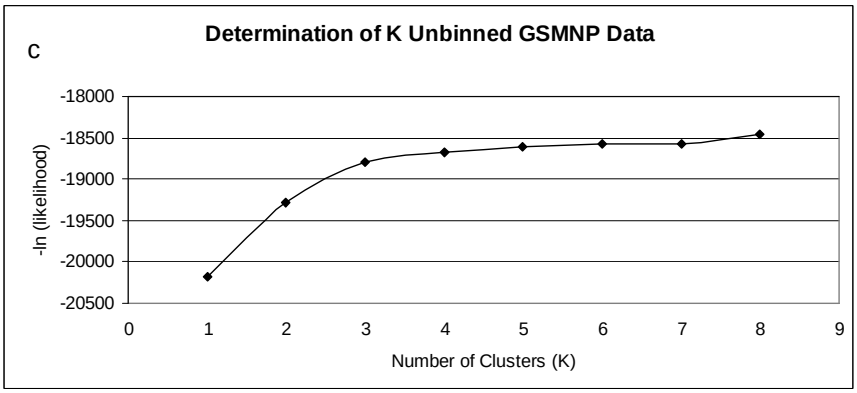
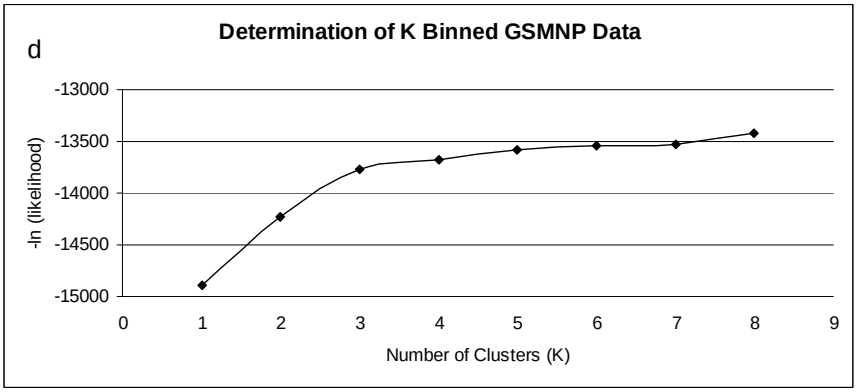

Alcoa
Maryville
Tennessee
Pigeon Forge
Pittman Center
Waterville
Gatlinburg
Townsend
Maggie Valley
Cherokee
Fontana Village
Bryson City
North Carolina
Study Site
Appalachian Trail
Other Trail
0
10
Miles

Figure 3-3. Delta K (Δk) calculations for unbinned (a and c) and binned (b and d) data. Determination of the number of clusters (k) vs. the second order rate of change in k (Δk) (upper 2 graphs). The clear maximum for Δk at k=2 indicates that two clusters fit the microsatellite data best for C*ornus florida* populations. The number of clusters (k) vs. the negative natural log of the likelihood [-ln (likelihood)] of the microsatellite data (lower 2 graphs).


a
Delta K Unbinned GSMNP Data
Delta K
100
80
60
40
20
0
-20
0 1 2 3 4 5 6 7 8
Number of Clusters (K)


b
Delta K Binned GSMNP Data
Delta K
200
150
100
50
0
-50
0 1 2 3 4 5 6 7 8
Number of Clusters (K)


c
Determination of K Unbinned GSMNP Data
-ln (likelihood)
-18000
-18500
-19000
-19500
-20000
-20500
0 1 2 3 4 5 6 7 8 9
Number of Clusters (K)


d
Determination of K Binned GSMNP Data
-ln (likelihood)
-13000
-13500
-14000
-14500
-15000
0 1 2 3 4 5 6 7 8 9
Number of Clusters (K)

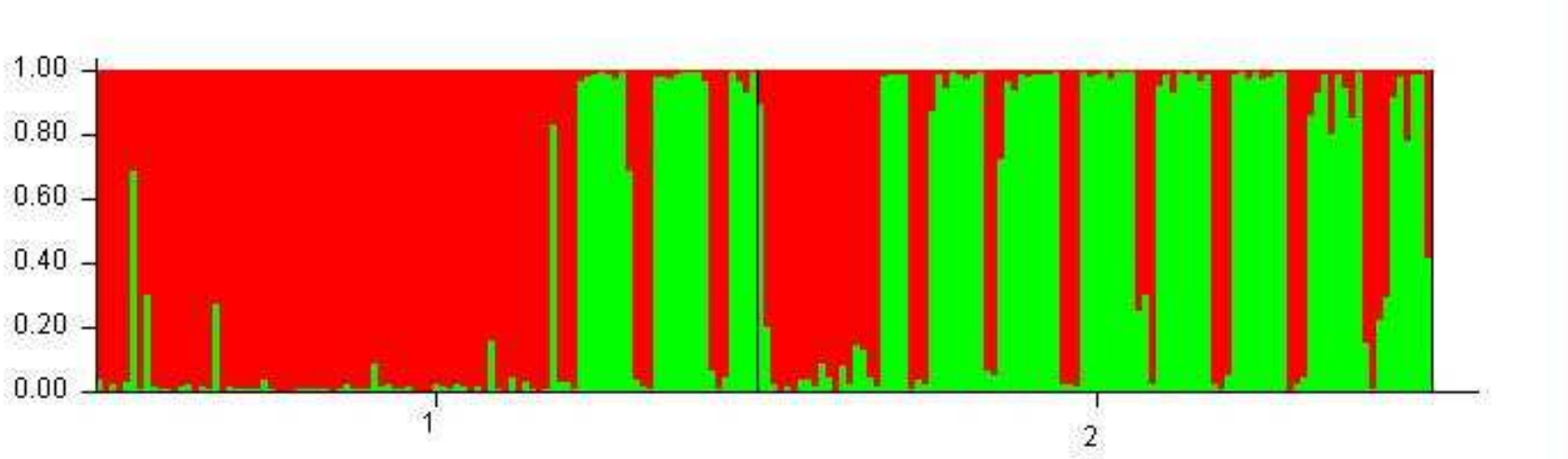


Figure 3-4. Bar plots showing Bayesian assignment probabilities using software STRUCTURE 2.2 for two clusters (K=2) (Cluster 1–north of Appalachian trail and Cluster 2 –south of Appalachian trail). The proportion of each bar that is red or green indicates the assignment probability of *Cornus florida* individuals to each of the two clusters, respectively.

Vita

Denita Hadziabdic was born in Sarajevo, Bosnia and Herzegovina on November 10, 1978. She attended and graduated from Prva Gimnazija, Sarajevo, Bosnia in 1997. Denita earned her B.S. in Agriculture-Horticulture in December of 2002 from Tennessee Technological University, Cookeville, TN. In 2005, Denita received M.S. in Plant Pathology from the University of Tennessee, Knoxville, TN under guidance and mentorship of Dr. Robert N. Trigiano. She received a Ph.D. in Plants, Soils, and Insects with concentration in Plant Pathology at the University of Tennessee, Knoxville, TN in 2010 after being mentored and supervised by Dr. Robert N. Trigiano.