5-2001

# Within and among population genetic variation in Fagus grandifolia within Great Smoky Mountains National Park

Ashley Lynn Brigham Morris

To the Graduate Council:

I am submitting herewith a thesis written by Ashley Lynn Brigham Morris entitled "Within and among population genetic variation in Fagus grandifolia within Great Smoky Mountains National Park." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Botany.

Mitchell B. Cruzan, Major Professor

We have read this thesis and recommend its acceptance:

Massimo Pigliucci, Jake Weltzin, Randy Small

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

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

To the Graduate Council:

I am submitting herewith a thesis written by Ashley Brigham Morris entitled "Within and among population genetic variation in *Fagus grandifolia* within Great Smoky Mountains National Park." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Botany.

Mitchell B. Cruzan, Major Professor

We have read this thesis and
recommend its acceptance:

Accepted for the Council:

Interim Vice Provost and
Dean of The Graduate School

# WITHIN AND AMONG POPULATION GENETIC VARATION IN *FAGUS GRANDIFOLIA* WITHIN GREAT SMOKY MOUNTAINS NATIONAL PARK

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Ashley Brigham Morris
May 2001

## DEDICATION

This thesis is dedicated to my entire family, who has been emotionally and financially supportive throughout all of my endeavors, academic and otherwise.

I don't care how poor a man is;

if he has family, he's rich.

~ M*A*S*H, Colonel Potter ~

## ACKNOWLEDGEMENTS

Research of any kind is always a group effort, and this work is no exception. I would first like to thank Greg Wiggins for introducing me to this study system almost three years ago when I worked as a research technician in the Department of Entomology at the University of Tennessee. His early assistance in identifying sampling sites and identification of beech bark disease symptoms was invaluable. I would also like to thank my thesis committee, Mitch Cruzan, Massimo Pigliucci, Jake Weltzin, Randy Small, and previously Hazel Delcourt, for their advice over the past two years. The discussions contained in this thesis are in no small part a result of the many conversations with my committee members.

I would like to thank all of those who gave their time to my field sampling efforts: Greg Wiggins, Chris Morris, Joe Moss, Stephanie Case, Dee Thonnard, Jason Nail, and John Cowden. I thank Jamie Estill, Gina Baucom and Courtney Murren for their discussions of molecular techniques, statistics, and their friendship. Shannon Datwyler was pivotal in the optimization of ISSR protocol. Matt Gitzendanner has been so helpful with both my thesis work and just questions in general during my time at Washington State University, and I thank him. I thank Markus Fischer for his discussions on the details of Harada and Iwasa's clonal identity probability.

I must thank the research staff at Great Smoky Mountains National Park, particularly Glenn Taylor and Mike Jenkins for their help in identifying and locating study sites matching the criteria that I specified. This research was funded by a Campbell

Fellowship for research in the Great Smoky Mountains through the Great Smoky Mountains Conservation and Natural History Associations.

Finally, without the support of my husband Chris this thesis would not have reached fruition. Our long-distance phone conversations have kept me going and have gotten me to the home stretch.

## ABSTRACT

In this study I use ISSR markers to assess levels of genetic variation within and among populations of *Fagus grandifolia* in Great Smoky Mountains National Park (GSMNP). I specifically test the following questions: 1) Do high elevation populations of *F. grandifolia* in GSMNP reproduce more frequently by root sprouting than by seed, resulting in fewer genets than ramets within a population? 2) Do low elevation populations of *F. grandifolia* in GSMNP reproduce more frequently by seed than by root sprouting, resulting in a relatively equal number of ramets and genets within a population? 3) Are high elevation populations of *F. grandifolia* in GSMNP genetically distinct from low elevation populations? and 4) Is an absence of beech scale (*Cryptococcus fagisuga*) infection on an individual tree in an otherwise infected stand correlated with genotype as a result of genetic resistance? The data collected in this research suggest that there is more genetic differentiation among high and low elevations than within high or low elevations sites, clonal reproduction is not limited to high elevation sites, and there is little evidence for genetic resistance in the trees sampled.

# TABLE OF CONTENTS

**Chapter I. Introduction to research** ..... 1

Paradigm of scale ..... 1
Inter-simple sequence repeats (ISSRs) ..... 4
Study system ..... 6
Research objectives ..... 12

**Chapter II. Analytical approaches in clonal population genetics** ..... 14

Defining clonal growth ..... 14
The costs of sexual reproduction ..... 15
Molecular markers ..... 17
Differentiating among genets ..... 20
Assessing size and distribution of genets ..... 26
Clonal identity probability ..... 27
Spatial autocorrelation analysis ..... 32
Future research ..... 33

**Chapter III. Population genetic structure in a clonal species, *Fagus grandifolia*, within Great Smoky Mountains National Park** ..... 34

Introduction ..... 34
Methods ..... 40
Results ..... 47
Discussion ..... 54

**Chapter IV. Investigating the potential for genetic resistance to beech bark disease in *Fagus grandifolia*** ..... 66

Introduction ..... 66
Methods ..... 71
Results ..... 77
Discussion ..... 78

**Chapter V. Conclusions** ..... 81

**Bibliography** ..... 84

**Vita** ..... 92

# LIST OF FIGURES

**Figure 1. The scale paradigm**.....2

**Figure 2. Inter-simple sequence repeats (ISSR) PCR**.....5

**Figure 3. Geographic distribution of *Fagus grandifolia***.....7

**Figure 4. Molecular marker efficiency**.....18

**Figure 5. Differentiating among genets**.....22

**Figure 6. Clonal identity probability**.....28

**Figure 7. Study site locations in Great Smoky Mountains National Park**.....42

**Figure 8. Results of locus resampling**.....50

**Figure 9. Mean clonal identity**.....51

**Figure 10. Genetic distance dendrogram of all sites sampled**.....56

# LIST OF TABLES

**Table 1. Study site elevations and sample sizes** ..... 43

**Table 2. Clonal diversity across sites** ..... 49

**Table 3. Genetic variation within and among sites** ..... 54

**Table 4. Nei's genetic distance across sites** ..... 55

# CHAPTER I

## INTRODUCTION TO RESEARCH

### The paradigm of scale

Our current knowledge of evolutionary patterns and processes in natural populations has led to the increasingly common approach of interdisciplinary research, in which it is often difficult to draw distinct lines defining any particular field of study. In addition, there has been an increasing focus on the relevant spatio-temporal scale necessary to describe changes in the flora, fauna, and landscape of interest. Delcourt and Delcourt (1988) describe the relationship between temporal and spatial scales as a way to hierarchically characterize environmental forcing functions, biological responses, and vegetation patterns (Figure 1). In their illustrative overview of what they refer to as the "operational scale paradigm", scale varies from micro- (site or population processes) to mega-scale (continental or global level processes). Although the boundaries defined in their scale paradigm are admittedly somewhat arbitrary, they provide a good starting point for successful research design.

My research directly encompasses both the micro- and meso-scales of the Delcourt scale paradigm. In this work, I sample genetic diversity within individual sites (or sub-populations) at the micro-scale, as well as the genetic variation among these sites across a altitudinal gradient on the landscape of the Great Smoky Mountains at the meso-

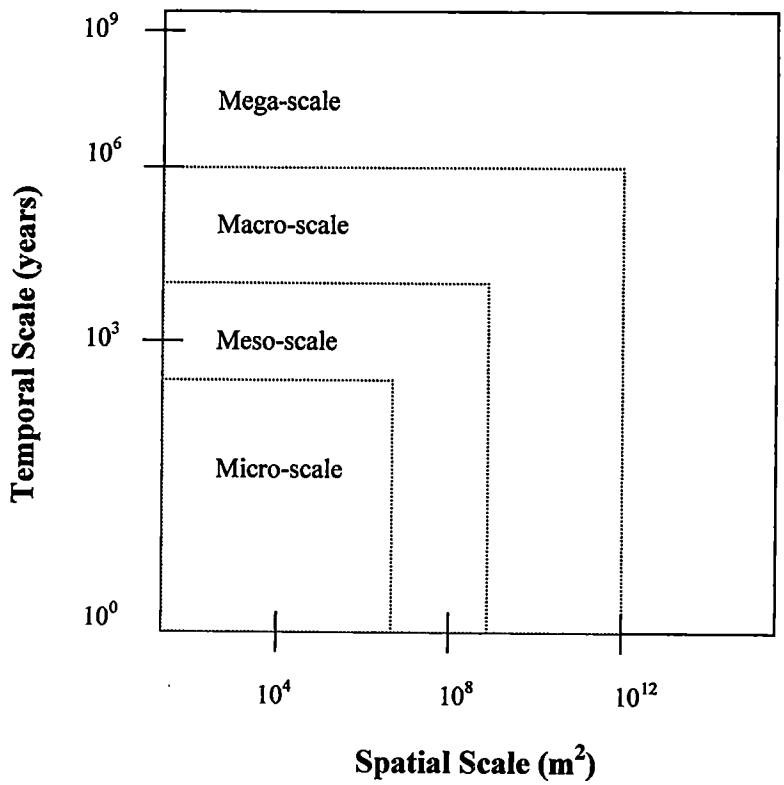


**Figure 1. The scale paradigm.** This illustration, redrawn from Delcourt and Delcourt (1988) represents the relationship between spatial and temporal scales relevant to research design. Micro-scale includes individual site (or population) studies; meso-scale includes multiple sites across some environmental gradient (landscape-level studies); macro-scale includes Quaternary studies, typically across a continent; and mega-scale encompasses the majority of geologic time.

scale. There are a number of reasons for using a multi-scale approach. First, the micro-scale domain as defined by the scale paradigm covers a temporal span of approximately 1 to 500 years, and a spatial span of $1m^2$ to 100ha (Delcourt and Delcourt 1988). At this level, population-specific processes such as seasonal trends, local to widespread disturbances of short duration, and succession can be identified. Population geneticists do not typically operate at this level. In fact, far fewer molecular studies have focused on within population genetic variation than among population genetic variation. Second, the meso-scale domain spans temporally from 500 to 10,000 years, and spatially from 1km wide to 100km wide. At this scale, direct inferences can be made about landscape-level processes such as patterns in species migrations and ecotone displacements across environmental gradients (e.g. latitudinal variation, moisture and nutrient regimes, and geological differentiation). It is at this level that population geneticists typically focus their research. The work in this thesis represents a study that incorporates both within-population structure and among-population differentiation with an emphasis on clonal structure, a combination that is rarely found in a single study. The primary biological question posed here is whether or not environmental conditions at high elevations in Great Smoky Mountains National Park result in a greater frequency of clonal reproduction in *Fagus grandifolia*. For this research I use inter-simple sequence repeats (ISSRs) to examine relationships among individuals and populations of *Fagus grandifolia* in Great Smoky Mountains National Park.

## Inter-simple sequence repeats (ISSRs)

Inter-simple sequence repeat markers were first described by Zietkiewicz et al. (1993), and shortly thereafter by Gupta et al. (1994) as a PCR-based fingerprinting technique designed from microsatellites. Microsatellites, also referred to as simple sequence repeats (SSRs), are short 2 to 5 base pair tandem repeats with a high levels of variation in the numbers of repeats among individuals and are found throughout the genomes of most organisms (Hillis et al. 1996). ISSR primers are designed from multi-nucleotide repeats (commonly di- or tri-nucleotide repeats) which are anchored in microsatellite regions using a variable number of repeated nucleotides (Gupta 1994; Wolfe and Liston 1998; Zietkiewicz 1993). These single primer reactions identify matching sequences on opposite strands of DNA, amplifying the genomic region flanked by two inversely oriented copies of the primer (Zietkiewicz 1993) (Figure 2). The length of the PCR product obtained varies with the distance between pairs of microsatellite loci. The most common means of visualizing ISSR banding patterns is with agarose gels stained with ethidium bromide, although there have been recent attempts in numerous labs to improve the accuracy and efficiency of these markers using automated sequencers. ISSRs appear as a series of bands within each lane on a gel, representing an individual multi-locus genotype. Each band is considered to be a locus, and these markers are inherited in a dominant Mendelian fashion. Thus, data are scored as present or absent, and an individual genotype is represented by a binary string of numbers. ISSR's, initially used in cultivar identification and genomic mapping of crop plants, have more recently been applied in natural populations to assess the relationships within

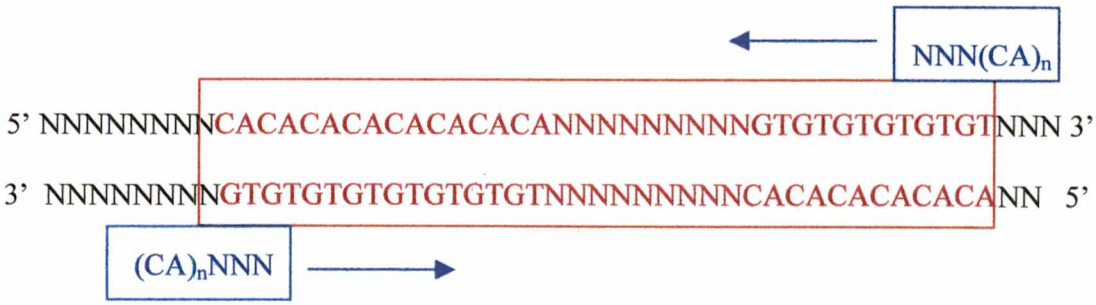


**Figure 2. Inter-simple sequence repeats (ISSR) PCR.** In this figure, redrawn from Zietkiewicz et al. (1994), a single primer designed from a $(CA)_n$ repeat anchored at the 5' end, is used to amplify genomic sequence flanked by two inversely oriented primers. In this illustration, primers are indicated in blue and the amplified region of the genome is indicated in red.

known hybrid complexes, clonal identity, and population structure (Wolfe and Liston 1998).

## Study system

### *Fagus grandifolia* Ehrh. (Fagaceae)

#### *Taxonomic history of Fagus grandifolia*

*Fagus grandifolia* (American beech) occurs over a broad geographic range in the eastern United States and Mexico, where it grows in a diversity of habitats (Figure 3). Although *F. grandifolia* is recognized as the only member of its genus in North America, There has been a continuing debate over division into potential subspecies that are based on the morphological and ecological variation observed across its range.

In 1951, W. H. Camp published what he referred to as a "paragenetic analysis" of beech. Based on such morphological characters as wood, bark, twigs, leaves, pubescence, flowers, and fruit. Camp concluded that there were three ecologically and morphologically distinct "types" of *F. grandifolia*: 1) white beech, which is found in low elevation river valleys; 2) red beech, which is found on well-drained slopes of the southern Appalachian Mountains on stabilized talus; and 3) gray beech, which is found at high elevations, as is the case with the beech "gaps" of the Great Smoky Mountains. Camp suggested that in some locations populations are easily divided into one of the three types. However, stands that are described as beech-maple climax forest are actually combinations of the three types of beech.

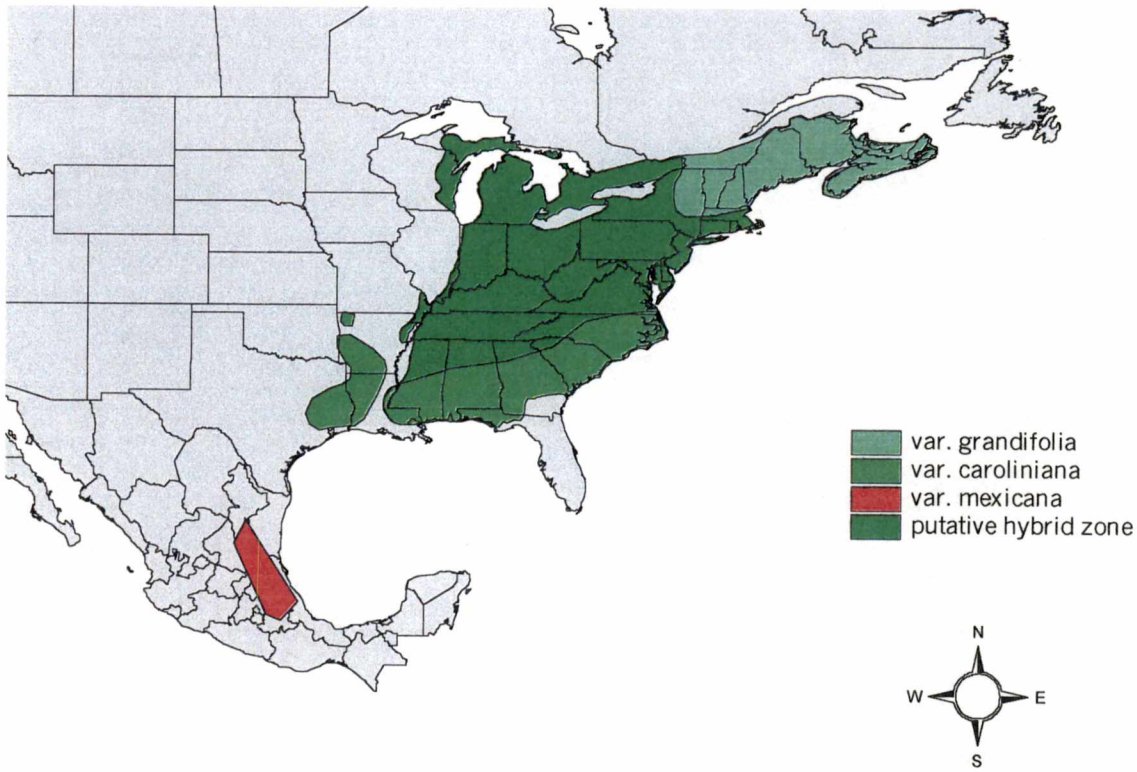


**Figure 3. Geographic distribution of *Fagus grandifolia***

Currently only two varieties of *F. grandifolia* are acknowledged north of Mexico: 1) var. *caroliniana* and 2) var. *grandifolia* (Gleason and Cronquist 1963; Radford et al. 1968; Shen 1992). According to Cooper and Mercer (1977), morphological variation within and among populations across North Carolina suggest that Camp's white race and var. *caroliniana* are equivalent, and Camp's red and gray beeches are equivalent to var. *grandifolia*. Shen (1992) describes a large amount of morphological variation within the species and notes that "purer" forms of var. *caroliniana* are found on the coastal plain and "purer" forms of var. *grandifolia* are found in the higher elevations of the southern Appalachian Mountains and in the northeastern portion of North America.

Current theory suggests that there were two distinct subspecies or species representing *Fagus* in North America prior to the Pleistocene. A series of north south oriented range expansions and contractions during the Pleistocene led to populations of mixed origin. Both Shen (1992) and Camp (1950) suggest that genetic and physical isolation between the original two subspecies was not sufficient to prevent hybridization. During the interglacial periods of the Pleistocene these hybrids inhabited areas previously unsuitable for the "purer" subspecies. Most of the beech in the central portion of the *F. grandifolia* range (which includes New York, Ontario peninsula, western Pennsylvania, and northern Ohio) is suspected to be of hybrid origin (Shen 1992).

***Reproductive strategies in Fagus grandifolia***

*Fagus grandifolia* is a shade tolerant species known to reproduce both sexually by seed and asexually by root sprouts (Rushmore 1961). The relative importance of these

two reproductive strategies is unclear. In the Great Smoky Mountains National Park (GSMNP), both Russell (1953) and Ward (1961) reported finding no beech seedlings in the park, suggesting that root sprouts are more important than are seeds for maintaining beech populations. However, Sain and Blum ( 1981) found beech seedlings in high elevation forests of the park, contradicting the findings of Russell and Ward. Jones and Raynal (1986b) suggested that root sprouts may be more important in the persistence of beech on a certain site, but seed dispersal likely drives establishment of new stands.

Numerous researchers have proposed that root sprouting is a product of injury to the root or tree (Held 1983; Jones 1986b; Melancon 1986; Whitney 1984). Jones and Raynal (1986b) listed mammals, freeze-thaw action, and human activity as potential sources for root injury. They stated that variation in the frequency of root injury might explain large differences in the number of sprouts found in stands with similar composition, as well as regional variation in relative abundance of sprouts reported by various researchers.

Held (1983) investigated beech reproduction strategies in the central portion of its range (including Wisconsin, Ohio, and Indiana), which Shen (1992) has suggested as a potential hybrid zone between two "purer" varieties. Held found no correlation between soil conditions (pH, calcium and magnesium levels) and the occurrence, abundance, survival, or growth of seedlings or root sprouts. However, he did find higher frequencies of seedling recruitment on northeast slopes and lower seedling recruitment on southern and western aspects. Held also found higher rates of root sprouting on south and west facing exposures. He concluded that there is segregation of root sprouts and individuals

of seed origin on separate slope aspects, and suggested that these results may be due to the more severe conditions of the southern exposures as described by Cantlon (1953). Cantlon noted an earlier start to spring activity on southern slopes, which may result in damage to plants from late frosts, which may kill any developing seedlings and eliminate them from the southern slopes. Freeze thaw damage to root systems may result in the production of root sprouts in place of sexually reproduced individuals. Understanding root sprouting in beech stands is a key factor for understanding patterns of sexual and asexual reproduction and the distribution of genetic variation among beech populations in GSMNP.

**Beech bark disease**

In recent years populations of *F. grandifolia* have been decimated across the species' range due to an insect mediated fungal pathogen complex known as Beech Bark Disease (BBD). The insect component of this complex was accidentally introduced into Nova Scotia in 1890 by way of infected nursery stock, and has since spread westward and southward to cover much of the current range of *F. grandifolia* (Ehrlich 1934; Houston et al. 1979). The woolly beech scale (*Cryptococcus fagisuga* Lindinger) infects beech bark, and through the damage it inflicts on trees provides a pathway for infection by *Nectria coccinea var. faginata* or *N. galligena*, two species of pathogenic fungi (Ehrlich 1934; Gavin and Peart 1993; Houston 1983).

The chronology of BBD has been well documented (Ehrlich 1934; Houston et al. 1979; Jones 1986a). The beech scale is parthenogenic and lays its eggs in mid- to late

summer before dying. The nymphs hatch in late summer or early fall and disperse locally on the same tree, with some nymphs being dispersed by wind and birds to other trees. The scale nymphs insert their stylets into the tree, molt before winter, and molt again the following spring. Adult scales secrete a white waxy substance that is easily detected by observers. The highest densities of scale occur in cracks and crevices of beech bark, which makes older trees with rougher bark more vulnerable to infection. Initial infestation by *C. fagisuga* provides a pathway for infection by *N. coccinea* var. *faginata*, an Ascomycete fungus. Signs of invasion often include tarry spots (which consist of a black or brown exudate on the bark of infected trees) and red perithecia, which is the sexual stage of the fungus. Mortality to the tree is a result of the vascular cambium becoming girdled, preventing the flow of nutrients throughout the tree.

As this disease spreads throughout the range of *F. grandifolia*, average mortality of beech within populations approaches 85% (Houston et al. 1979). Such high levels of mortality have been partially attributed to past logging practices (Houston et al. 1979). Historically, *F. grandifolia* was not considered a valuable timber species in northern hardwoods and was often left behind in logged stands. This has resulted in the rise of *F. grandifolia* as a major component in many northern forests. This overabundance of beech stems is thought to be a contributing factor in the widespread dispersal of beech bark disease. Damage incurred by logging practices is thought to have produced a niche for *C. fagisuga* on otherwise healthy stems (Houston 1982). In addition to the BBD complex, *F. grandifolia* is also subjected to *Xylococcus betulae*, which is another scale species that is often associated with young sprout thickets. Early infection by this species

may lead to later infection by BBD in individuals that would otherwise be resistant (Blozan 1995; Houston et al. 1979). Houston (1983) reported that although there are known predators of the beech scale, their effectiveness has been limited. Little is known about other potential predators or pathogens. Currently there is no known effective treatment for BBD.

BBD is described as having three chronological stages (Shigo 1972). The advancing front is the initial stage and is characterized by increasing populations of *C. fagisuga*. The second stage is referred to as the killing front, and is characterized by large populations of *C. fagisuga*, severe *Nectria* attacks, and high rates of tree mortality. Finally, the aftermath zone consists of gnarled trees, hypothesized to be resistant, defective sprout thickets, and endemic populations of the disease complex. The Great Smoky Mountains is currently encompassed by the killing front. Although some potential for genetic resistance to beech bark disease has been shown, no direct correlation between genotype and level of infection has been made.

## Research objectives

In this study I use ISSR markers to assess levels of genetic variation within and among populations of *Fagus grandifolia* in Great Smoky Mountains National Park (GSMNP). I specifically test the following questions: 1) Do high elevation populations of *F. grandifolia* in GSMNP reproduce more frequently by root sprouting than by seed, resulting in fewer genets than ramets within a population? 2) Do low elevation populations of *F. grandifolia* in GSMNP reproduce more frequently by seed than by root

sprouting, resulting in a relatively equal number of ramets and genets within a population? 3) Are high elevation populations of *F. grandifolia* in GSMNP genetically distinct from low elevation populations? and 4) Is an absence of beech scale (*Cryptococcus fagisuga*) infection on an individual tree in an otherwise infected stand correlated with genotype as a result of genetic resistance? The data collected in this research suggest that there is little or no genetic differentiation among high and low elevation sites, clonal reproduction is not limited to high elevation sites, and there is little evidence for genetic resistance in the trees sampled.

# CHAPTER II

## ANALYTICAL APPROACHES TO CLONAL POPULATION GENETICS

Clonal organisms have been of interest to ecologists and evolutionary biologists for decades and has been recognized as an important mode of reproduction in a broad range of organisms, including plants, fungi, and coral (Carbone et al. 1999; Coffroth and Lasker 1998; Ellstrand 1987). With the advent of molecular markers, the rapid and accurate genetic delineation of individuals has become possible, facilitating the development of more refined experiments on clonal populations. These methods allow researchers to determine the size and distribution of genets, and to assess the relative importance of sexual and asexual reproduction within and among populations. The statistical analyses involved in such studies are fairly new and the approaches taken vary with each new publication. My goal is to describe and critique a range of the methods that have been used in an attempt to identify the most informative and efficient methods for clonal molecular ecology.

### Defining clonal growth

Clonal growth, also called clonal reproduction, vegetative reproduction, and vegetative growth, is a form of asexual reproduction resulting in offspring genetically identical to the parent (Klimes 1997; Mclellan et al. 1997; Silvertown and Doust 1993). There are various mechanisms by which this occurs, but the final product is always a

potentially independent offspring capable of acquiring its own resources. Because 'clone' is used with various meanings, I will use the terms 'genet' and 'ramet' throughout the rest of this paper. A genet is a genetically unique unit comprised of varying numbers of individuals, or ramets. For example, in the case of Quaking Aspen (*Populus tremuloides*), each individual stem represents a ramet, and all stems that share the same genetic identity comprise a genet . In the case of a gregorian coral (*Plexara kuna*) a colony represents a ramet, and all colonies of identical genetic composition represent a genet (Coffroth and Lasker 1998). Ramets of the same genet may result from a variety of mechanisms. Quaking aspen undergoes root sprouting to reproduce asexually, whereas the coral previously mentioned may undergo fragmentation, budding, or fission to reproduce asexually. In all cases, ramets of the same genet will be nearly genetically identical, save some degree of somatic mutation, which is a consequence DNA replication errors during mitosis. Thus, in different species the concept of a clone may be difficult for those unfamiliar with the organism, but discussing populations in terms of genets and ramets should be familiar to all clonal ecologists.

## The costs of sexual reproduction

Clonal growth is relatively common across numerous plant taxa but less so among animals. Interestingly, most asexual species tend to be derived from sexual ancestors, indicating that asexual reproduction has immediate evolutionary advantages, but in the long run is less successful (Crow 1988). The primary advantage of sexual reproduction is considered to be the resulting genetic diversity, which allows for continuing evolutionary

change (Bell 1982; Crow 1988; Richards 1997). High levels of allelic diversity in many outcrossing populations allow adaptation to changing environments. Other advantages of sexual reproduction include the ability to combine beneficial mutations from different lineages and eliminate deleterious alleles through recombination (Crow 1988; Mclellan et al. 1997). Disadvantages to sexual reproduction include the expenditures of time and energy toward the production and rearing of offspring. The production of sexually produced offspring represents the two-fold disadvantage due to the transmission of only one set of chromosomes (compared with two under asexual reproduction) known as the cost of meiosis. In completely sexual populations, reduced fitness may occur due to segregation and recombination loads (Bell 1982; Crow 1988; Maynard Smith 1989).

Many of the costs associated with sexual reproduction are avoided through clonal reproduction. Since offspring are identical to the parent, the parent does not waste energy on the care of unfit zygotes. High-fitness genotypes can rapidly spread to dominate a population and are not lost through segregation and recombination. However, this may also become a disadvantage in rapidly changing environments if these genotypes are not phenotypically plastic. In general, clonal growth is considered a "low risk mechanism of proliferation" for genets of many flowering plants because of the minimal reproductive expenditure required by the parent.

In plants there is the added advantage of clonal reproduction because it is not limited by successful pollination or seasonal constraints as is sexual reproduction. Advantages of clonality in many plant species include rapid increases in population size and reproductive potential, the ability to capture resources over a heterogeneous

landscape (for species that retain connections among their ramets), and the ability to rapidly colonize and displace other organisms in new habitats (Silvertown and Doust 1993). Clonality allows for a fit individual to more rapidly spread throughout a population, increasing its potential to contribute to the gene pool in subsequent generations.

## Molecular markers

Because there are so many different kinds of molecular markers available, it is often necessary to compare the efficiency of different markers for the detection of genetic variation in different species. This is particularly true in studies of clonal organisms where it is necessary to obtain the most accurate estimate of individual identity to avoid underestimating the number of genets or overestimating genet size. Two methods for comparing marker efficiency that have been proposed are the assay efficiency index (*Ai*) (Pejic et al. 1998) and the genotypic index (*GI*) (McGregor et al. 2000) (Figure 4). These two indeces are fairly simple to calculate, although the assay efficiency index requires more information. The assay efficiency index is calculated by dividing the number of effective alleles detected (*Ne*) by the number of assays performed. If preliminary marker assessment based on published data were desired, this calculation may be less feasible because *Ne* is rarely reported in publications using DNA-based markers. The *GI* may be easier to apply as it is based on the mean number of genotype profiles generated per assay unit. In this case it may be possible to gather the data needed for calculation of the index from publications to pre-screen the efficiency of particular markers. However, if

Two methods for assessing genetic marker efficiency for the delineation of genets are the assay efficiency index (*Ai*) (Pejic 1998) and the genotypic index (*GI*) (McGregor 2000). *GI* may be more effective than *Ai* if there is an abundance of rare alleles in the population because *Ai* is based partly on the effective number of alleles ($N_e$) present. $N_e$ will tend to be low if there are a lot of rare alleles in the population, but rare alleles are actually more useful for identifying genets than are common alleles. The probability of two ramets having the same genotype by chance is much lower for rare alleles than for a common allele. *GI* is based on actual genotypes, which is the "unit" of interest in clonal studies.

***Assay efficiency index* (Ai)**

*Ai* combines the effective number of alleles identified per locus ($N_e$) and the number of polymorphic bands ($P$) detected in each assay:

$$Ai = N_e / P$$

**Figure 4. Molecular marker efficiency**

**Genotypic index (*GI*)**

*GI* is the mean number of genotype profiles generated per assay-unit for a marker system ($G_{AV}$) as a fraction of the number of individuals typed ($N$):

$$GI = G_{AV} / \text{total number of assays across all marker systems}$$

**Figure 4. Molecular marker efficiency (continued)**

published data were not available for the species of interest, and preliminary data collection is required, either index may be appropriate. Although neither index incorporated cost efficiency, this could be taken into account by dividing each index by the cost per assay. This slight modification may provide a more practical estimate of overall marker efficiency.

## Differentiating among genets

Prior to the application of molecular markers, there were few attempts to identify individual genets within a population and they were likely to be inaccurate, particularly in plants, where tracing connections among individuals was the typical approach taken to identify genets. However, this method cannot account for the possibility of physical disintegration of connections among ramets, which would result in overestimating genet number and underestimating genet size. Molecular markers now allow researchers to obtain individual genetic fingerprints, which may provide a more accurate depiction of genet distribution and population structure.

Although molecular markers typically provide more information than previously available, difficulties remain with the interpretation of such data. One problem associated with the molecular approach is the possibility of detecting two ramets sharing the same multi-locus genotype but belonging to two different genets. This is less likely to happen as the number of alleles and the number of loci included in the analysis increase. Several authors have presented methods for testing the probability of such an

occurrence. Aspinwall and Christian (1992) present a method for detecting the probability that ramets with the same multi-locus genotype actually belong to the same genet that is based on the frequency of each single-locus genotype (Figure 5). This probability, referred to as P bar, increases with larger numbers of alleles (at one locus or at separate polymorphic loci), resulting in greater confidence in genet determination.

Parks and Werth (1993) present a slightly different approach that requires two separate calculations to determine the probability of encountering two identical genotypes resulting from independent zygote formation (Figure 5). The first equation simply tests if ramets with identical genotypes represent the same genet ($p_{gen}$). The second equation calculates the probability of a second encounter of the same genotype ($p_{se}$). The latter equation is for more distantly spaced ramets, although distance is never explicitly defined in either instance.

Berg and Hamrick (1994) present yet another approach to identifying individual genets (Figure 5). They calculate two separate parameters: 1) the probability of randomly drawing two identical genotypes from a population and 2) the proportion of genotype pairs that differ by two or more loci. The second calculation indicates the number of pair-wise comparisons that would have to be mis-scored at two or more loci to wrongly identify two distinct genotypes as members of the same genet. With increasing numbers of loci, this proportion should increase.

One problem associated with any genetic marker approach for the identification of genets is the possibility of detecting two ramets sharing the same multi-locus genotype that actually represent two different genets. Several authors have presented methods for testing the probability of such an occurrence.

**Aspinwall and Christian (1992)**

Empirical calculation of the average probability of ramets with the same multilocus genotype belonging to the same genet - calculated for each population:

$$\overline{P} = 1 - \frac{\left[\sum_{Q=1}^{N}\prod_{D=1}^{M}\left[\frac{X_{DQ}}{P_D}\right]\right]}{N}$$

where $Q$ is the particular individual in a sample, $N$ is the total number of individuals in a sample, $M$ is the number of polymorphic loci in the sample, $D$ is an individual polymorphic locus, $X_{DQ}$ is the number of individuals in the sample of the same $D$th locus genotype as the $Q$th individual, and $P_D$ is the total number of individuals examined for each polymorphic locus. The relationship between the

**Figure 5. Differentiating among genets**

average probability of ramets with the same multilocus genotype belonging to the same genet, or P bar, and the numbers of polymorphic loci for a population are highly correlated.

**Parks and Werth (1993)**

Estimates probability that identical genotypes could result from independent formation of zygotes. Formula 1 applies to clusters of two or more closely spaced ramets:

$$P_{gen} = \left(\prod p_i\right) 2^h$$

where $P_{gen}$ is the probability that a zygote acquires a given diploid genotype, $p_i$ is the frequency in the population of each allele represented in the genotype, and $h$ is the number of loci that are heterozygous.

Formula 2 applies to distantly spaced ramets with identical genotypes:

$$p_{se} = 1 - \left(1 - P_{gen}\right)^G$$

where $p_{se}$ is the probability that the genotype is NOT encountered subtracted from

**Figure 5. Differentiating among genets (continued)**

unity, $G$ is the number of separate genets genotyped, and $n$ is the number of separated fragments with identical genotype to some previously encountered ramet.

**Berg and Hamrick (1994)**

Calculates the probability of randomly drawing two identical genotypes by calculating the sum of squares of the expected genotype frequencies and obtaining the multilocus probability by multiply all of the individual-locus probabilities. Calculates the proportion of genotype pairs which differed by two or more loci for all $N(N\text{-}1)/2$ possible pairs of $N$ genotypes. If this proportion equals 100% then all comparisons two or more loci would have to be mis-scored in order to incorrectly identify two distinct genotypes as members of the same genet.

**Figure 5. Differentiating among genets (continued)**

There are several assumptions made with these calculations that must be considered. First, all of these methods assumes that each different genotype encountered results directly from a zygote (no somatic mutation). Second, mating is random. This will be of particular concern for populations of sedentary organisms where ramets remain in close proximity to one another. For example, if there were very little sexual reproduction, then production of new ramets will be dominated by the most successful genets. Thus in either case, the assumption of random mating may be a difficult one to support. Furthermore, a sexually reproduced individual will get half of its alleles from the local dominant genet. The third assumption is that there is independent assortment of loci and no linkage disequilibrium. Assumption 1 (no somatic mutation) over the time-scale of an individual's life is fair in that mutation rates at individual loci are small, and assumption 3 (linkage disequilibrium) can be tested (although this could be difficult with many species that are long-lived or difficult to breed). Assumption 2 (random mating), although testable with adherence to Hardy-Weinberg, may be a more difficult assumption to apply in populations with local mating and small effective population size, which is probably particularly common in clonal populations.

With the exception of Aspinwall and Christian's P bar, these calculations assume that all genotypes are equally possible. Particularly in clonal populations, the genotypes present will represent only a non-random subset of all those possible. Thus, P bar may result in a more accurate approximation of genet identification than either of the other two calculations. Unfortunately, there is not enough appropriate data available in the literature to compare these different approaches. When preparing to analyze a new clonal

data set, it may be useful to complete all three calculations and compare the results to evaluate the utility of each measure.

## Assessing size and distribution of genets

In recent years clonal population studies have relied heavily on statistics originally applied by Ellstrand and Roose (1987) to estimate the size and distribution of genets. The simplest calculation of these is the "proportion distinguishable" (*PD*). This is calculated by dividing the number of genets detected by the sample size (i.e., the number of ramets). This index produces a value that ranges from $1/N$ to 1, where a value of 1 indicates each individual sampled represents a unique genotype and a value of $1/N$ indicates a single genet is present. Ellstrand and Roose suggest this is a good overall descriptor of genet diversity both within and among populations.

Ellstrand and Roose (1987) also tested the sensitivity of their methods to the detected levels of genetic diversity. Essentially they estimated the number of loci necessary to gain adequate levels of variation to appropriately distinguish genets. Kendall's coefficient of rank correlation ($\tau$) was used to measure the relationship between the total number of genotypes detected and the two variables of sample size and number of characters (loci) scored. The same analysis was conducted using *PD* in place of the total number of genets detected. This approach allowed the authors to show that the number of loci used had a greater affect on genetic diversity than d sample size. The application of such power analyses is an excellent method for determining when additional data collection is necessary.

Finally, Ellstrand and Roose suggest that individual genotypes can be local or widespread. A local genotype is one that occurs in only one population whereas a widespread genotype occurs in more than 75% of the populations sampled. These measures can provide a clear estimate of genotype size and distribution as well as a way to approach the issue of long-distance versus short-range dispersal of asexual propagules.

## Clonal identity probability

The simplest measure of the importance of clonal reproduction is *PD*, as discussed above. However, clonal identity probability (referred to as *F(r)*) has proven effective in assessing the relative importance of sexual and asexual reproduction within and among populations, and provides more information than does *PD* (Harada and Iwasa 1996; Harada et al. 1997) (Figure 6). This method, which depends on the area and size of clusters, is applicable to any mapped genotype data and is not limited to evenly spaced or linear collections. There are three points to address regarding this method. First, the authors note that their model does not account for spatial heterogeneity or local adaptations that could explain spatial clumping of genotypes, suggesting that the importance of such causal relationships requires more study. Second, as the number of loci used increases, the reliability of clonal identity increases, as has been indicated by previous approaches discussed in this paper. Finally, *F(r)* is calculated as an equilibrium distribution. With temporally spaced samples, this method can be used to illustrate clonal succession in response to local changes in environment. Thus, any single measure of *F(r)*

Clonal identity probability *F(r)*, the probability that two randomly chosen plants separated by distance of *r* belong to the same clone. The first step is to estimate the relationship between all possible ramet pairs. If the two ramets are members of the same genet, then the pair is assigned a clonal identity of 1; if the two ramets are not members of the same genet then the pair is assigned a clonal identity of 0. This data is then fitted to the following function using least squares regression:

$$F(r) = k\,(e^{-cr}/sqrt\ r)$$

where *F(r)* is clonal identity, *r* is pairwise distance, and *k* and *c* are intrinsic constants of the model. By fitting this function to data with non-linear regression, the constants c and k can be determined. The constant *c* is the exponential rate of decay in *F(r)* and can be used to estimate the ratio of sexual to asexual reproduction:

$$c = \sqrt{\frac{2u}{v\sigma^2}}$$

**Figure 6. Clonal identity probability**

where $\sigma^2$ is the range of dispersal by vegetative propagation, $u$ is the probability of being produced by seed, and $v$ is the probability of having been produced by vegetative propagation. The rate of decrease of the decay constant $c$ reflects the relative importance of sexual and asexual reproduction. The larger $c$ is, the greater the influence of seed production over the population.

In theory, the *F(r)* is summarized as follows: 1) two adjacent samples are more likely to be members of the same genet than are two distantly spaced samples; 2) *F(r)* declines more slowly if the sampled population includes only a few large genets as opposed to many small genets; average genet size should be larger if vegetative reproduction is more important than sexual reproduction; 4) thus the rate of decrease of *F(r)* (denoted as $c$) is indicative of the relative importance of the two modes of reproductions The actual calculations are dependent on three demographic parameters: 1) seed recruitment rate ($u$), 2) vegetative propagation rate ($v$), and 3) the variance of vegatatively spreading distance ($\sigma^2$). Of these, $\sigma^2$ will be easiest to estimate in the field, while the other variables can be estimate from the probability calculation itself. The authors note that it is only possible to estimate the relative importance of the two modes of

**Figure 6. Clonal identity probability (continued)**

reproduction from $F(r)$ when reproduction is mainly clonal because the slope ($c$) is sensitive to the ratio of the two modes of reproduction.

**Figure 6. Clonal identity probability (continued)**

only estimates the relationship of the two modes of reproduction at a given moment in time.

The clonal identity probability method has been applied in a limited number of clonal plant population studies to examine the relative importance of clonality in those systems (Holderegger et al. 1998; Schlapfer and Fischer 1998). Schlapfer and Fischer (1998) used the clonal identity probability to estimate the rates of sexual and vegetative recruitment in different populations of *Brachypodium pinnatum* (Family), a perennial grass species that reproduces both sexually by seed and asexually by underground rhizomes. In this study, the mean clonal identity was plotted against the distance between tillers to reflect the diminishing likelihood of clonal identity with increasing distance, which they suggest reflects spatial aggregation of genets . In a study of population structure in *Anemone nemorosa* (Ranunculaceae), a perennial woodland herb, Holderegger et al. (1998) used the same approach and found a much more severe decline in clonal identity with increasing distance. When plotted against distance, the clonal identity probability approaches zero within 2m distance between ramets, suggesting extreme aggregation of genets. Such plots can be compared among populations to assess the level of aggregation in different microhabitats.

In addition, the decay constant *c* (Figure 6) can be used to compare the percentages of sexual recruitment within and among different populations. Previous work on similar species had shown that 3% sexual recruitment is sufficient for the maintenance of genetic variation in clonal populations (Holderegger et al. 1998). Holderegger et al. (1998) identified values of *c* ranging from 2.8% to 6.3% indicating that these populations

have a sufficient level of sexual reproduction to maintain genetic variation. Thus the methods introduced by Harada and Iwasa (1996) present an empirical way to assess the importance of clonal reproduction in natural populations without the necessity of more time-intensive demographic studies.

## Spatial autocorrelation analysis

Spatial autocorrelation, which was originally a tool for geographers, was first introduced to biologists by Sokal and Oden (1978a; 1978b) as a way to analyze microgeographic variation of allozyme frequencies. The patterns of dependence between observations can be assessed by examining a correlogram, which displays the association between pairs of individuals at increasing distances.

Spatial autocorrelation has been applied to a number of clonal systems as a way to estimate clone size and distribution (Hossaert-McKey et al. 1996; Ueno et al. 2000; Weller et al. 2000). Hossaert-McKey et al. (1996) use these methods to identify spatial genetic structure in *Lathyrus sylvestris* (Fabaceae), a widespread perennial legume that reproduces both sexually by seed and asexually by underground rhizomes. This study was unique in that spatial autocorrelation was applied to age classes to estimate temporal variation of clonal structure. The authors found a spatial correlation between the frequency of clonemates and age of population. From this they suggested that the older portion of the population was initially established by one or a few founders and has been maintained primarily by clonal growth. Thus, spatial autocorrelation can be a useful tool for identifying evolutionary processes in clonal systems.

## Future research

The analytical methods reviewed here provide a great deal of information specific to clonal populations. Each method can be applied to different types of data, answering a variety of questions, and thus may not be appropriate in all situations. More comparative studies are needed to allow us to assess the advantages and disadvantages of each method. Perhaps the best approach is one in which multiple markers and multiple sampling strategies are used. In addition, work with familiar data sets in which the genetic structure is well known will be the best way to test the efficiency and reliability of these different approaches. Beyond that, to truly understand clonal systems, future research must include multiple seasons of data collection and analysis. Clonal populations can be quite dynamic, and one moment in time cannot provide a thorough understanding of the processes leading to patterns of genet distributions within and among populations.

# CHAPTER III

## POPULATION GENETIC STRUCTURE IN A CLONAL TREE SPECIES, *FAGUS GRANDIFOLIA* EHRH. (FAGACEAE), WITHIN GREAT SMOKY MOUNTAINS NATIONAL PARK

### Introduction

In the broadest interpretation of the concept of clonality, all plant species are essentially potentially clonal. Modular growth allows for the propagation of an individual from cuttings of a single module (e.g., a leaf or branch). However, the strict sense of the term clonality encompasses those plants that can asexually produce offspring genetically identical to the parent (Groenendael et al. 1997). The difference between pure modular growth and clonal (also termed vegetative) reproduction is that ramets have the potential for physiological independence (Groenendael et al. 1997). Clonality has been documented in a number of plant species, most of which retain some level of sexual reproduction (Ellstrand 1987). Many researchers have assumed that clonal reproduction limits population genetic variation, but numerous molecular studies have provided evidence that negates this assumption (see review by Ellstrand and Roose 1987). When assessing patterns of genetic diversity in plant populations with both sexual and asexual reproduction, it is necessary to consider the relative role of each strategy in determining the distribution of genetic variation within and among populations.

Advantages of clonality in many plant species include rapid increases in population size and reproductive potential, the ability to capture resources over a heterogeneous landscape (for species that retain connections among their ramets), and the

ability to rapidly colonize and displace other organisms in new habitats (Silvertown and Doust 1993). Clonality allows for a fit individual to more rapidly spread throughout a population, increasing its potential to contribute to the gene pool in subsequent generations. High-fitness genotypes can rapidly spread to dominate a population and are not lost through segregation and recombination. However, this may also become a disadvantage in rapidly changing environments if these genotypes are not phenotypically plastic, particularly in response to rapidly spreading disease.

An important implication of clonal population studies has been the inability to account for lost connections between ramets. Such previously unavoidable sampling errors potentially overestimate the number of clones (the number of genets) in a population, as well as underestimate clone size (the number of ramets per genet). Recent advances in molecular genetics have provided the tools needed for accurate genet delineation, even in species that are subject to clonal disintegration. Molecular markers including allozymes, RAPDs, microsatellites, AFLPs, and ISSRs allow clonal ecologists to more accurately describe population sub-structuring without the necessity of genet excavation (Burke et al. 2000; Fay et al. 1999; Reusch et al. 2000; Suzuki et al. 1999; Tani et al. 1998). The primary objective of the research presented here is to determine the relative importance of sexual and asexual reproduction in *Fagus grandifolia* at high and low elevations within Great Smoky Mountains National Park.

### *Fagus grandifolia* as a model system

In this study, I examined the importance of clonal reproduction in shaping population genetic structure in *Fagus grandifolia* Ehrh. (Fagaceae) and the potential for genetic resistance to beech bark disease. *Fagus grandifolia* is a shade tolerant deciduous tree known to reproduce both sexually by seed and asexually by root sprouts (Rushmore 1961). In Great Smoky Mountains National Park (GSMNP) the importance of these two strategies is unclear. *Fagus grandifolia* is relatively abundant at high and low elevations in GSMNP, but is less common at mid elevations (Russell 1953; Whittaker 1956). High elevation populations have been called "beech gaps" or "beech orchards" in reference to the high densities of small trees, often no greater than 12 inches in diameter and 20 feet in height (Whittaker 1956). In contrast, low elevation populations are typically far less dense, and trees may grow to be quite large (up to 120 ft tall under optimal conditions, but typically average 60 to 80 feet, Rushmore 1961). The small size of high elevation stems is not likely a product of age; Blozan (1995) reports high elevation stems ranging in age from 50 to 220 years.

### Morphological variation among high and low elevations

There are at least three hypotheses that could explain the differences observed between high and low elevation stands of *F. grandifolia* in GSMNP. First, stem density at high elevations may be a result of predominately clonal reproduction. Both Russell (1953) and Ward (1961) reported finding no beech seedlings in high elevation beech gaps

in GSMNP, suggesting that root sprouts are more important than are seeds for maintenance of these populations. However, Sain and Blum (1981) found beech seedlings in the high elevation forests, contradicting the previous study by Russell and Ward. Jones and Raynal (1986b) suggest that root sprouts may be more important in the persistence of *F. grandifolia* on a given site, but seed dispersal likely drives colonization of new sites.

***Clonal reproduction***

Numerous researchers propose that root sprouting is a product of injury to the root or tree (Held 1983; Jones 1986b). In the high elevations of GSMNP, sources of such injury include more extreme winter weather, which results in frequent frost-heaving that damages roots and stems. In addition, soil disturbance by wild boars (*Sus scrofa*) has been implicated in root injury to *F. grandifolia*, resulting in increased frequency of root sprouting (Jones 1986b; Lacki and Lancia 1986). Annual spring migrations of high concentrations of feral hogs to the high elevation beech gaps have been documented by radio-telemetry studies (Singer et al. 1981). However, it is important to note that feral hogs have only been present in GSMNP since the 1940's. Therefore, the disturbance resulting from these animals has only contributed to population structure for a relatively short time. In general, large differences in the number of root sprouts found in stands of similar composition have been attributed to variation in the frequency of root injury (Jones 1986a). If this were true, then greater frequencies of root injury as a result of freeze-thaw action and herbivore disturbance (which is more common at high elevations

as a result of spring migrations for beechnut foraging) should induce greater frequencies of root sprouting at high elevations than at low elevations where such conditions are less severe.

***Independent evolutionary histories***

The second hypothesis is that high elevation populations of *F. grandifolia* in GSMNP are genetically distinct from low elevation populations as a result of independent evolutionary histories. Camp (1950) distinguishes between three types of beech with within *F. grandifolia*: white beech, found mainly at low elevations along the Gulf and Atlantic coastal Plains; red beech, which is associated with the Piedmont region north of the Coastal Plain; and gray beech, which is a representative of high elevation spruce-fir associations in the Southern Appalachians, and is found again above 2000 feet in the Adirondack Mountains and west to the Great Lakes. Some have suggested that the Southern Appalachians served as a glacial refugium during the Pleistocene serving as the primary source of many of the plant taxa that recolonized the northeastern United States after glacial retreat (Braun 1950; Sharp 1972). The debate continues as to whether habitat diversity in the Southern Appalachians during the last glacial period was sufficient to serve as refugia to deciduous tree species (Delcourt and Delcourt 1984). Shen (1992) lends some support to this theory in his monograph of the genus *Fagus* by suggesting that populations of *F. grandifolia* in the high elevations of the Southern Appalachians are most morphologically similar to those populations in New England than to those located on the Atlantic and Gulf Coastal Plains. Coastal Plain populations

are suspected to have survived glaciation in southern refugia along the uplands of the Mississippi River Valley (Delcourt and Delcourt 1984; Shen 1992). Thus the potential exists for the high elevation beech gaps in GSMNP to be genetically distinct from lower elevation populations as a result of being derived from separate Pleistocene refugia.

***Phenotypic plasticity***

The third hypothesis is that phenotypic plasticity has resulted in the differences observed among high and low elevation stands. Although this hypothesis cannot be directly tested without common garden experiments and detailed field studies, estimates of genetic relatedness among stands using molecular techniques can provide evidence that will help support or refute this claim. For example, if genetic marker assays demonstrate that stands at the same elevation are as closely related as between the low and high elevation regions then environmentally induced differences becomes more likely. From these estimates conclusions regarding the roles of asexual and sexual reproduction, isolation by distance, and local adaptation can be inferred.

**Research objectives**

In this study I use inter-simple sequence repeat (ISSR) markers to understand population structure in *F. grandifolia* at high and low elevations in GSMNP. Specifically, I address the following questions related to patterns of clonal growth and the distribution of genetic variation: 1) What is the relative importance of clonal reproduction high elevation populations and low elevation populations in GSMNP? 2) Does genetic

variation within and among populations of *F. grandifolia* support the theory of separate northern and southern refugia during the Pleistocene? My results indicate that there is little genetic differentiation among high and low elevation sites, and clonal reproduction is not as common as previously suspected.

## Methods

### Study Organism

American beech (*Fagus grandifolia* Ehrh. (Fagaceae)) occurs over a broad geographic range in the eastern United States and Mexico, where it is found in a diversity of ecological habitats (Figure 3). Beech is a shade tolerant, late successional, mesophytic tree species associated with at least 20 forest cover types in North America (Rushmore 1961). Beech is a monoecious species, with male flowers occurring in catkins and female flowers occurring in clusters. An individual tree can produce a substantial seed mast at 40 years of age, and good seed crops occur at intervals of 2 to 8 years. The primary dispersal agent for beechnuts is the bluejay, which can transport seeds across several kilometers.

### Study Site

Great Smoky Mountains National Park (GSMNP) is an International Biosphere Reserve that encompasses over 521,000 acres in Tennessee and North Carolina and is the most frequently visited national parks in North America with over 9,000,000 tourists annually. In GSMNP, elevation ranges from 800 to 6643 feet, resulting in a diversity of

forest cover types, including spruce-fir forest (above 4500 feet), northern hardwood forest (3500 to 5000 feet), pine-oak forest (dry ridges), hemlock forest (along stream banks), and cove hardwood forest (low valleys). Beech is a minor component in cove hardwoods, hemlock, and pine-oak forest but a major component in northern hardwood and spruce-fir forests in the park.

**Data collection**

*Field sampling methods*

Data were collected from five sites that were selected to be representative of both low and high elevation populations and based on long term projects managed by GSMNP staff (Figure 7). Three sites, Greenbriar (GRB), Grassy Branch (GRS), and the Chimneys (CHM), will be referred to collectively as the low elevation sites (LOW; Table 1). LOW sites are all below 3500 feet in elevation and are associated with hemlock (*Tsuga canadensis*) forests. Two sites, Balsam Mountain Road (BMR) and Double Spring Gap (DSG), will be referred to collectively as the high elevation sites (HIGH; Table 1). HIGH sites are all above 4500 feet elevation and are representatives of "beech gap" forests, which are located within topographic gaps in areas otherwise dominated by spruce-fir forests. Due to differences in stem densities at all of the sites, different sampling strategies were employed. At GRB and GRS, a 10m x 20m plot was established, within which all *F. grandifolia* stems were sampled. At CHM, a 10m x 10m area of a 20m x 20m permanent study plot was established within which all *F. grandifolia* stems were sampled. At BMR and DSG, a 90m transect was established,

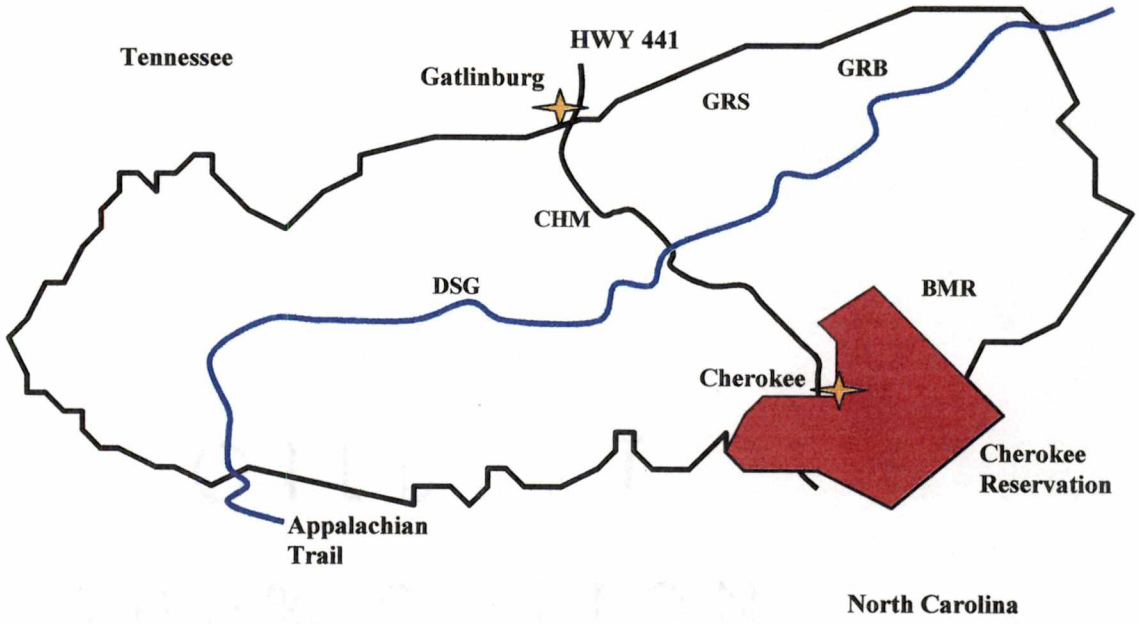


**Figure 7. Study site locations within Great Smoky Mountains National Park.** Study site locations are indicated by site name abbreviations (DSG = Double Spring Gap; CHM = Chimneys; GRS = Grassy Branch; GRB = Greenbriar; BMR = Balsam Mountain Road). The Appalachian Trail is indicated in blue, and the Cherokee Reservation is indicated in red. Cities are indicated by yellow stars.

**Table 1. Study site elevations and sample sizes.**

| Site | Elevation | N |
|---|---|---|
| Balsam Mountain Road (BMR) | 5000 | 30 |
| Chimney Tops (CHM) | 3600 | 19 |
| Double Spring Gap (DSG) | 5200 | 35 |
| Grassy Branch (GRS) | 2000 | 16 |
| Greenbriar (GRB) | 1800 | 23 |

along which samples were collected at 10m intervals. At each sampling interval the five closest stems were sampled.

Leaf material was collected from each stem sampled, placed in separate 1.5-mL microcentrifuge tubes, and stored on ice while in the field. Upon returning to the lab, all samples were snap frozen in liquid nitrogen and stored at - 70° C.

***DNA Extraction***

Total genomic DNA was extracted from each sample using a protocol adapted from Edwards et al (1991) as described in Martin and Cruzan (1999). Leaf material was ground in 100 ul of extraction buffer (200 mM Tris HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS) for 15 seconds using disposable grinders. An additional 400 ul of extraction buffer was added to this mixture and ground again. Each sample was then centrifuged at 13000 rpm for 2 minutes. Following centrifugation, 400 ul of supernatant was extracted and placed in a sterile 1.5 ml microcentrifuge tube, to which 400 ul of isopropanol was added. After standing at room temperature for 2 minutes, samples were centrifuged at 13000 rpm for 10 minutes. All liquid was then poured off, and the samples were allowed to dry. The remaining DNA was then resuspended in 100 ul of $dH_20$. Each sample was cleaned by binding the DNA to DEAE-cellulose as described by Marechal-Drouard and Guillemaut (1995) and then stored at –20° C.

### *ISSR Protocol*

ISSR primer 810 ($(GA)_8T$) was selected from UBC primer set number 9 based on the brightness and consistency of the bands produced during screening. Single-primer reaction conditions were carried out in 15 ul reactions as follows: 2.5 mM MgCl2, 200 μM dNTPs, 1 unit of *Taq* polymerase, 0.10 uM primer, and 0.5 ul DNA. The thermal cycler profile was adopted from Huang and Sun (2000) and is as follows: 1 cycle of 94° C for 5 minutes, followed by 45 cycles of 94° C for 5 seconds, 50° C for 45 seconds, 72° C for 1.5 minutes; and a final 7 minute extension at 72° C. PCR products were separated on 2% agarose gels in 1X TBE buffer until bromphenol blue marker migrated 10 cm from the origin. Gels were stained with ethidium bromide and were documented digitally using Kodak 1D Biomax software. Bands with the same molecular weight were treated as identical loci. A data matrix was compiled in which band presence was scored as 1, band absence was scored as 0. Ambiguous bands were excluded from the analysis.

## Data analysis

### *Data efficiency*

Two methods were used to determine if more data were necessary to detect all possible variation. The first method, introduced by Aspinwall and Christian (1992) calculates the average probability of ramets with the same multi-locus genotype belonging to the same genet:

$$\overline{P} = 1 - \frac{\left[\sum_{Q=1}^{N}\prod_{D=1}^{M}\left[\frac{X_{DQ}}{P_D}\right]\right]}{N}$$

where $Q$ is the particular individual in a sample, $N$ is the total number of individuals in a sample, $M$ is the number of polymorphic loci in the sample, $D$ is an individual polymorphic locus, $X_{DQ}$ is the number of individuals in the sample of the same $D$th multi-locus genotype as the $Q$th individual, and $P_D$ is the total number of individuals examined for each primer. This probability and the number of polymorphic loci in a population are highly correlated, such that as the probability approaches 1, additional assays will result in diminishing returns.

In the second method varying numbers of loci (1 through 24) were randomly resampled to infer the effects of increasing the number of markers on the number of genets detected. The number of genets was calculated for one thousand replicates of each data set size and these data were used to examine the relationship between number of loci in the data set and the number of genets detected. I assumed that an asymptotic curve (i.e., the number of new genets detected approaches zero as the number of loci in the data set is reached) would indicate that adding additional loci would not result in an appreciable change in the total number of genets detected.

***Clonal diversity***

Clonal diversity was estimated using an index presented by Ellstrand and Roose (1987). As an overall descriptor of diversity, the 'proportion distinguishable' (PD = G/N, where G is the number of genets and N is the total number of ramets) was calculated for each site. The relative importance of sexual and asexual reproduction was estimated by calculating mean clonal identity (Harada et al. 1997) at 5 meter intervals for each site

(Schlapfer and Fischer 1998). This was done by comparing all possible pairs of ramets within a site. If the pair shared the same multi-locus genotype, a clonal identity of 1 was assigned. If the pair did not share the same multi-locus genotype, a clonal identity of 0 was assigned. The mean clonal identity was then calculated for each of 6 distance classes of 5 meters each. The relationship between mean clonal identity and distance was then plotted for each population.

### *Genetic diversity*

Gene diversity within and among populations was estimated using the program POPGENE (Yeh et al. 1997), which handles both codominant and dominant marker data. This program was used to calculate Nei's (1987) gene diversity statistics in subdivided populations ($G_{ST}$), diversity among groups ($G_{CS}$), and Nei's genetic distance. The number of polymorphic loci and percent polymorphism were calculated for each site.

## Results

### Data efficiency

A total of 123 ramets were assayed for genetic variation across five sites (Table 1). The mean number of ramets per site was 24.6, ranging from a minimum of 16 to a maximum of 35. One ISSR primer revealed 24 loci, 23 of which were polymorphic. One hundred unique genets were identified across all sites. The mean number of genets per site was 20, ranging from a minimum of 8 to a maximum of 34.

Analyses of marker efficiency indicate the genet detection was close to saturation. Aspinwall and Christian's Pbar is the average probability that ramets with the same multi-locus genotype belong to the same genet. This value was 0.99 across all sites, and was never less than 0.98 within a site (Table 2). Resampling of loci indicated that the relationship between the number of loci used and the number of genets detected was asymptotic (Figure 8). Both methods indicate that a sufficient number of polymorphic loci were obtained to assess clonal genetic diversity within and among sites.

**Clonal diversity**

There were substantial levels of clonal diversity both within and among sties of American beech (Table 2). The mean percent polymorphism within sites is 80%, ranging from a minimum of 45.83% to a maximum of 91.67%. The total percent polymorphism across all sites is 95.8%. The proportion distinguishable (PD) ranges from a minimum of 0.42 (CHM) to a maximum of 1.00 (GRS). All genets are local, meaning that no individual genet occurs in more than one population.

Mean clonal identity was estimated for three (CHM, GRB, and BMR) of the five sites. The two remaining sites (GRS and DSG) are predominately sexual (PD = 1.0 and 0.97, respectively), making the calculation of clonal identity unnecessary. Mean clonal identity was plotted against distance at 5 meter intervals for each of the three sites analyzed (Figure 9). The greatest distance between two ramets of the same genet is 16.25 (BMT). Mean clonal identity is greatest within the 15 meter distance class for CHM with a value of 0.357. This indicates that the identification of clonal identity beyond the

**Table 2. Clonal diversity across sites.** *N* = sample size (number of ramets), *G* = number of genets, *PD* = *N/G*, and P bar is the probability that two ramets sharing the same multi-locus genotype belong to the same genet (Aspinwall and Christian).

| Site | *N* | *G* | *PD* | *P bar* | % Polymorphic |
|---|---|---|---|---|---|
| BMR | 30 | 22 | 0.73 | 0.99 | 83.33 |
| CHM | 19 | 8 | 0.42 | 0.98 | 45.83 |
| DSG | 35 | 34 | 0.97 | 0.99 | 91.67 |
| GRS | 16 | 16 | 1.00 | 0.99 | 91.67 |
| GRB | 23 | 20 | 0.87 | 0.99 | 87.50 |

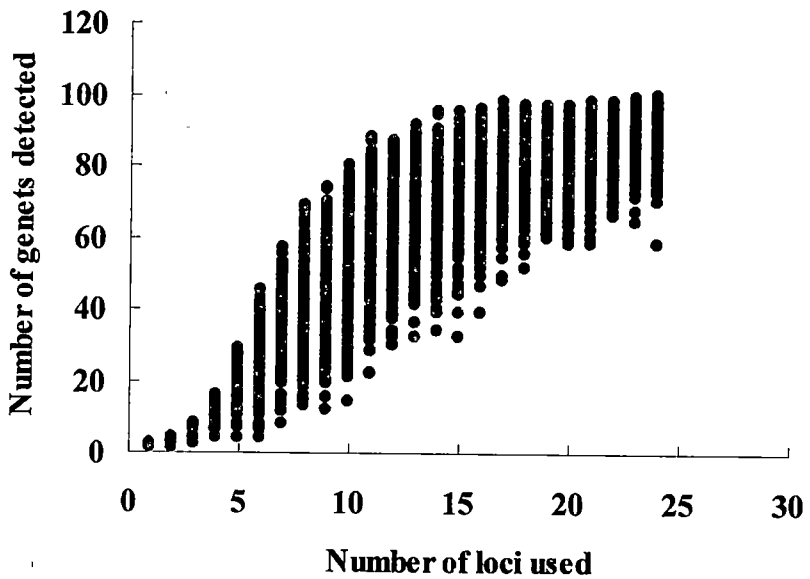


**Figure 8. Results of locus resampling.**

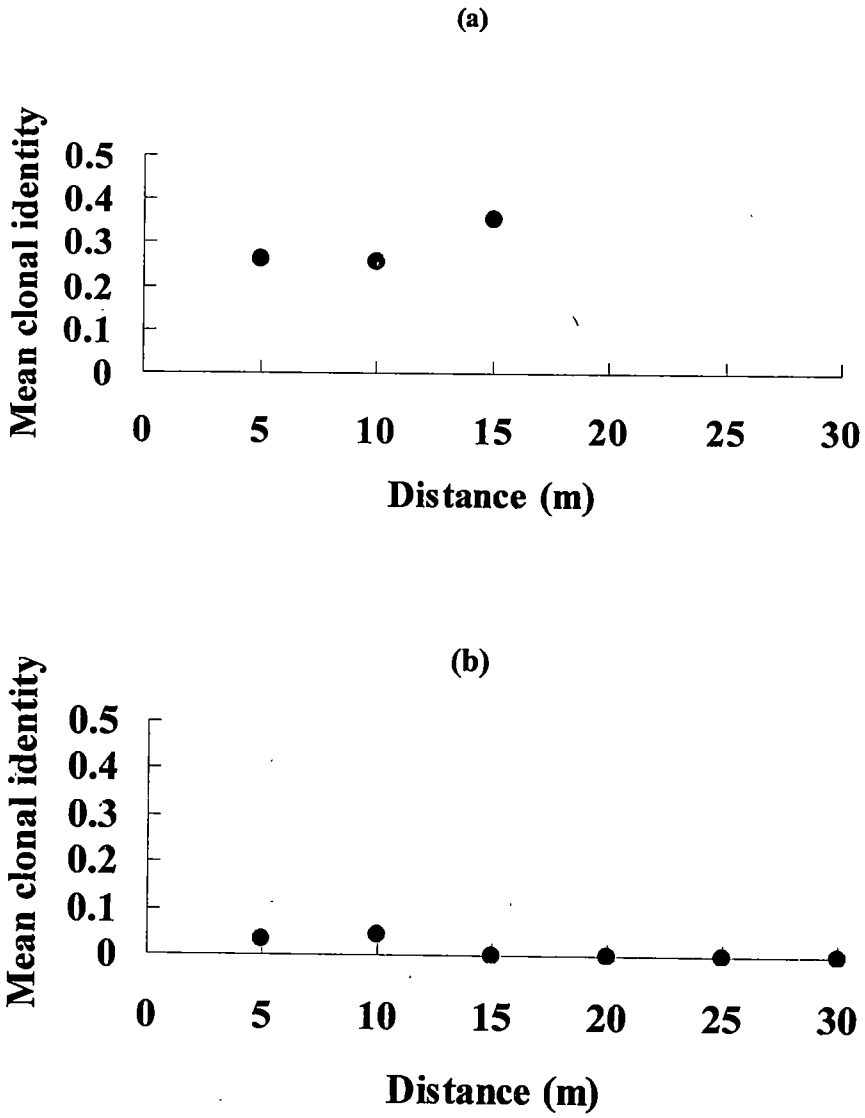


**Figure 9. Mean clonal identity.** Mean clonal identity is plotted against distance class for each of three sites (CHM = (a); GRB = (b); BMR = (c)).

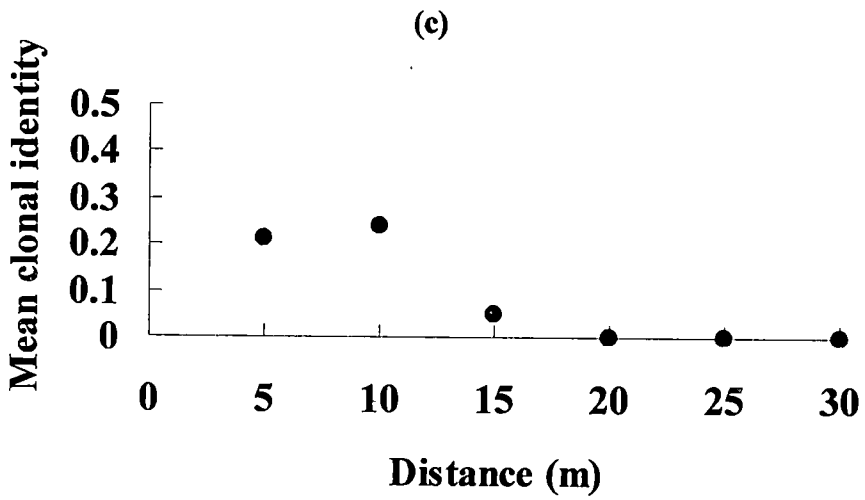


**Figure 9. Mean clonal identity (continued).** Mean clonal identity is plotted against distance class for BMR (c).

sampling area is likely. At both GRS and BMT, mean clonal identity is greatest within the 10 meter distance class, with values of 0.044 and 0.240, respectively. Beyond the 10 meter distance class, mean clonal identity is 0 at both sites.

### Genetic diversity

Genetic diversity is summarized in Table 3. Analysis of diversity among groups (HIGH and LOW) revealed a $G_{ST}$ value of 0.0243 and a $G_{CS}$ of 0.1979.Analysis of diversity within groups(HIGH and LOW) revealed a $G_{ST}$ of 0.1428 for HIGH sites and 0.2347 for LOW sites. Analysis of genetic distances indicate sites did not group by elevation (Table 4; Fig. 10).

## Discussion

### Data efficiency

Molecular markers provide ecologists and evolutionary biologists with a wealth of information regarding population structure and the processes shaping that structure. One consideration faced in all molecular studies is adequacy of the data collected for accurately assessing population structure. An insufficient number of markers can have the greatest impact on studies of clonal structure. The greater the number of alleles assayed, the greater the probability of differentiating among unique genets. Insufficient data results in the underestimate of the number of genets in a population, artificially inflating the size and range of the genets identified.

**Table 3. Genetic variation within and among sites**

| | $G_{ST}$ | $G_{CS}$ |
|---|---|---|
| ALL sites | 0.0243 | 0.1979 |
| LOW sites | 0.2347 | |
| HIGH sites | 0.1428 | |

**Table 4. Nei's genetic distance across sites.**

| Site | GRB | BMR | CHM | DSG | GRS |
|---|---|---|---|---|---|
| **GRB** | ***** | | | | |
| **BMR** | 0.2968 | ***** | | | |
| **CHM** | 0.2945 | 0.2010 | ***** | | |
| **DSG** | 0.0904 | 0.1977 | 0.1682 | ***** | |
| **GRS** | 0.2073 | 0.0688 | 0.1893 | 0.1229 | ***** |

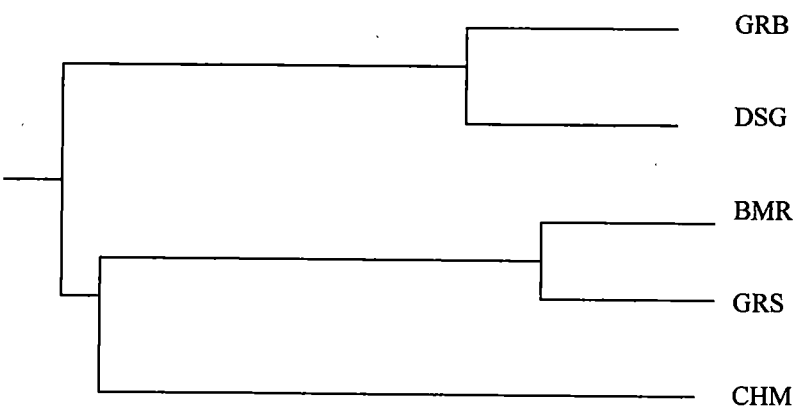


**Figure 10. Genetic distance dendrogram of all sites sampled.**

Two methods were used here to test the adequacy of the number of ISSR loci assayed to detect genetic variation and clonal structure in *Fagus grandifolia* in Great Smoky Mountains National Park. Both methods indicate that the 23 polymorphic loci used in analyses are sufficient to detect the amount of clonal variation present in those stems sampled for this research. The value of Pbar is 0.98 or greater within all sites (Table 2), which indicates that there is less than a 2% chance that two ramets with the same multi-locus genotype actually originated from different zygotes. The results from locus resampling also reflect the adequacy of the data collected here (Figure 8). Both of these approaches represent useful tools for population geneticists, particularly those studying clonal systems.

## Morphological variation among low and high elevations

Three specific hypotheses were presented as potential explanations for the observable differences among high and low elevation beech forests in Great Smoky Mountains National Park. The results of this research clearly reject two of these hypotheses and provide indirect support for the third.

### *Clonal reproduction*

The first hypothesis presented in this paper is that prolific clonal reproduction due to microenvironmental variation and herbivore disturbance is responsible for the visible differences in stand architecture among high and low elevations. Previous work by Held (1983) supports a shift in reproductive strategies from completely sexual reproduction to

increased frequencies of asexual reproduction in high elevations or northern latitudes where environmental conditions are more severe. His results were based on data from eight sites across six states (WI, OH, KY, IN, TN, and NC).

My work does not support this hypothesis, but is important to note that data on microsite variation as described in Held (1983) was not collected in this study. In GSMNP, high elevation stands persist under more extreme weather conditions than those found at low elevations. Shanks (1954) reports a decrease in mean temperature at an average rate of 2.23° F per 1000 feet increase in elevation, and that the climate of the high elevations in GSMNP is most similar to that of Maine or New Brunswick, approximately 1000 miles to the north. Because of this, high elevation sites are typically 10 to 15° F cooler than low elevations during the growing season. This temperature difference increases during winter. In addition, these high elevations reportedly receive more annual precipitation (as much as 203 cm) than any other location in the continental United States with the exception of the Pacific Northwest (Whittaker 1956). Alternatively, low elevations in GSMNP, mean annual precipitation ranges from 127 to 152 cm. Even with this extreme difference in environmental conditions, high elevation stands of *F. grandifolia* are no more clonal than are low elevation stands.

Clonal reproduction appears to have played the greatest role in structuring genetic variation at CHM, which is categorized here as a low elevation site (Table 1). At this site, the largest distance between two members of the same clone is 13.7 meters, which is considerably greater than distances reported between "parent" trees and sprouts by Jones and Raynal (1986b). In their study, trees with a diameter of more than 15 cm were

somewhat arbitrarily designated as parent trees and therefore the center of a 24m x 24m plot, within which all relationships among beech stems were determined by excavation. The greatest distance between a parent and sprout was 10.15 m, which is 3.55 m less than the greatest distance found in this study. However, Jones and Raynal do not specify the greatest distance among all ramets of clone. Therefore it is difficult to make a direct comparison between their results and mine. Houston and Houston (1994) assayed 9 polymorphic allozymes to estimate clonal structure in two beech stands, one in Massachusetts and one in West Virginia. They found stems with identical allozyme patterns separated by as much as 50 meters. However, the number of allozymes assayed may not have been sufficient to reflect the true level of genetic diversity present in those two populations. Because of this, clonal associations can only be suggested and are not strongly supported by the data. In my work, as previously discussed, two analytical methods were applied to the data to test this. So, I feel that the distance of 13.7 meters between clonemates as reported here is perhaps the largest distance reported in the literature for *Fagus grandifolia* that is well supported by the data presented. In addition, mean clonal identity at CHM is greatest in the 15 meter distance class, beyond which few data points were collected. This suggests that the CHM clone is potentially larger in both numbers of ramets and area covered.

The second largest clone detected in this study was at BMR, a high elevation site. This clone consisted of eight ramets, with the greatest distance between ramets being 10.3 meters. Because this site was sampled along a transect, the area covered by this clone (approximately 11 square meters) is likely underestimated. This raises the question

of what the appropriate sampling design must be to assess clonal structure at these sites. Recall that high elevation sites were sampled every 10 meters along a 90 meter linear transect, with leaf material sampled from the five closest stems at each of these 10 m intervals. This was done to cover as much distance as possible in such densely wooded areas proposed to be highly clonal in nature. Low elevation sites were sampled by complete census within rectangular plots. This was done because stem density was far less than at the high elevation sites, requiring a greater sampling area for a smaller sample size. The question, then, is how these two different sampling strategies may have effected the number of genets identified. The easiest way to test this without resampling in the field is to use the map data for CHM, which has the largest clone, and overlay transects for artificial resampling of clonal diversity. Because of the smaller sample size at this site relative to that of the high sites, it is not possible to sample the five closest stems every 10 meters as was done at the high sites. The total length of the overlaid transect can only reach a maximum distance of 22.7 meters depending on how it is drawn through the existing plot. At least at this scale, it is quite obvious that sampling 5 stems every 10 meters will miss the majority of the clonal variation present, suggesting that a plot-based approach may be more appropriate. In order to test the validity of this conclusion for high elevation sites it will be necessary to apply plot-based sampling in the field. Until these additional measures can be taken, it must be concluded from the data presented that there is no evidence to suggest that prolific root sprouting at high elevation sites is responsible for the phenotypic variation observed relative to low elevation sites.

***Independent evolutionary histories***

The second hypothesis initially proposed to explain the differences observed among high and low elevation stands of *Fagus grandifolia* in GSMNP was isolation by distance as a result of separate glacial refugia. Approximately 18000 years before present (BP), the climatic boundary separating boreal and warm-temperate vegetation reached far into the southeastern United States (1984). As a result, the Appalachian Mountains and the Atlantic Coast were predominately uninhabitable for cool-temperate hardwood species common to the region today. The question facing paleoecologists for decades is how displaced species survived during full glacial times. As reviewed in Delcourt and Delcourt (Delcourt and Delcourt 1984), there were two predominating theories in the early 1950's: 1) the diversity present in the modern Southern Appalachian landscape must represent relicts of ancient communities established as early as 24 million years ago, with endemic species persisting within intact hardwood communities throughout the Pleistocene; 2) cool-temperate hardwood species could not possible survive the dramatic cooling occurring in the Southern Appalachians during the Pleistocene, and must therefore have been displaced southward to Florida and Mexico. Fossil pollen evidence collected throughout the southeastern United States now provides support for the latter hypothesis, but no hard evidence exists to either reject or support the former.

The suggestion that the beech gaps of the high elevations of the Southern Appalachians are remnants of a second glacial refugium separate from that known to occur along the southeastern coastal plain and alluvial valleys of the Mississippi River is supported to some extent by the data collected in my research. $G_{ST}$ and $G_{cs}$ estimates

among and within groups (HIGH and LOW) suggest that there is more genetic variation among elevations than within elevations (Table 3). If these two regions originated from separate glacial refugia, genetic variation among populations should be greater than within populations, which is not the case here. However, when genetic distance was analyzed, sites did not group by elevation (Table 4, Figure 6). There is no clear pattern to the grouping of sites as illustrated by the distance dendrogram. Therefore, I conclude that the evidence presented here lends some support to separate evolutionary histories for high and low elevation sites. This should be further explored through detailed phylogeographic analyses across the range of the species.

***Phenotypic plasticity***

The third and final hypothesis presented to explain the variation observed among high and low elevation stands of beech in GSMNP is phenotypic plasticity. As previously stated (see the section on separate glacial refugia), there is a distinct difference in climate across the elevational gradient of the Smoky Mountains. As reviewed in DeWitt et al. (1998), phenotypic plasticity is the ability to produce a "better" phenotype-environment match across more environments than is possible by producing only a single phenotype across all environments. I suggest that the stunted growth of the beech gaps of the high elevations is a response to the sub-alpine climate that occurs there. With no significant difference in clonal variation and genetic variation among high and low sites, it is necessary to consider the possibility of plasticity.

Lacki (1985) provides some potential support for this idea with his study of variation in radial growth of beech stems in response to climate at high elevations in GSMNP. In his work, Lacki randomly sampled tree cores from three stands at each of two sites, one of which is the same as DSG in my research. After carefully documenting the patterns in annual growth rings for each core (cross-referenced to provide a 40 year series), he tested the significance of dependence of radial growth on climatological variation, include mean monthly temperature and total monthly precipitation for the 40 year series (1942 – 1982). Lacki then compared his results with those of Friesner (1941) and Fritts (1962), who conducted similar studies in Indiana and Ohio, respectively, where beech is reportedly morphologically distinct from beech in the high elevations of GSMNP. The findings of Friesner and Fritts suggest that growth of beech is limited in spring by the frequency of high temperature days, in summer by the amount of rainfall with greater rainfall resulting in larger ring widths. Lacki, in contrast, found no significant relationship between summer rainfall and radial growth of beech at high elevations. Lacki also found no relationship between radial growth and temperatures of the previous growing season as reported by Fritts. Lacki suggests that the differences in growth response of beech in the Ohio Valley and the high elevations of GSMNP supports the suggestion by Camp (1950) and Russell (1953) that beech found at the high elevations of the Great Smoky Mountains represent a distinct ecotype of *Fagus grandifolia.* However, based on the results obtained in my research, I propose that Lacki's work lends more support to the hypothesis of phenotypic plasticity than to ecotypic variation. It is important here to note that the hypothesis of phenotypic

plasticity as the structuring factor in these high elevation sites can only truly be tested by greenhouse experiments, and thus warrants further study. However, the data presented from my work suggest that phenotypic plasticity is the most plausible of the three hypotheses originally presented to describe the differences in phenotype observed among low an high elevation stands of *Fagus grandifolia* in Great Smoky Mountains National Park.

**Conclusions and future directions**

The primary objective of this research was to determine potential causes for the documented morphological variation among high and low elevation beech stands in Great Smoky Mountains national Park. Three hypotheses were proposed to explain these differences: 1) prolific clonal reproduction at high elevations; 2) genetic differentiation among high and low elevation sites as a result of origination from separate glacial refugia; and 3) phenotypic plasticity. The results obtained here negate the first hypothesis, lends inconclusive evidence to the second, but does not provide evidence either way for the third. Further investigation of this question will require a number of approaches. First, the hypothesis of clonality must be retested with different sampling strategies. It is apparent from map data at CHM that transect-based sampling designs may not adequately encompass the range of clonal variation, although this can only be tested by applying multiple sampling strategies at each site studied. In addition, the potential importance of microsite variation must be examined. Second, the hypothesis of separate glacial refugia must be studied on a larger scale across the range of the species.

The recently reviewed use of gene genealogies in such a study will provide direct genetic evidence for historical relationships among populations (Cruzan and Templeton 2000; Schaal and Olsen 2000). Third, the hypothesis of phenotypic plasticity must be further investigated by greenhouse experimentsusing seeds from across the range of morphological variation of the species. With each of these additional studies, the processes shaping both genetic and morphological variation in *Fagus grandifolia* may be identified.

In general, this work represents one of few studies that incorporate both within and among site variation in species known to reproduce both sexually and asexually. Such an approach can provide more information than either single site or among site studies alone.

# CHAPTER IV

## INVESTIGATING THE POTENTIAL FOR GENETIC RESISTANCE TO BEECH BARK DISEASE IN *FAGUS GRANDIFOLIA*

## Introduction

### The ecological importance of forest pathogens

Forest pathogens, although considered by managers to be detrimental to forest health, are essential for the maintenance of ecological balance in natural (Castello 1995). Pathogens may serve to selectively remove individual trees that are less fit, genetically or otherwise. There is a large body of literature on forest pathogens, but rarely does the literature focus on the interaction between pathogens and abiotic disturbance, and the role that this interaction plays in forest succession. Castello et al. (1995) review a number of cases encompassing both within population (gaps) and among population (landscapes) level studies. They suggest that the impact of pathogens on species distribution, forest structure and composition, and succession is strongly dependent on land use history and the resulting landscape patterns. One well-known example of this is the explosion of bark beetle populations as a result of long-term fire suppression in regions dominated by lodge-pole pine. Fire suppression results in dramatically increased stand basal area and increased stand age, both of which can lead to greater susceptibility to bark beetles as well as a number of other pathogens, such as dwarf mistletoe, western gall rust, and comandra rust (Castello 1995). Under natural fire regimes pathogen outbreaks are suppressed by the natural thinning of otherwise healthy forests.

In altered forests where pathogen outbreaks are extensive, human management may be necessary to sustain the infected species. An integrated approach to such intervention should include considerations of genetic structure and potential genetic resistance to the pathogen in question. In this work I use inter-simple sequence repeat (ISSR) markers to determine if populations of *Fagus grandifolia* in Great Smoky Mountains National Park contain individuals that are genetically resistant to beech bark disease.

**Beech bark disease in *Fagus grandifolia***

In recent years populations of *Fagus grandifolia* (American Beech) have been decimated across the species' range due to an insect mediated fungal pathogen complex known as Beech Bark Disease (BBD). The insect component of this complex was accidentally introduced into Nova Scotia in 1890 by way of infected nursery stock, and has since spread westward and southward to cover much of the current range of American Beech (Ehrlich 1934, Houston 1979). The woolly beech scale (*Cryptococcus fagisuga* Lindinger) infects beech bark, which then provides a pathway for infection by *Nectria coccinea var. faginata* or *N. galligena*, two species of pathogenic fungi (Ehrlich 1934, Houston 1983, Gavin and Peart 1993). The chronology of the disease has been well documented Jones 1986, Ehrlich 1934, Houston 1979). As this disease spreads throughout the range of American Beech, average mortality of beech within a population is estimated at 85% (Houston 1979). Such high levels of mortality have been partially attributed to past logging practices (Houston 1979). Historically, beech was not

considered a valuable timber species in northern hardwoods and was often left behind in logged stands. This has resulted in the rise of beech as a major component in many northern forests. This overabundance of beech stems is thought to be a contributing factor in the widespread dispersal of beech bark disease. Damage incurred by logging practices is thought to have produced a niche for *C. fagisuga* on otherwise healthy stems (Houston 1982). Houston (1983) reports that although there are known predators of the beech scale, their effectiveness has been limited. Currently there is no known effective treatment for BBD.

**Potential for genetic resistance**

Resistance to BBD in North America was first studied by Houston (1982), who introduced *C. fagisuga* onto apparently resistant beech trees in aftermath forests to test susceptibility over a period of three years. On susceptible trees the insect was able to complete its life cycle and produce massive amounts of eggs. Alternatively, the insect failed to become established on stems that Houston perceived to be resistant. Based on these trials, Houston produced two hypotheses regarding the source of resistance in these trees: 1) the inability of overwintered insects to complete their life cycle on disease free trees may be a result of the presence of some toxin or lack of some necessary chemical in the tree itself; 2) the complete lack of infestation on some trees and consistently low levels of infestation on others may be the result of either a complete or partial anatomical barrier in these stems. In more recent work, Houston and Houston (1994 and 2000) study

a third hypothesis: the inability of the scale insect to establish itself on some trees may be the result of genetic resistance.

The potential for resistance has been reported, although observed resistant trees are rarely recorded (Houston 1983). Houston and Houston (1994) observed resistance in less than 1% of all beech trees that they sampled. They also note that resistant stems often occur in discrete clumps or groups, suggesting the possibility of resistance clonal genotypes. At least for *F. sylvatica*, resistance in planted orchards has been associated with genotype (Wainhouse and Deeble 1980). For *F. grandifolia*, no known relationship between genotype and resistance has yet been found. Houston and Houston (1994, 2000) used isozymes to identify resistant genotypes and found that in some instances individual stems sharing the same genotype could be either infected or not infected. However, they do report finding higher levels of observed heterozygosity in susceptible stands, suggesting some level of genetic differentiation between susceptible and resistance trees (Houston and Houston 2000). It is important to note that allozymes often do not display adequate levels of variation to resolve genetic identities in clonal populations (Cruzan 1998, Escaravage 1998, Waycott 1998). Whereas allozyme studies are commonly limited to less than 10 polymorphic loci, studies using markers based on DNA variation can include many times that number of loci for the resolution of individual genotypes. In this study, I use inter-simple sequence repeats (ISSRs) to investigate the potential of genetic resistance to beech bark disease in *Fagus grandifolia* within Great Smoky Mountains National Park.

**Beech bark disease in Great Smoky Mountains National Park**

Beech bark disease (BBD) was first noted in Great Smoky Mountains National Park (GSMNP) in 1993, and since initial discovery, areas of nearly complete stand mortality have been located (Blozan 1995). West of Clingman's Dome, high elevation beech forests are continuous along the North Carolina/ Tennessee state boundary for several miles (Blozan 1995). This area is likely the most susceptible to BBD due to high densities of individuals within populations and relative proximity of populations to one another, and is feared to be the site of origin of BBD in the park (Blozan 1995). Evidence suggests that BBD is most prevalent in moist, shaded areas, and that larger trees in higher densities appear to be more susceptible (Gavin and Peart 1993). In addition, high incidence of BBD has been reported for lower elevation areas where hemlock (*Tsuga canadensis*) is common (likely due to high shade and moisture regimes) suggesting that BBD is not limited to higher elevation (Blozan 1995). However, the high density of stems and consistently moist conditions at higher elevation sites in GSMNP create ideal conditions for BBD infestation and spread (Blozan 1995). In this study, I sample genetic variation within and among three sites to estimate the relationship between level of scale infestation and genotype in GSMNP.

# Methods

## Study system

### *Fagus grandifolia*

American beech (*Fagus grandifolia* Ehrh. (Fagaceae)) occurs over a broad geographic range in the eastern United States and Mexico, where it is found in a diversity of ecological habitats (Figure 3). Beech is a shade tolerant, late successional, mesophytic species associated with at least 20 forest cover types in North America. Beech is a monoecious species, with male flowers occurring in catkins and female flowers occurring in clusters. An individual tree can produce a substantial seed mast at 40 years of age, and good seed crops occur at intervals of 2 to 8 years. The primary dispersal agent for beechnuts is the bluejay, which can transport seeds across several kilometers. Asexual reproduction by root sprouting appears to be abundant under stressful or extreme environmental conditions and is likely a product of injury to the root or tree (Held 1983, Jones 1986) Causes of such injury include foraging mammals, freeze-thaw action, and human activity. Variation in the frequency of root injury may explain large differences in the number of sprouts found in stands with similar composition, as well as regional variation in relative abundance of sprouts reported by various researchers.

### Beech bark disease

The beech scale is parthenogenic and lays its eggs in mid to late summer before dying. The nymphs hatch in late summer or early fall and disperse locally on the same tree, with some nymphs being dispersed by wind and birds to other trees. The scale

nymphs insert their stylets into the tree, molt before winter, and molt again the following spring. Adult scales secrete a white waxy substance that is easily detectable by observers. The highest densities of scale occur in cracks and crevices of beech bark, which makes older trees with rougher bark more favorable for infection. Initial infestation by *C. fagisuga* provides the niche for infection *by N. coccinea* var. *faginata*, an Ascomycete fungus. Signs of invasion often include tarry spots (which consist of a black or brown exudate on the bark of infected trees) and red perithecia, which is the sexual stage of the fungus. Mortality to the tree is a result of the vascular cambium being girdled, preventing the flow of nutrients throughout the tree.

BBD is described as having three chronological stages (Shigo 1972). The advancing front is the initial stage and is characterized by increasing populations of *C. fagisuga*. The second stage is referred to as the killing front, and is characterized by large populations of *C. fagisuga*, severe *Nectria* attacks, and high rates of tree mortality. Finally, the aftermath zone consists of defective trees, hypothesized to be resistant, often defective sprout thickets, and endemic populations of the disease complex. The Great Smoky Mountains is currently encompassed by the killing front.

**Study Site**

Great Smoky Mountains National Park (GSMNP) is an International Biosphere Reserve that encompasses over 521,000 acres shared between Tennessee and North Carolina, and is the most visited National Park with over 9,000,000 visits annually. In GSMNP, elevation ranges from 800 to 6643 feet, resulting in a diversity of forest cover

types, including spruce-fir forest (above 4500 feet), northern hardwood forest (3500 to 5000 feet), pine-oak forest (dry ridges), hemlock forest (along stream banks), and cove hardwood forest (low valleys). Beech is a minor component in cove hardwoods, hemlock, and pine-oak forest and a major component in northern hardwood and spruce-fir forests in the park.

**Sampling Methods**

Data were collected from three sites, which were selected based on long-term monitoring projects managed by GSMNP staff. Two of these sites, Balsam Mountain Road (BMR), and Double Spring Gap (DSG), are above 4500 feet in elevation and are representatives of "beech gap" forests, which are located within topographic gaps in areas otherwise dominated by spruce-fir forests. The third site, the Chimneys (CHM), occurs below 3500 feet and is associated with eastern hemlock (*Tsuga canadensis*) forests. Because stem densities vary greatly between high elevation (BMR and DSG) and low elevation (CHM) sites, different sampling strategies were applied to these two areas (described in the previous chapter).

Disease infection was assessed by estimating the percent cover of scale on the trunk of each tree (0 = none, 1 = < 25%, 2 = 26 – 75%, 3 = >75%). Leaf material was collected from each individual, placed in separate 1.5-mL microcentrifuge tubes and stored on ice while in the field. Upon returning to the lab, all samples were snap frozen in liquid nitrogen and stored at -70° C.

**DNA Extraction**

Total genomic DNA was extracted from each sample using a protocol adapted from Edwards et al (1991) and modified by Martin and Cruzan (1999). Leaf material was ground in 100 ul of extraction buffer (200 mM Tris HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS) for 15 seconds using disposable grinders. An additional 400 ul of extraction buffer was added to this mixture and ground again. Each sample was then centrifuged at 13000 rpm for 2 minutes. Following centrifugation, 400 ul of supernatant was extracted and placed in a sterile 1.5 ml microcentrifuge tube, to which 400 ul of isopropanol was added. After standing at room temperature for 2 minutes, samples were centrifuged at 13000 rpm for 10 minutes. All liquid was then poured off, and the samples were allowed to dry. The remaining DNA was then resuspended in 100 ul of $dH_20$. Each sample was cleaned by binding the DNA to DEAE-cellulose as described by Marechal-Drouard and Guillemaut (1995) and then stored at –20° C.

**ISSR Protocol**

One ISSR primer was selected from UBC primer set number 9 based on the brightness and consistency of the bands produced during screening. Single-primer reaction conditions were carried out in 15 ul reactions as follows: 2.5 mM MgCl2, 200 dNTPs, 1 unit of Taq polymerase, 0.10 μM primer, and 0.5 μl DNA. The thermal cycler profile was adopted from Huang and Sun (2000) and is as follows: 1 cycle at 94° C for 5 minutes, followed by 45 cycles at 94° C for 5 seconds, 50° C for 45 seconds, 72° C for 1.5 minutes; and a final 7 minute extension at 72° C. PCR products were electrophoresed

on 2% agarose gels in 1X TBE buffer until bromphenol blue marker migrated 10 cm from the origin. Gels were stained with ethidium bromide and were documented digitally using Kodak 1D Biomax software. Bands with the same molecular weight were treated as identical loci. A data matrix was compiled in which band presence was scored as 1, band absence was scored as 0. Ambiguous bands were eliminated from the analysis.

**Data Analysis**

***Data efficiency***

Two methods were used to determine if more data were necessary to detect all possible variation. The first method, introduced by Aspinwall and Christian (1992) calculates the average probability of ramets with the same multi-locus genotype belonging to the same genet:

$$\overline{P} = 1 - \frac{\left[\sum_{Q=1}^{N}\prod_{D=1}^{M}\left[\frac{X_{DQ}}{P_D}\right]\right]}{N}$$

where Q is the particular individual in a sample, $N$ is the total number of individuals in a sample, $M$ is the number of polymorphic loci in the sample, $D$ is an individual polymorphic locus, $X_{DQ}$ is the number of individuals in the sample of the same $D$th multi-locus genotype as the $Q$th individual, and $P_D$ is the total number of individuals examined for each primer. This probability and the number of polymorphic loci in a population are

highly correlated, such that as the probability approaches 1, additional assays will result in diminishing returns.

In the second method varying numbers of loci (1 through 24) were randomly resampled to infer the effects of increasing the number of markers on the number of genets detected. The number of genets was calculated for one thousand replicates of each data set size and these data were used to examine the relationship between number of loci in the data set and the number of genets detected. I assumed that an asymptotic curve (i.e., the number of new genets detected approaches zero as the number of loci in the data set is reached) would indicate that adding additional loci would not result in an appreciable change in the total number of genets detected.

**Genetic resistance**

Genetic structure was analyzed by calculating the number of ramets, number of genets, and number of polymorphic loci assayed. Ellstrand and Roose's proportion distinguishable (PD) was used to estimate clonal diversity within sites. Due to the small number of clones detected, statistical analysis of the relationship between genotype and level of infection was not possible. Thus, this relationship is qualitatively assessed.

**Gene diversity**

Gene diversity within and among populations was estimated using the program POPGENE (Yeh et al. 1999). This program calculates Nei's (1987) gene diversity statistics in subdivided populations ($G_{ST}$).

# Results

## Data efficiency

A total of 84 individuals across three sites were assayed for genetic variation . One ISSR primer resulted in 24 loci, of which were polymorphic. Sixty four unique genets were identified across sites. Two methods were used to determine if the number of polymorphic loci detected is sufficient to detect the actual number of genets within and across sites. Aspinwall and Christian's Pbar is the average probability that ramets with the same multi-locus genotype actually belong to the same genet. This value is 0.99 across sites and never less than 0.98 within sites. Locus resampling resulted in an asymptotic curve. Both methods indicate that sufficient genetic data was obtained.

## Genetic structure and resistance

The number of ramets per site was 19 (CHM), 35 (DSG), and 30 (BMR). The number of genets per site was 8 (CHM), 34 (DSG), and 22 (BMR). The proportion distinguishable (PD) was calculated as 0.42 (CHM), 0.97 (DSG), and 0.73 (BMR). All genets are local, meaning that no individual genet occurs in more than one population. There is no apparent relationship between genotype and the level of beech scale infection. Ramets of the same genet exhibit ranges of infection from 0 – 3, encompassing all possible rating categories. Analysis of diversity among all three sites revealed a $G_{ST}$ value of 0.2209.

# Discussion

The objective of this research is to determine if there exists some evidence for genetic resistance to beech bark disease in stems of *Fagus grandifolia* at high elevations in Great Smoky Mountains National Park. In my work using ISSR markers, clonemates were observed to be both infected and uninfected with *Cryptococcus fagisuga*, suggesting that genetic resistance is not a factor in the sites studied. This is consistent with previous work by Houston and Houston (1994, 2000) in the central and northeastern United States. In light of this, and in light of previous evidence in support of some form of resistance (Houston 1985, Wainhouse and Deeble 1980), it is important to discuss the validity of the methods used both here and in the Houston and Houston studies.

Genetic resistance is often associated with one or a few resistance genes in the plant and avivirulance (*avr*) genes in the pathogen (Caicedo et al. 1999). Resistance is thus an interaction among genes of plant and pathogen, and if either member of this gene-pair interaction is absent, disease may result. Because allozymes represent protein loci and ISSRs represent arbitrarily primed DNA fragments, specific resistance genes would be difficult to detect with such assays. Therefore, a truly detailed analysis of resistance requires genetic mapping, which is both more expensive and more time intensive than either of the two markers discussed here.

Other considerations for methodology concern the estimation of infection in the field. More detailed information regarding the range of infection observed beyond "susceptible" and "resistant" categorization. In my research, stems were qualitatively

categorized from infection to heavy infection. Such an approach is a good starting point, but it will be necessary in future research to obtain more quantitative estimates. Wiggins describes a rating scale of 0 to 6 in which a rating square measuring 33 cm x 33 cm placed on the bark of the tree 122 cm above the ground. Both scale and perithicia coverage were measured and distribution described within the rating square, and this was done on both the north and south sides of the tree. Such a technique, although highly informative, may be too time consuming when additional research (e.g. genetic assays) is required. Thus some intermediate approach between that described in Wiggins and that used in my research may be the most efficient way to assess BBD infection in the field.

An important consequence my research, as well as that of Houston and Houston (1994, 2000) is that in the absence of genetic resistance, there is no known way to control the explosion of beech bark disease in GSMNP, or the rest of the species range. This may have far reaching implications for the structure and composition of infected stands across GSMNP. As reviewed in Castello et al. (1995), similar pathogenic outbreaks in other species have dramatically altered the forested landscape. The American Chestnut blight has eliminated the American chestnut from forest canopies in the eastern United States, where it was previously a dominant component of many forest communities. These communities are now described as more diverse in species composition. However, no single species has replaced the role of canopy dominant previously occupied by American chestnut.

Although the increase in diversity across these aftermath stands may be viewed by some as an improvement, it is imperative to realize that not only the tree species within a

site are affected by forest pathogens. Widespread mortality of canopy species results in increasing light gaps, which in turn may change the composition of the vegetation on the forest floor. In addition, vertebrate as well as invertebrate species that are dependent on either the fruits of the eliminated dominant species or the shelter that it provides may no longer persist in these aftermath stands. Granted, not all species will be negatively effected. However, the preservation of unique communities such as the American chestnut forests of the past or the beech gaps endemic to the high elevations of the Great Smoky Mountains sustains diversity at the level of the community as opposed to at the level of a single site. If the beech gaps of GSMNP are to be sustained as an ecologically unique unit, additional research into the potential for genetic resistance is required.

# CHAPTER V

## CONCLUSIONS

In clonal plant populations, the emphasis of research goals appears to be shifting from questions of resource translocation and ramet interdependence to questions of population genetic structure and evolution. With this shift in goals has necessarily come the need for new analytical approaches. Work in clonal populations over the past decade has resulted in a considerable number of statistics specific to clonal organisms. The review of these statistics presented in Chapter 2 of this thesis indicates that each formula is potentially useful to varying degrees under numerous circumstances. It is thus necessary for future research to apply multiple statistics to the data to determine which are the most appropriate, both in the level of conservatism of the result and in the utility in answering the questions posed.

The goal of this research was to assess genetic structure both within and among populations of a clonal tree species, *Fagus grandifolia*, across elevations in Great Smoky Mountains National Park. Specifically, four questions were posed to approach this goal: 1) Do high elevation populations of *F. grandifolia* in GSMNP reproduce more frequently by root sprouting than by seed, resulting in fewer genets than ramets within a population? 2) Do low elevation populations of *F. grandifolia* in GSMNP reproduce more frequently by seed than by root sprouting, resulting in a relatively equal number of ramets and genets within a population? 3) Are high elevation populations of *F. grandifolia* in GSMNP genetically distinct from low elevation populations? and 4) Is an absence of

beech scale (*Cryptococcus fagisuga*) infection on an individual tree in an otherwise infected stand correlated with genotype as a result of genetic resistance?

The results of the data collected here indicate that clonal reproduction is limited at most sites, both at high and low elevations, suggesting that sexual reproduction dominates in *F. grandifolia* within Great Smoky Mountains National Park. However, the potential importance of microsite variation in controlling the frequency of clonal reproduction should not be overlooked. Specific data relevant to soil types, moisture regimes, and aspect were not collected in this study. If these factors are determinants of clonality, then the transect sampling applied at high elevations could potentially miss patches of clones. Thus important considerations for future clonal studies in this species as well as others include appropriate sampling scale and the importance of environmental heterogeneity.

The data collected here also indicates that there is more genetic differentiation among high and low elevation sites than within, suggesting the potential for separate evolutionary histories of high and low elevation sites. There is no evidence either way to support or negate phenotypic plasticity as a factor explaining observed morphological variation among elevations. Direct evidence for plasticity must be acquired from greenhouse experiments and potentially from the analysis of quantitative trait loci.

Finally, the data collected in this research indicate that there is no genetic resistance to beech bark disease exhibited by *F. grandifolia* stems in the high elevation sites at GSMNP. It is important to note here that disease resistance is commonly the result of a single or few resistance genes in the plant interacting with specific genes in the

pathogen. Such a relationship may go undetected when using ISSR markers (but not necessarily so), which target arbitrary DNA fragments rather than specific genes. Future work on this question will require using markers with a larger clone sample size. If resistance is established, genetic mapping can be used to determine the specific genes involved.

In general, the study of clonal organisms requires a variety of approaches, both within and among populations, in order to understand the importance of clonality in structuring the species. Future studies will surely include some temporal aspect in an attempt to assess the long-term stability of clonal structure, from which evolutionary patterns may be inferred.

# BIBLIOGRAPHY

## BIBLIOGRAPHY


Aspinwall, N., and T. Christian. 1992. Clonal structure, genotypic diversity, and seed production in populations of *Filipendula rubra* (Rosaceae) from the northcentral United States. American Journal of Botany 79:294-299.

Bell, G. 1982. The masterpiece of nature: the evolution and genetics of sexuality. University of California Press.

Berg, E. E., and J. L. Hamrick. 1994. Spatial and genetic structure of two sandhills oaks: *Quercus laevis* and *Quercus margaretta* (Fagaceae). American Journal of Botany 81:7-14.

Blozan, W. 1995. Dendroecology of American beech stands infested with beech bark disease. Resources Management and Science Division, Great Smoky Mountains National Park, Gatlinburg.

Braun, E. L. 1950. Deciduous forests of eastern North America. Blakiston, Philadelphia.
Burke, J. M., M. R. Bulger, R. A. Wesselingh, and M. L. Arnold. 2000. Frequency and spatial patterning of clonal reproduction in Louisiana iris hybrid populations. Evolution 54:137-144.

Caicedo, A. L., B. A. Schaal, and B. N. Kunkel. 1999. Diversity and molecular evolution of the RPS2 resistance gene in *Arabidopsis thaliana*. Proceedings of the National Academy of Sciences, USA 96:302-306.

Camp, W. H. 1950. A biogeographic and paragenetic analysis of the American beech. Year book of the American Philosophical Society:166-169.

Cantlon, J. E. 1953. Vegetation and microclimates on north and south slopes of Cushetunk Mountain, New Jersey. Ecological Monographs 23:241-270.

Carbone, I., J. Anderson, and L. Kohn. 1999. Patterns of descent in clonal lineages and their multilocus fingerprints are resolved with combined gene genealogies. Evolution 53:11-21.

Castello, J. D., D. J. Leopold, and P. J. Smallidge. 1995. Pathogens, patterns, and processes in forest ecosystems. BioScience 45:16-24.

Coffroth, M. A., and H. R. Lasker. 1998. Population structure of a clonal gorgonian coral: the interplay between clonal reproduction and disturbance. Evolution 52:379-393.

Cooper, A. W., and E. P. Mercer. 1977. Morphological variation in *Fagus grandifolia* Ehrh. in North Carolina. The Journal of the Mitchell Society Fall:136-149.

Crow, J. 1988. The importance of recombination. *in* R. E. Michod and B. R. Levins, eds. The evolution of sex. Sinauer Associates, Inc.

Cruzan, M., and A. Templeton. 2000. Paleoecology and coalescence: phylogeograhic analysis of hypotheses from the fossil record. Trends in Ecology and Evolution 15:491-496.

Cruzan, M. B. 1998. Genetic markers in plant evolutionary ecology. Ecology 79:400-412.

Delcourt, H. R., and P. A. Delcourt. 1984. Ice age haven for hardwoods. Natural History 93:22 - 25.

Delcourt, H. R., and P. A. Delcourt. 1988. Quaternary landscape ecology: Relevant scales in space and time. Landscape Ecology 2:23-44.

DeWitt, T. J., A. Sih, and D. S. Wilson. 1998. Costs and limits of phenotypic plasticity. Trends in Ecology and Evolution 13:77-81.

Edwards, K., C. Johnstone, and C. Thompson. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Research 19:1349.

Ehrlich, J. 1934. The beech bark disease. A *Nectria* disease of *Fagus*, following Cryptococcus fagi (Baer.). Canadian Journal of Research 10:593-692.

Ellstrand, N. C., and M. L. Roose. 1987. Patterns of genotypic diversity in clonal plant species. American Journal of Botany 74:123-131.

Escaravage, N., S. Questiau, A. Pornon, B. Doche, and P. Taberlet. 1998. Clonal diversity in a *Rhododendron ferrugineum* L. (Ericaceae) population inferred from AFLP markers. Molecular Ecology 7:975-982.

Fay, M. F., M. D. Lledo, M. M. Kornblum, and M. B. Crespo. 1999. From the waters of Babylon? *Populus euphratica* in Spain is clonal and probably introduced. Biodiversity and Conservation 8:769-778.

Friesner, R. C. 1941. A preliminary study of growth in the beech, *Fagus grandifolia*, by the dendrographic method. Butler University Botanical Studies 5:85-94.

Fritts, H. C. 1962. The relation of growth ring widths in American beech and white oak to climate. Castanea 25:2-10.

Gavin, D. G., and D. R. Peart. 1993. Effects of beech bark disease on the growth of American beech (*Fagus grandifolia*). Canadian Journal of Forest Research 23:1566-1575.

Gleason, H. A., and A. Cronquist. 1963. Manual of vascular plants of northeastern United States and adjacent Canada. D. van Nostrand Co., Princeton, New Jersey.

Groenendael, J. M. v., L. klimes, J. Klimesova, and R. J. J. Hendriks. 1997. Comparative ecology of clonal plants. Pp. 191-209 *in* J. Silvertown, M. Franco and J. Harper, eds. Plant life histories: Ecology, phylogeny, and evolution. Cambridge University Press.

Gupta, M., Y.-S. Chyi, J. Romero-Severson, and J. L. Owen. 1994. Amplification of DNA markers from evolutionarily diverse genomes using single primers of simple-sequence repeats. Theoretical Applied Genetics 89:998-1006.

Harada, Y., and Y. Iwasa. 1996. Analyses of spatial patterns and population processes of clonal plants. Researches on Population Ecology 38:153-164.

Harada, Y., S. Kawano, and Y. Iwasa. 1997. Probability of clonal identity: inferring the relative success of sexual versus clonal reproduction from spatial genetic patterns. Journal of Ecology 85:591-600.

Held, M. E. 1983. Pattern of beech regeneration in the east-central United States. Bulletin of the Torrey Botanical Club 110:55-62.

Hillis, D. M., C. Moritz, and B. K. Mable. 1996. Molecular Systematics. Sinauer Associates, Inc., Sunderland, Massachusetts.

Holderegger, R., I. Stehlik, and J. J. Schneller. 1998. Estimation of the relative importance of sexual and vegetative reproduction in the clonal woodland herb *Anemone nemorosa*. Oecologia 117:105-107.

Hossaert-McKey, M., M. Valero, D. Magda, M. Jarry, J. Cuguen, and P. Vernet. 1996. The evolving genetic history of a population of *Lathyrus sylvestris*: evidence from temporal and spatial genetic structure. Evolution 50:1808-1821.

Houston, D. B., and D. R. Houston. 2000. Allozyme genetic diversity *among Fagus grandifolia* trees resistant or susceptible to beech bark disease in natural populations. Canadian Journal of Forest Research 30:778-789.

Houston, D. B. a. D. R. Houston. 1994. Variation in American beech (*Fagus grandifolia* Ehrh.): isozyme analysis of genetic structure in selected stands. Silvae Genetica 43:277-284.

Houston, D. R. 1982. American beech resistance to *Cryptococcus fagisuga*. IUFRO Beech bark disease working party conference proceedings, Hamden, CT 38-42.

Houston, D. R. 1983. Developments in biological control of beech bark disease. 10 International Congress of Plant Protection, Brighton, England 1035-1041.

Houston, D. R., E. J. Parker, R. Perrin, and K. J. Lang. 1979. Beech bark disease: a comparison of the disease in North America, Great Britain, France, and Germany. International Journal of Forest Pathology 9:199-211.

Huang, J. C., and M. Sun. 2000. Genetic diversity and relationships of sweetpotato and its wild relatives in *Ipomoea* series *Batatas* (Convolvulaceae) as revealed by inter-simple sequence repeat (ISSR) and restriction analysis of chloroplast DNA. Theoretical and Applied Genetics 100:1050-1060.

Jones, R. H. 1986a. Initiation, spatial distribution, and demography of root sprouts in American beech (*Fagus grandifolia* Ehrh.). Environmental Science and Forestry. New York, Syracuse.

Jones, R. H. a. D. J. R. 1986b. Spatial distribution and development of root sprouts in *Fagus grandifolia* (Fagaceae). American Journal of Botany 73:1723-1731.

Klimes, L., Jitka Klimesova, Rob Hendriks, and Jan van Groenendael. 1997. Clonal plant architecture: a comparative analysis of form and function. Pp. 1-29 *in* H. d. K. a. J. v. Groenendael, ed. The Ecology and Evolution of Clonal Plants. Backhuys, Leiden.

Lacki, M. J. 1985. Variation in radial growth of American beech (*Fagus grandifolia*) at high elevations in the Great Smoky Mountains. Bulletin of the Torrey Botanical Club 112:398-402.

Lacki, M. J., and R. A. Lancia. 1986. Effects of wild pigs on beech growth in Great Smoky Mountains National Park. Journal of Wildlife Management 50:655-659.

Marechal-Drouard, L., and P. Guillemaut. 1995. A powerful but simple technique to prepare polysaccharide free DNA, quickly and without phenol extraction. Plant Molecular Biology 13:26-30.

Martin, L. J., and M. B. Cruzan. 1999. Patterns of hybridization in the *Piriqueta caroliniana* complex in central Florida: evidence for an expanding hybrid zone. Evolution 53:1037-1049.

Maynard Smith, J. 1989. Evolutionary Genetics. Oxford University Press.

McGregor, C. E., C. A. Lambert, M. M. Greyling, J. H. Louw, and L. Warnich. 2000. A comparative assessment of DNA fingerprinting techniques (RAPD, ISSR, AFLP and SSR) in tetraploid potato (*Solanum tuberosum* L.) germplasm. Euphytica 113:135-144.

Mclellan, A. J., D. Prati, O. Kaltz, and B. Schmid. 1997. Structure and analysis of phenotypic and genetic variation in clonal plants. Pp. 185-210 *in* H. d. Kroon and J. v. Groenendael, eds. The Ecology and Evolution of Clonal Plants. Backhuys, The Netherlands.

Melancon, S. a. M. J. L. 1986. Differences in the damage caused by glaze ice on codominant *Acer saccharum* and *Fagus grandifolia*. Canadian Journal of Botany 65:1157-1159.

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

Parks, J. C., and C. R. Werth. 1993. A study of spatial features of clones in a population of bracken fern, *Pteridium aquilinum* (Dennstaedtiaceae). American Journal of Botany 80:537-544.

Pejic, I., P. Ajmone-Marsan, M. Morgante, V. Kozumplick, P. Castiglioni, G. Taramino, and M. Motto. 1998. Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theoretical and Applied Genetics 97:1248-1255.

Radford, A. E., H. E. Ahles, and C. R. Bell. 1968. Manual of the vascular flora of the Carolinas. University of North Carolina Press, Chapel Hill.

Reusch, B. H., W. T. Stam, and J. L. Olsen. 2000. A microsatellite-based estimation of clonal diversity and population subdivision in *Zostera marina*, a marine flowering plant. Molecular Ecology 9:127-140.

Richards, A. J. 1997. Plant Breeding Systems. Chapman and Hall.

Rushmore, F. J. 1961. Silvical characteristics of beech (*Fagus grandifolia*). United States Forest Service Northeast Forest Experiment Station Paper 161.

Russell, N. H. 1953. The beech gaps of the Great Smoky Mountains. Ecology 34:366-374.

Sain, R. E. a. K. E. Blum. 1981. Seedling production in the high-elevation beech (*Fagus grandifolia*) forests of Great Smoky Mountains National Park. Castanea 46:217-224.

Schaal, B., and K. Olsen. 2000. Gene genealogies and population variation in plants. Proceedings of the National Academy of Science 97:7024-7029.

Schlapfer, F., and M. Fischer. 1998. An isozyme study of clone diversity and relative importance of sexual and vegetative recruitment in the grass *Brachypodium pinnatum*. Ecography 21:351-360.

Shanks, R. E. 1954. Climates of the Great Smoky Mountains. Ecology 35:354-361.

Sharp, A. J. 1972. Epilogue *in* P. C. Holt, ed. The distributional history of the biota of the Southern Appalachians. Part II. Flora.

Shen, C.-F. 1992. A monograph of the genus *Fagus* Tourn. ex L. (Fagaceae). Ph.D. Dissertation. The City University of New York. 390 pp.

Shigo, A. L. 1972. The beech bark disease today in the northeastern U.S. Journal of Forestry 70:286-289.

Silvertown, J. W., and J. L. Doust. 1993. Introduction to plant population biology. Blackwell Science, Oxford.

Singer, F. J., D. K. Otto, A. R. Tipton, and C. P. Hable. 1981. Home ranges, movements, and habitat use of European wild boar in Tennessee. Journal of Wildlife Management 45:343-353.

Sokal, R. R., and N. L. Oden. 1978a. Spatial autocorrelation in biology 1. Methodology. Biological Journal of the Linnean Society 10:199-228.

Sokal, R. R., and N. L. Oden. 1978b. Spatial autocorrelation in biology 2. Some biological implications and four applications of evolutionary and ecological interest. Biological Journal of the Linnean Society 10:229-249.

Suzuki, J.-I., T. Herben, F. Krahulec, and T. Hara. 1999. Size and spatial pattern of *Festuca rubra* genets in a mountain grassland: its relevance to genet establishment and dynamics. Journal of Ecology 87:942-954.

Tani, N., N. Tomaru, Y. Tsumura, M. Araki, and K. Ohba. 1998. Genetic structure within a Japanese stone pine (*Pinus pumila* Regel) population on Mt. Aino-Dake in Central Honshu, Japan. Journal of Plant Research volume:6-15.

Ueno, S., N. Tomaru, H. Yoshimaru, T. Manabe, and S. Yamamoto. 2000. Genetic structure of *Camellia japonica* L. in an old-growth evergreen forest, Tsushima, Japan. Molecular Ecology 9:647-656.

Wainhouse, D., and R. Deeble. 1980. Variation in susceptibility of beech (*Fagus* spp.) to beech scale (*Cryptococcus fagisuga*). Annales Sciences Forestieres 37:279-290.

Ward, R. T. 1961. Some aspects of the regeneration habits of the American beech. Ecology 42:828-832.

Waycott, M. 1998. Genetic variation, its assessment and implications to the conservation of sea grasses. Molecular Ecology 7:793-800.

Weller, S. G., K. H. Keeler, and B. A. Thomson. 2000. Clonal growth *of Lithospermum caroliniense* (Boraginaceae) in contrasting sand dune habitats. American Journal of Botany 87:237-242.

Whitney, H. E. and W. C. Johnson. 1984. Ice storms and forest succession in southwestern Virginia. Bulletin of the Torrey Botanical Club 111:429-437.

Whittaker, R. H. 1956. Vegetation of the Great Smoky Mountains. Ecological Monographs 26:1-80.

Wiggins, G. J. 1997. Temporal incidence, progression, and impact of beech scale, *Cryptococcus fagisuga* Lindinger, and beech bark disease in the Great Smoky Mountains National Park. Department of Entomology and Plant Pathology. The University of Tennessee, Knoxville.

Wolfe, A., and A. Liston. 1998. Contributions of PCR-based methods to plant systematics and evolutionary biology. Pp. 43-86 *in* D. Soltis, P. Soltis and J. Doyle, eds. Plant Molecular Systematics II. Kluwer, Boston.

Yeh, F., R.-C. Yang, T. Boyle, Z.-H. Ye, and J. Mao. 1997. POPGENE, the user friendly shareware for population genetic analysis. Molecular biology and biotechnology centre, University of Alberta, Canada.

Zietkiewicz, E., Antoni Rafalski, and Damian Labuda. 1993. Genome fingerprinting by simple sequence repeat (SSR)-Anchored Polymerase Chain Reaction Amplification. Genomics 20:176-183.

# VITA

Ashley Lynn Brigham was born on 24 October 1974 in Mobile, Alabama. At age five, she and her family moved to Alexander City, Alabama. Ashley attended the Alexander City public school system, graduating from Benjamin Russell High School in May 1993. In May 1997, she graduated from the University of the South, Sewanee, Tennessee, with a BS in Natural Resources and a minor in Biology. The following August Ashley married Christopher G. Morris and officially changed her name to Ashley Brigham Morris. Ashley spent a year working as a research technician for the Department of Entomology at the University of Tennessee, Knoxville. She began her graduate career in August 1998 in the Department of Botany at the University of Tennessee. She was awarded a Master of Science in Botany in May 2001. She is currently a PhD student in the Department of Botany at the University of Florida in Gainesville.