ABSTRACT

PETERS, MELINDA DEAN. Genetic analyses of the federally endangered *Echinacea laevigata* using amplified fragment length polymorphisms (AFLP) – Inferences in population genetic structure and mating system.

(Under the direction of Dr. Qiu-Yun (Jenny) Xiang)

Echinacea laevigata is a federally endangered species and a close relative of the medicinally important *E. purpurea*. It was listed by the US Fish and Wildlife Service on October 8, 1992. The species has 24 recognized populations restricted to four states (VA, NC, SC, GA). Since prior knowledge of the mating system is limited, an estimation of outcrossing will provide information if gene flow can occur across the range of the population. Information about the genetic structure and mating system will guide future management efforts of this species. To determine the population structure and outcrossing rate across the range of the species, I conducted AFLP analysis using four primer combinations for 22 populations. The genetic diversity of this species was found to be high based on the level of polymorphic loci (200 of 210 loci; 95.24%) and Nei's gene diversity (ranging from 0.1398 to 0.2606; overall 0.2611). There was significant population genetic differentiation (G_{ST} of 0.2941), suggesting possible adaptation to local environments. Results from AMOVA analysis suggest that a majority of the genetic variance is attributed to variation within populations (70.26%). These results are congruent with a previous allozyme study that examined the genetic makeup of 11 of the 24 populations, excluding the Georgia populations. An isolation by distance analysis indicated that genetic differentiation among populations is a function of geographic distance, although long-distance gene dispersal (LDGF) between some populations was suggested based on population relationships from phylogenetic analysis. An estimate of

the outcrossing rate based on genotypes of progenies from six of the 22 populations using a method of multilocus estimate ranged from 0.833-1.2, where 1.2 is considered complete outcrossing, suggesting that the species is predominantly outcrossing. The data provide an estimate of genetic diversity and structure, which can be used to assist in the conservation of the species.

GENETIC ANALYSES OF THE FEDERALLY ENDANDERED *ECHINACEA LAEVIGATA* USING AMPLIFIED FRAGMENT LENGTH POLYMORPHISMS (AFLP) – INFERENCES IN POPULATION GENETIC STRUCTURE AND MATING SYSTEM

by
MELINDA DEAN PETERS

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Master of Science

BOTANY

Raleigh

2005

APPROVED BY:

_____ _____

_____
Chair of Advisory Committee

BIOGRAPHY

I was born in Roanoke, Virginia on Novemeber 22, 1980 to Rebecca and William Peters. I grew up in the small town of Clifton Forge, Virginia located in the Allegheny Mountains. At the time I did not appreciate the beauty of this town, but now I realize that growing up there made me appreciate the wonders of nature, which I feel has played a role in where I am today. I graduated from Alleghany High School located in Covington, Virginia during June of 1999. I then went on to attend James Madison University, where I received a Bachelor of Science degree in Biology (2003). As an undergraduate at JMU, I worked on a flora project, "Noteworthy Additions to the Vascular Flora of Rockingham County, Virginia," with Dr. Conley K. McMullen, which was eventually published in the journal *Castanea* (March 2005). In the fall of 2003, I enrolled in the botany department of North Carolina State University, under the direction of Dr. Jenny Xiang, to pursue a master of science with interests in conservation and population genetics. After graduation, I plan to continue working in the field of plant science and conservation, while preparing to pursue a Ph.D. degree in this area.

ACKNOWLEDGMENTS

I would first like to begin by thanking my committee members, Dr. Jon Stucky, Dr. Ed Vargo, and Dr. Jenny Xiang. Special thanks to Jenny, my major advisor, for taking a chance on me and guiding me through the process of graduate school. To Dr. Stucky who initiated an interest in this plant, and to Dr. Vargo for letting me use his lab plus helping with technical and conceptual problems. Huge thanks to our lab tech, David Thomas, for putting up with my continuous questions and teaching me the lay of the land. I would still be floundering about if it were not for your patience, understanding, and willingness to help me even on the simple problems. To all the other faculty members whose classes I have attended, and especially the two graduate directors, Dr. Nina Allen and Dr. Becky Boston.

My undergraduate advisor, Doc, thanks for continuing to push me to do better and open doors that I would not have figured out on my own.

There are many others to thank for helping get this project underway, and I have to first start with Dale Suiter. I call him MY "Fish and Wildlife Man" only because he is the leading USFWS botanist for this species and has helped with collecting efforts and seeking funding to support the study.

Thanks to all of the state and federal offices that approved my permits (I had a total of 5), and to others who have helped with field collecting and various other aspects. The long list of people starts from Virginia: Mike Leahy, Mountain Region Steward with the Division of Natural Heritage, and Ben Asma; from North Carolina: Laura Gadd, and Andy Walker, both from NCSU, and Sudie Davis; from South Carolina: Robin Roecker at US Forest Service; and from Georgia: James Sullivan.

Special thanks to the NC State Phytotron Staff for taking care of my plants and allowing me the space for part of this project.

To Kathryn Cherry, my undergraduate assistant, thank you for your help with some of the fieldwork and labwork, which included DNA extractions. I particularly would also like to thank Kathy McKeown for helping with the germination procedure of this plant, and thanks to Rob Sutter for taking some of the seedlings I harvested and transplanting them at Penny's Bend Nature Preserve. Thanks to Jeff Essic, the GIS "master," for helping me to make my maps and determine geographic distance between my populations.

Thanks to my fellow-lab mates, Wenheng Zhang and Jennifer Modliszewski, for their help in the lab and more importantly out of the lab. I cannot leave out my fellow graduate students, thanks for suffering along with me; it would not have been as much fun without you all. To the "ladies" of the Botany office, it was so nice to have five other mothers to brighten my day and appreciate my cute shoes. To Sue Vitello, the goddess of making graduate school easy on students, I hope you know that we could not make it without you! To Meghan Rothenberger, my roommate, I will always remember all of the support, laughs, and of course great meals that we shared.

Thanks to Dr. Xiang and the Virginia Nature Conservancy office through Jennifer Allen, Conservation Ecologist, for providing support for this project and to the Botany and Biology departments for providing my stipend.

Lastly, but not least, to my parents, thank you for always being supportive of everything I always wanted to pursue. I feel so fortune to have both of you in my life,

and I am so thankful and proud to be your daughter. It takes wonderful parents to deal with a daughter in graduate school, so thanks a ton!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AMOVA – analysis of molecular variance
bp – base pair
DNA – deoxyribonucleic acid
dNTPs – dinucleotide triphospates
GA – Georgia
Km – Kilometers
LDGD – Long distance gene dispersal
NC – North Carolina
NCSC – North Carolina State University Herbarium
NCSU – North Carolina State University
PCR – polymerase chain reaction
SC – South Carolina
VA – Virginia

Genetic analyses of the federally endangered *Echinacea laevigata* using amplified fragment length polymorphisms (AFLP) – Inferences in population genetic structure and mating system

Melinda D. Peters[1], Qiu-Yun (Jenny) Xiang[1], David T. Thomas[1], Jon Stucky[1], and Edward Vargo[2]

1 *Department of Botany, North Carolina State University*

2 *Department of Entomology, North Carolina State University*

ABSTRACT

Echinacea laevigata is a federally endangered species and a close relative of the medicinally important *E. purpurea*. The species has 24 recognized populations restricted to four states (VA, NC, SC, GA). To determine the population structure and outcrossing rate across the range of the species, we conducted AFLP analysis using four primer combinations for 22 populations. The genetic diversity of this species was high based on the level of polymorphic loci (200 of 210 loci; 95.24%) and Nei's gene diversity (ranging from 0.1398 to 0.2606; overall 0.2611). There was significant population genetic differentiation (G_{ST} of 0.2941), suggesting restricted gene flow and possible adaptation to local environments. Results from the AMOVA analysis suggest that a majority of the genetic variance is attributed to variation within populations (70.26%). These results are congruent with a previous allozyme study that examined the genetic makeup of 11 of the 24 populations, excluding any Georgia populations. An isolation by distance analysis indicated that genetic differentiation among populations is a function of geographic distance, although long-distance gene dispersal (LDGF) between some populations was suggested based on the NJ tree. An estimate of the outcrossing rate based on genotypes of progenies from six of the 22 populations using a multilocus estimate was 0.833-1.2, where 1.2 is considered complete outcrossing, suggesting that the species is predominantly outcrossing. These results are encouraging, suggesting that there is abundant genetic diversity in the existing populations of *E. laevigata* which may be sustained by the predominant outcrossing mating system of the species if there is no decrease in populations size; therefore, management efforts can focus on other issues instead of increasing genetic diversity.

Keywords:

Echinacea laevigata, genetic structure, outcrossing rate, AFLP

INTRODUCTION

Many factors that influence the continued existence of plant populations become more important when the species is endangered. A partial list includes the consequence of human interactions, availability of habitat, fragmentation and isolation of habitat, type of mating system, population size, and stochastic factors (Frankham, 2004; Takami, 2004; Zawko, 2001). An important factor that influences the long-term survival of threatened and endangered plant species is genetic variation, which in turn, is affected by population size, genetic drift, inbreeding depression, bottleneck effects, and natural selection (Frankham, 2004; Ribeiro, 2002). A major goal in plant conservation is, thus, to preserve existing genetic variation and minimize the processes that reduce this variation (Neel et al., 2001). Therefore, conservation genetics is motivated by the need to reduce current rates of extinction and to preserve biodiversity, not only for economic reasons, but for bioresources, ecosystem services, aesthetics, and for the rights of living organisms to exist (Frankham, 2004).

Genetic variation and population structure have been the key variables in assessing long-term viability of rare or fragmented populations, and hence impact the conservation strategies for these species (Kephart, 2004). These variables have provided essential information to the development of sound management strategies for the long-term conservation of many rare plant species (Falk and Holsinger, 1991). While genetic variation and population structure indicate the present status of a species, mating system, the manner in which gametes combine to form the next generation, defines the future pattern of genetic diversity of the species, and is crucial to the sustainability of populations and the species as a whole. Therefore, knowledge of mating systems of a rare species is important in predicting

future genetic diversity of the species and assisting in the development of appropriate conservation strategies (Neel et al., 2001). An example of partitioning resources in plant conservation could be to protect the highly diverse or differentiated populations while management efforts to restore diversity could be focused on the reduced populations (Neel, 2003). Flowering plants have evolved different mating systems, from strictly selfing (self-compatible) to strictly outcrossing (self-incompatible). The mating system of species is closely related to the level and pattern of gene flow (Barrett, 2003); thus, permitting the prediction of these variables in a species if the mating system is known. Outcrossing species typically have higher levels of gene flow within and among populations than selfing species, and hence, are able to maintain a higher level of genetic diversity within species and populations and show lower genetic differentiation among populations (Zawko et al., 2001). In contrast, selfing species usually show the reverse and suffer various levels of inbreeding depression (Neel et al., 2001). However, many flowering plants display a mixed model mating system, which is a combination of selfing and outcrossing (Kephart, 2004). Low outcrossing rates and inbreeding in such species may lead to decreased genetic diversity if no special management strategy is taken to promote outcrossing.

Background of the study species — *Echinacea* is a genus of the sunflower family, Asteraceae, and contains the medicinally important species *E. purpurea* (L.) Moench. The genus has been through a number of classification changes since it was first described in the 18th century by European explorers (Binns et al., 2004; McGregor, 1968). It was not until 1968 when Ronald McGregor published "The Taxonomy of the Genus *Echinacea* (Compositae)" that a better understanding of the genus began to divulge. McGregor recognized nine species and two varieties in *Echinacea*. In 2002 Binns et al. recognized four

species and six varieties within two subgenerea based on a morphometric numerical analyses. One of the species, *Echinacea laevigata* (Boynton and Beadle) Blake, commonly known as smooth coneflower, was first listed as a federally endangered species by the U.S. Fish and Wildlife Services on October 8, 1992 (USFWS, 1995). The species is still listed as federally endangered and is now found only in the states of Virginia, North Carolina, South Carolina, and Georgia (USFWS, 1995), with 24 extant populations. USFWS (1995) reported that *E. laevigata* historically had a total of 62 populations from 26 counties in eight states (Pennsylvania, Maryland, Virginia, North Carolina, South Carolina, Georgia, Alabama, and Arkansas). Later, the reports from Alabama and Arkansas were thought to be erroneous (Gaddy, 1991). Once this species was listed, a recovery plan was developed and published by the U.S. Fish and Wildlife Service in 1995. This plan provides a summary of the species and outlines reasonable actions that would allow this species to be "recovered and/or protected" (USFWS, 1995). Current management efforts are underway in all states as they monitor populations that are owned by federal or state agencies. However, whether all the known populations will be self-sustaining under the recovering/protection plan remains to be assessed. Key information on genetic structure and mating patterns of the species needs to be gathered to allow the evaluations of current and future status of the species. A genetic study on *E. laevigata* can inform management decisions, including reintroduction and translocation of populations within the species "historical" range (task 2, priority 2) and for the preparation of a Propagation Plan for *E. laevigata*, pursuant to the USFWS Policy on Controlled Propagation of Listed Species (USFWS, 2000). Any seeds collected for such a study will also assist in the maintenance of a seed bank (task 6, priority 1) at the North Carolina Botanical Garden.

Echinacea laevigata (Figure 2) is an herbaceous perennial (Gaddy, 1991) with an erect thick, rarely branched, glabrous stem of up to 1.5 m tall (McGregor, 1968). The leaves form a basal rosette and the blade can become up to 2-6.5 cm x 20 cm with a long petiole (McGregor, 1968). The flowering heads are usually solitary, with a disk 1.5-3.4 cm in diameter and ray flowers 3.5-8 cm long that are purplish to light pink in color (McGregor, 1968).

The species grows in various habitats including open woods, cedar barrens, roadsides, dry limestone bluffs, power line rights-of-way, and other sunny to partly sunny situations, usually on magnesium and calcium-rich soils associated with amphibolite, dolomite, or limestone (VA); gabbro (NC and VA); diabase (NC and SC); and marble (SC and GA) (USFWS, 1995). The species flowers from May through July and the fruits develop in late June and mature in September (Gaddy, 1991). Little is known regarding the pollinators and breeding system of *E. laevigata*. However, a preliminary list of potential pollinators has been suggested from the South Carolina populations and includes bees, butterflies, and beetles (Order Hymenoptera, Lepidoptera, and Coleoptera; respectively) (USFWS, 1995). It has also been suggested that the seeds of *E. laevigata* are dispersed by seed-eating birds or small mammals (USFWS, 1995). Vegetative reproduction via new shoots emerging from underground rhizomes was reported (USFWS, 1995; Apsit and Dixon, 2001). The taxonomy study by McGregor in 1968, which included all species of *Echinacea*, reported that all plants of the genus are self-sterile based on a studied including 500 bagged heads. A more robust study is being conducted that examines the pollination biology of *E. laevigata* (Gadd and Stucky, 2005). This study involved bagging individual flowering heads and performing artificial pollinations among the flowers of bagged heads to test for self-compatibility and

self-pollination. The unpublished results indicate that no viable seeds were produced in either of the two pollination treatments suggesting that *E. laevigata* is self-incompatible and requires pollinators for effective pollination (Laura Gadd, NCSU, personal communication). It was also observed that the floral development of *E. laevigata* displays a dichogamous system (Laura Gadd, NCSU, personal communication). Dichogamy is defined as a temporal separation of the maturation of the pollen and the stigma presentation within flowers on the same flowering head (Barrett, 2003). This type of system helps to promote outcrossing, or cross-pollination among flowering heads.

It is noteworthy that among the 24 extant populations, a few are found in powerline rights-of-way. At these powerline sites, establishment of woody vegetation is precluded by mowing and managed burns; therefore, allowing the flora to be more diverse compared to sites that are not as well mowed or burned (USFWS, 1995). *Echinacea laevigata* tends to be more abundant and flowers more in burned or mowed locations than in unmanaged sites.

Major threats to the smooth coneflower include habitat loss via highway right-of-way maintenance, urbanization and suburbanization of the habitat, over-collection, fire suppression, encroachment by exotic species, possible predation by insects, inadequacy of existing protection afforded by State laws, and small population size (USFWS, 1995).

To evaluate the genetic diversity and population genetic structure of the species, Apsit and Dixon (2001) examined 28 allozyme loci for 11 populations of *E. laevigata* from Virginia, North Carolina, and South Carolina. The study found that the 11 populations contained moderate levels of genetic diversity (H_e = 0.178) and exhibited significant population structure (G_{ST} = 0.109). A significant, positive correlation between pairwise genetic and geographic distances among the 11 populations (ρ = 0.38; $p \leq 0.025$) was also

detected in the allozyme study, which suggests that gene flow in the species is restricted by geographic distance. The hierarchical analysis of molecular variance (AMOVA) of the allozyme data indicated that significant genetic variation was partitioned among states, among populations within states, and within populations (Φ_{CT} = 0.105; Φ_{SC} = 0.127; Φ_{ST} = 0.219, respectively); with the greatest genetic variance partitioned within populations. Based on these results, the authors suggested that *E. laevigata* might be adapting to local environmental heterogeneity. They further indicated that analysis of mating system would be very useful to assess levels of inbreeding and outcrossing.

The goals of this study were 1) to evaluate the genetic variation and structure of *E. laevigata* with a more complete sampling of populations using more variable molecular markers, amplified fragment length polymorphisms (AFLP) and 2) to estimate the outcrossing rate of the species using the AFLP data from progeny grown from known mothers. This study will provide a finer estimate of genetic diversity, gene flow within and among populations, and structuring across the range of the species to assist in conservation management of the species.

MATERIALS AND METHODS

AFLP Markers — AFLP (amplified fragment length polymorphism) markers are employed in the present study to examine patterns of genetic variation and population structure. AFLP markers provide genetic data from a large number of loci without requiring prior knowledge of DNA sequences and are relatively "fast and easy" to obtain (Kjølner et al., 2004). The analysis of AFLP loci generates multi-locus genotypic data for comparison across individuals and populations and permits the survey of numerous DNA loci across the

entire nuclear genome by using multiple PCR primers to allow the detection of genetic variations between closely related individuals (Vos et al., 1995). AFLP markers are reproducible and have been widely accepted (Kjølner et al., 2004). These markers provide valuable insight to rare species management (Coart et al., 2005; Gaudeul et al., 2000; Palacios et al., 1999; Schmidt and Jensen, 2000; Tero et al., 2003; and Zawko et al., 2001).

Plant material — Leaf material was collected from 22 of the 24 reported populations for genotyping using AFLP analysis (Figure 1). A total of 420 individuals were chosen randomly across populations. The number of individuals sampled per population ranged from 10-23 (Table 1).

Material from six of the 22 populations was used for outcrossing rate estimation (Table 2). Ten to 30 heads each from a different randomly chosen individual (mother plant) were collected from each of these six populations. Fruits were soaked for 24 hours in a 1mM ethephon solution and stratified on wet blotters in Petri dishes at 4ºC for two weeks (McKeown and Widrlechner, 2003). After stratification the fruits were planted in small Styrofoam cups and placed in a misting bed in the North Carolina State University Phytotron Greenhouse for up to 3 weeks (Figure 3). Leaf tissue was collected for DNA extraction. One greenhouse-raised plant was collected as a voucher and deposited in NCSC (North Carolina State University Herbarium).

DNA extraction and AFLP genotyping — Genomic DNA was isolated from fresh leaf material collected from field and greenhouse specimens using a modified cetyltrimethylammonium bromide (CTAB) method of Doyle and Doyle (1987) with modifications described in Cullings (1992). To estimate the quality and quantity of DNA, a GeneQuant spectrophotometer (Amersham Pharmacia Biotech, Cambridge, England) was

used to measure the 260 and 280 UV absorbance, which is an estimate of DNA, and each sample was ran on a 1.0% agarose gel to compare band densities across individuals. Samples with 260:280 absorbance ratios of 1.8-1.9 and had consistent band densities were used for AFLP analysis.

AFLP reactions were preformed as described by Vos et al. (1995) with minor modifications. The restriction digest and ligation steps were done as separate reactions. For the digestion, approximately 500ng of genomic DNA was incubated at 37 °C for 3 hours in a 10 μL volume reaction containing 1X NEBuffer (New England Biolabs), 5 U *Eco*RI, 5 U *Mse*I, and 4.5 μg/mL BSA. Next, 5 μL of a ligation mix including 10X T4 DNA Ligase Buffer, 1 μM *Eco*RI-adapter, 5 μM *Mse*I-adapter, and 40 U T4 DNA Ligase was added to the sample and kept at room temperature for approximately 24 hours. After ligation the samples were diluted 10-fold with sterile deionized water (dH_20). A pre-selective polymerase chain reaction (PCR), using a Robocycler thermocycler (Stratagene, LaJolla, California, USA) was done using primer pairs with a single selective nucleotide extension (Table 3). The reaction mix (total volume of 20 μL) included 5 μL template DNA from the restriction/ ligation step, 4 U Taq DNA Polymerase, along with 10X Buffer B (Promega, Madison, Wisconsin, USA), 200 μM of each dNTP, 1.5 mM $MgCl_2$, 1 μg/mL BSA, 50 ng/μL *Eco*RI primer, and 50 ng/μL *Mse*I primer. After an initial incubation at 72 °C for 2.5 min, 30 cycles at 94 °C for 45 s, 56 °C for 45 s, and 72 °C for 2 min were preformed with a final extension at 60 °C for 10 min. Samples were diluted 10-fold with dH_20 and stored at –20 °C.

Preliminary tests were preformed to identify the most variable selective primer extensions, which included 12 primer combinations. Four primer combinations were selected (Table 3). Selective PCRs (total volume of 10 μL) included 3 μL of template from

the pre-selective step, 4 U Taq DNA Polymerase, with 10X Buffer B, 200 μM of each dNTP, 1.5 mM MgCl2, 1.5 μg/mL BSA, 1 μM *Eco*RI selective primer labeled with a fluorescent marker (LI-COR, Lincoln, Nebraska, USA), and 8.3 μM *Mse*I primer. PCRs were temperature cycled in 96-well plates using a Robocycler thermocycler with the following protocol: an initial incubation at 94 °C for 5.5 min, then 34 cycles at 94 °C for 1 min, 56 °C for 1 min, and 72 °C for 2 min with a final extension at 60 °C for 10 min.

After selective amplification, 8 μL of loading buffer (95% deionized formamide, 20 mM EDTA, 0.8 mg/mL bromphenol blue) was added to each sample. Samples were denatured at 90 °C for 4 min, placed on ice until loaded onto a 6.5% KB Plus polyacrylamide gel (LI-COR). Samples were electrophoresed for 3 h (45°C, 1500 V) on a LI-COR 4300L automated DNA sequencer. To provide additional control for comparing bands across gels, a few DNA samples were included in all gels of the same primer combination. Gels were scored using the AFLP Quantar Software (Keygene, 2003) and each fragment was treated as a separate locus and scored as "+" for the presence of a band and as "-" for the absence of a band. Ambiguous presence or absence of a band in a sample was scored as unknown "?". Overall, only fragments between 564 bp and 63 bp were scored.

Statistical Analyses — Linkage disequilibrium was calculated overall for all individuals using the ARLEQUIN program. To estimate the genetic diversity and the overall genetic structure of *E. laevigata*, various measures were calculated for the AFLP data using the computer program POPGENE (v.1.31, Yeh et al., 1999). The percentage of polymorphic loci within and among populations was determined. A G_{ST} estimate (Weir, 1990), which is equivalent to the traditional F-statistics, was computed to estimate the genetic differentiation

among populations. Nei's (1978) gene diversity was calculated for each population, using the Popgene program, based on unbiased estimates for the different loci.

To estimate the relationships among populations, Nei's unbiased pairwise genetic distance data (after Lynch and Milligan, 1994) was calculated by the program AFLP-SURV 1.0 (Vekemans, 2002) and imported into PAUP 4.0* (Swofford, 2002) for phylogenetic analysis using the Neighbor-Joining method. The AFLP-SURV program was also used to calculate 1000 bootstrap distance matrices and these were opened with the NEIGHBOR-JOINING program in PHYLIP (Felsenstein, 1993) in order to obtain a tree file and elucidate evolutionary relationships among populations. The CONSENSE program from the PHYLIP software was used to compute a majority-rule tree to estimate support for the Neighbor-Joining tree.

The computer program ARCVIEW (Version 3.3) was used to determine geographic distance using GPS coordinates for all of the 22 populations. These data were used in conjunction with the genetic data for a Mantel test to determine if there was significant isolation by distance (Bohonak, 2002). To further examine the partitioning of genetic variation, the program ARLEQUIN (Schneider et al., 2000) was used to perform an analysis of molecular variance (AMOVA) to assess the hierarchical genetic structure among populations and within populations. The program TRANSFORMER-3b.01 was used to convert the AFLP data matrix into the appropriate ARLEQUIN input matrix (Caujapé-Castells and Baccarani-Rosas, 2005). To facilitate comparing the AMOVA results with the previous allozyme study (Apsit and Dixon, 2001), an AMOVA was also done by partitioning the populations among states, among populations within states, and within populations, as done in the allozyme study. The AMOVA analyses are based on the pairwise squared

Euclidean distances (Excoffier et al., 1992) and an estimated population differentiation Φ_{ST}, which is equivalent to F_{ST} was calculated using ARLEQUIN. The AMOVA with dominant markers assumes the mating pattern to be the same in all populations (Tero et al., 2003).

These analyses were also run by excluding individuals that had missing data for two or more loci, and the results are not presented because no significant difference was found. Only six individuals total (two from PH, one from GH, two from DC, and one from GA-006) from the 22 populations had missing data for two or more loci.

Outcrossing Rate Estimate — The MLTR (Multilocus Mating System Program) software from Ritland (2002) estimates parameters of the outcrossing rates based on a multilocus outcrossing theory (Ritland and Jain, 1981). We estimated the multilocus outcrossing rate (t_m) using this method with AFLP data from progenies of six populations (Table 2) using the mixed mating model described by Ritland and Jain (1981). Since this is a dominant marker, by default the program encodes a diploid locus by a single allele rather than a pair of alleles (Ritland, 2002). Standard errors were estimated based on 1000 bootstraps between individuals within progeny arrays.

RESULTS

Gene diversity, gene flow, and population structure — Linkage disequilibrium was examined between all pairs of polymorphic loci across all individuals regardless of population or state. There were 28,681 combinations, of which 9790 were significant at least at the 5% level. The overall percentage of linkage disequilibrium was 34%, and is higher than what is expected by chance. This suggests that some loci may be linked, but without

further analysis into the separate populations the specific loci cannot be identified and may be causing this inflation.

A total of 210 loci were scored from 420 individuals based on the results from four primer pairs. Of these loci, 200 (95.24%) were polymorphic across all populations (Table 1). The percentage of polymorphic loci per population ranged from 39.05% to 70.00% (SHP and PH populations, respectively) as shown in Table 1. Gene diversity within populations based on Nei's gene diversity (h) ranged from 0.1398 to 0.2606 (SHP and PH populations, respectively), with the overall level of 0.2611 (Table 4).

The overall estimate of genetic structure (G_{ST}) was 0.2941. Genetic differentiation based on pairwise Φ_{ST} comparisons between populations ranged from 0.069 (lowest distance between populations PH and DC) to 0.529 (highest distance between populations GH and RT 76-001) (Table 5). All of the genetic distances in the matrix were significantly different from zero at the $p < 0.05$ levels except between populations GA-006/GA-001 and populations CCS/RT 76-001. A positive and significant ($p < 0.002$) correlation between genetic differentiation, Φ_{ST}, and geographic distance (Km), was detected (Figure 4). The neighbor-joining phylogram constructed for all populations based on Nei's unbiased pairwise genetic distance data was overall not well supported by bootstrap values (10 out of 19 nodes had bootstrap values above 50%; Figure 5). However, populations clustered basically by geographic region (Figure 6) and a finding congruent with the AMOVA results (see below). Exceptions were found for three populations, SH from NC, JC from VA, and GA-022 from GA. SH and JC united with populations from GA, while GA-022 grouped with populations from SC. Results of all AMOVA analyses indicate that most of the variance occurred within

populations (Table 6). There was significant variance among populations (29.74%) from the analysis, by grouping of within and among populations.

Outcrossing rate estimates — The multilocus outcrossing rate (t_m) estimated for the six populations were high, ranging from 0.833 to 1.2 (GH to PCPL/PCWL/SH) respectively, Table 7). The three populations sampled from North Carolina all displayed an outcrossing estimate of being complete outcrossing (1.2). The three populations from Virginia had lower outcrossing estimates, which indicate that more than 80% of the mating system relies on outcrossing. Small sampling size could contribute to the low estimates obtained from the Virginia populations because the program suggests that 20-30 progeny per mother are generally needed for robust estimates. Our sampling of progeny was low, for example, the entire GH population had only 24 individuals (Table 2). Thus, population size could have contributed to the discrepancy of t_m estimates between the North Carolina and Virginia populations. Nonetheless, values suggest that these populations are predominantly outcrossing.

DISCUSSION

Genetic variation — The overall level (34%) of linkage disequilibrium found for the species suggests that some loci are linked. Causes of linkage disequilibrium, or non-random association of alleles, are due to population bottlenecks, recent mixing of different populations and selection (Frankham et al., 2004). Usually in large outcrossing populations, alleles randomly associate and are in linkage equilibrium (Frankham et al., 2004). Hill and Robertson (1968) proposed that under genetic drift and mutation, the estimate of linkage disequilibrium is expected to be small in a large population even between closely linked loci.

Since this is an endangered species, and the results suggest that some populations are becoming genetically isolated, it is expected that some of the loci might be linked due to small population size (Slatkin, 1994). Another reason for the high percentage estimated in this study is likely to be inflated due to differentiation among populations and further analyses into individual population estimates are needed to determine if only a few loci are contributing to this estimate.

It was predicted that *E. laevigata* would display lower levels of genetic diversity, but based on the results, it is evident that diversity levels are higher than expected for an endangered species. The overall genetic diversity estimate and percentage of overall polymorphic loci in *E. laevigata* populations based on AFLP (0.2611 and 95.24% polymorphic loci; Table 4 and Table 1, respectively) is higher but comparable to the estimates from a previous allozyme study (0.178 and 63.6% polymorphic loci; Apsit and Dixon, 2001). Other species that are comparable to *E. laevigata* and display a predominantly outcrossing system also had high percentages of polymorphic loci (92.07% in Cardoso et al., 2000; 64% in Schmidt and Jensen, 2000; 82% in Shrestha et al., 2005; and 89% in Zawko et al., 2001). Population size and breeding system are factors that contribute to higher levels of diversity within populations along with the effects of processes like selection, genetic drift, mutation, and migration (Hamrick and Godt, 1996; Frankham et al., 2004). Total population sizes are not reported here, but the largest population is PCPL and one of the smaller populations is RT 76-001 (personal observation). The percentage of polymorphic loci for both populations respectively was 44.76% and 50.48% (Table 1). Based on these results, population size does not seem to be a factor affecting the levels of diversity. Sizes of populations can change due to migrants into or out of populations, and intermediate and

lower levels of gene flow are affected by varying number of migrants into a population (Frankham et al., 2004). Wright (1969) reported that a single migrant was enough to prevent complete differentiation and fixation and that higher migration rates in small populations counteract their greater loss of variation due to drift.

Recall that breeding systems can also affect the levels of diversity in populations. Mostly outcrossing species usually have higher levels of diversity due to the nature of their breeding system and because they rely on some sort of pollinator to enable reproduction. *E. laevigata* can reproduce clonally, which increases the risk of lower genetic diversity and could become evident in smaller populations because flowering heads are forming from the same vegetative rootstock. So even though *E. laevigata* is thought to be self-incompatible, Frankham et al. (2004) reports that self-incompatibility is regulated by one or more loci that may have up too 50 alleles in large populations and in small populations these alleles are lost by random sampling, which may reduce the number of plants. An example of this can be found in the Lakeside daisy (*Hymenoxys acaulis* (Pursh) K.F. Parker var. *glabra* (A. Gray) K.F. Parker), which showed decreased fitness due to the loss of self-incompatible alleles (Demauro, 1993). Therefore, lower genetic diversity is not evident in small populations of *E. laevigata*, and the risk of losing genetic diversity in the future in small populations may occur via random loss of alleles.

Genetic differentiation (G_{ST} 0.2941) among populations of *E. laevigata* was significant suggesting populations are genetically isolated. However, the G_{ST} from this study is higher than that from the allozyme data for *E. laevigata* (0.109; Apsit and Dixon, 2001). The differentiation estimate from other studies working on predominantly outcrossing species show a range of being genetically isolated to mixing of the gene pool among

populations (F_{ST} = 0.426 in Cardoso et al., 2000; G_{ST} = 0.27-0.89 in Schmidt and Jensen, 2000; and in Zawko et al., 2001, they reported low differentiation). Gaudeul et al. (2000) reported that the endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), which displays a mixed mating system, had a G_{ST} estimate of 0.42. Current literature reports (mostly isozyme data), that genetic differentiation F_{ST} for outbreeding species is 0.2 and for inbreeding or selfing species, 0.5 (Loveless and Hamrick, 1984; Hamrick and Godt, 1990). Of course this could be highly variable and it depends on the spatial scale plus the degree of isolation among populations. It is evident from the examples given and the results from this predominantly outcrossing species, *E. laevigata*, that gene flow across the range of populations plays a role in defining how genetically differentiated they are from another.

The isolation by distance analysis (Figure 4) shows a significant ($p < 0.05$) correlation between genetic and geographic distances except for two intercepts (Table 5). The intercepts of the populations that deviate from the others are between populations GA-001/GA-006 and the intercept between populations RT 76-001/CCS. An explanation for these two intercepts is that they are geographically close; therefore, they are also genetically close and are acting as one large population. Overall, the results are similar to the allozyme study, which also found a significant relationship between distance and genetic differentiation results ($p \leq 0.025$), suggesting that the populations further away geographically are also more genetically isolated. When compared to other studies on rare species it has been reported that they follow a significant isolation by distance model (Cardoso et al., 2000; Gaudeul et al., 2000; and Schmidt and Jensen, 2000). This indicates that geographic distance may be a factor in the distribution of genetic variation among populations of rare or endangered species.

The populations based on the results from the neighbor-joining tree clustered together by geographic region (Figure 6). The three exceptions (Figure 5) that did not cluster by geographic region are SH from North Carolina, JC from Virginia, and GA-022 from Georgia. One explanation could be due to long distance gene flow by pollinators or these populations have a share a common ancestor. The phylogram is congruent with the isolation by distance model and suggest that populations spatial related are more similar genetically. Pollinators are expected to have the potential to move between the populations that are geographically close, which is an explanation for the populations within close range to group together on the phylogram. Historically, it is known that more habitats were available and populations had fewer barriers such as human designed structures, which most likely allowed pollinators an easier route to populations across the range of the species.

Even though the AMOVA revealed that most of the molecular variance resides within populations (70.26%; Φ_{ST} = 0.297), there was still significant ($p < 0.0001$; 1023 permutations) variation among populations (29.74%), which suggests slight population differentiation (Table 6). Results of the AMOVA that included the among states source of variation (Table 6), also found that most of the variance occurs within populations. This is comparable to results previously reported for *E. laevigata* in the allozyme study (Apsit and Dixon, 2001; among states, 10.53 %, Φ_{CT} = 0.105; among populations within states, 11.43 %, Φ_{SC} = 0.127; and within populations, 78.04 %, Φ_{ST} = 0.219). It was important to compare these results to those including a natural hierarchical structuring among populations and within populations because states have no natural geographic barriers and provide arbitrary boundaries. When populations were grouped by regions in the AMOVA, the results also showed that the majority of variance was within populations (Table 6).

There are several previous reports each indicating that the majority of genetic variation occurs within populations of rare species. Gaudeul et al. (2000) found that 43.5% of the variation was found among populations whereas 56.5% was detected within populations; Tero et al. (2003) found 36.92% among and 63.08 within populations; Schmidt and Jensen (2000) reported 43.61% among and 51.76% within subpopulations; Zawko et al. (2001) found 10.3% among and 89.7% within populations. A reason that the within population diversity is so high for these species could be due to the frequency of outcrossing, which is often the most important determinant of population genetic structure (Hamrick and Godt, 1990).

Outcrossing Rates – The multilocus outcrossing rate estimated for six populations *E. laevigata* (t_m) was high suggesting that this species displays a predominantly outcrossing mating system. A species that displays complete outcrossing, based on the estimates from Ritland's program (MLTR, 2004), would be equal to 1.2. The range of estimates given by this program is zero to 1.2, where zero is complete selfing and 1.2 is outcrossing. The outcrossing rates estimated in this study for the North Carolina populations were 1.2 and those for the Virginia populations were 0.833, 0.844, and 0.876. A reason for the Virginia populations straying from what was expected could be due to the sampling size. Recall that the program suggests having 20-30 progeny per mother and that was not the case for the populations from Virginia (Table 2). However, the estimates are still considered high enough when compared to other studies (Neel et al., 2001 and Gaiotto et al., 2003) to suggest that this is a predominantly outcrossing species. Other support that this is an outcrossing species are based on the results from the bagging experiments of single heads by McGregor (1968), the unpublished data from Laura Gadd (2005), and the floral Dichogamy, which all

suggest that this species is self-incompatible. Based on this support and the outcrossing estimates it is likely that *E. laevigata* relies on a cross-pollination mating system.

Other studies have used this method to determine outcrossing rates and found similar results. A study on the endangered *Eriogonum ovalifolium* var. *vineum* (Small) Jepson (Polygonaceae), reports that the species is largely outcrossed based on family multilocus rates (t_m = 0.88 {SE 0.03}; Neel et al., 2001). However, they did note that the family estimates were based on relatively small within-family sample sizes; therefore, they were not certain if the plant is self-compatible with a mechanism to promote outcrossing or if the plant is self-incompatible and outcrossing events went unnoticed due to a small number of loci examined (Neel et al., 2001). In 2003, Neel et al. published another study that described the genetic diversity of the species, but also explained more about the flowering of the plant. *E. ovalifolium* var. *vineum* was described as having perfect flowers clustered into inflorescences, and that the flowers on a plant open simultaneously, which provides ample opportunities for geitenogamous pollination, or fertilzation between flowers on the same plant. Another study looked into the mating system of *Euterpe edulis* Mart., commonly known as heart-of-palm, and found that it displayed a predominantly outcrossed mating system (t_m = 0.94) based on the multilocus estimate (Gaiotto et al., 2003). A previous study in 2000 on the same species used AFLPs to estimate the genetic diversity, but also reported that the opening of male flowers before female flowers, protandrous dicogamy, in *E. edulis* seemed to benefit outcrossing (Cardoso et al., 2000).

CONCLUSION

The remaining populations of *E. laevigata* have significant levels of population diversity, but do exhibit slight signs of population differentiation or isolation among populations. It is encouraging that gene flow is occurring to maintain the genetic diversity because the long-term survival of this species will depend on moderate to high levels of genetic diversity. The downfall is that populations seem to be clustering by regions and are showing early signs of becoming isolated from the each other. This is important for conservation agencies to be aware of because future management efforts can be geared to dealing with this problem and undergo preventive measures to ensure the continued high levels of diversity across the range of this species. Population size can affect future diversity levels if populations continue to decline. Estimates of genetic diversity based on population size in this study suggest that the number of flowering heads is not an immediate factor as long as habitat is available. A suggestion for management efforts in smaller populations would be to clear any encroachment and if possible do a prescribed burn. It is evident from the populations in the powerline rights-of-way that disturbance is very helpful and preferred by this species. Maintenance of populations would also allow and potentially increase the visitation of pollinators for this species because flowering heads would be more abundant and more accessible. If the species is to survive, persist, and eventually become delisted, priorities have to be set in order to meet the goals described by the recovery plan, which include maintaining genetic diversity.

ACKNOWLEDGMENTS

The authors thank Dale Suiter, Mike Leahy, James Sullivan, Sudie Davis, Kathryn Cherry, Andy Walker, Laura Gadd, and Ben Asma for helping with the collection of material. The NCSU Phytotron for providing space for the plants during the study. Others who aided in some aspect of this study were Jeff Essic, Kathy McKeown, and Kermit Ritland. This study was partial supported by a grant funded by the VA Nature Conservancy.

REFERENCES

APSIT, V.J. AND P.M. DIXON. 2001. Genetic Diversity and Population Structure in *Echinacea laevigata* (Boynton and Beadle) Blake, and Endangered Plant Species. *Natural Areas Journal* 21: 71-77.

BARRETT, S.C.H. 2003. Mating strategies in flowering plants: the outcrossing-selfing paradigm and beyond. *The Royal Society* 358: 991-1004.

BINNS, S.E., J.T. ARNASON, AND B.R. BAUM. 2004. Taxonomic History and Revision of the Genus Echinacea. *In* S.C. Miller [ed.], Echinacea: the genus Echinacea, 3-11. CCR Press, Boca Raton, FL, USA.

BINNS, S.E., B.R. BAUM, AND J.T. ARNASON. 2002. A taxonomic revision of the genus *Echinacea* (Heliantheae; Asteraceae). *Systematic Botany* 27: 610-632.

BOHONAK, A.J. 2002. IBD (Isolation By Distance): a program for analyses of isolation by distance. *Journal of Heredity* 93: 153-154.

CARDOSO, S.R.S., N.B. ELOY, J. PROVAN, M.A. CARDOSO, AND P.C.G. FERREIRA. 2000. Genetic differentiation of *Euterpe edulis* Mart. Populations estimated by AFLP analysis. *Molecular Ecology* 9: 1753-1760.

CAUJAPÉ-CASTELLS J. AND M. BACCARANI-ROSAS. 2005. Transformer-3b.01: A program for the analysis of molecular population genetic data. Exegen software.

COART, E., S. VAN GLABEKE, R.J. PETIT, E. VAN BOCKSTAELE, AND I. ROLDÁN-RUIZ. 2005. Range wide versus local patterns of genetic diversity in hornbeam (*Carpinus betulus* L.). *Conservation Genetics* 6: 259-273.

CULLINGS, K.W. 1992. Design and testing of plant-specific PCR primer for ecological and evolutionary studies. *Molecular Ecology* 1: 233-240.

DEMAURO, M.M. 1993. Relationship of breeding system to rarity in the lakeside daisy (*Hymenoxys acaulis* var. *glabra*). *Conservation Biology* 7: 542-550.

DOYLE, J.J. AND J.L. DOYLE. 1987. A rapid DNA isolation procedure from small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19, 11-15.

EXCOFFIER, L., P.E. SMOUSE, AND J.M. QUATTRO. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479–491.

FALK, D.A. AND K.E. HOLSINGER. 1991. Genetics and conservation of rare plants. Oxford University Press, New York, USA.

FELSENSTEIN, J. 1993. PHYLIP (Phylogeny Inference Package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle.

FRANKHAM, R., J.D. BALLOU, AND D.A. BRISCOE. 2004. Introduction to Conservation Genetics. Cambridge University Press, New York, USA

GADD, L.E. and J.M. STUCKY. 2005. Published abstract: Habitat loss and its effect on the pollination biology of *Echinacea laevigata*, smooth coneflower. North Carolina State University. *North Carolina Academy of Science.*

GADDY, L.L. 1991. The status of *Echinacea laevigata* (Boynton and Beadle) Blake. Unpublished report to the U.S. Fish and Wildlife Service, Asheville, NC. 24pp., plus appendices and maps.

GAUDEUL, M., P.TABERLET, AND I. TILL-BOTTRAUD. 2000. Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers. *Molecular Ecology* 9: 1625-1637.

GAIOTTO, F.A., D. GRATTAPAGLIA, AND R. VENCOVSKY. 2003. Genetic Structure, Mating System, and Long-Distance Gene Flow in Heart of Palm (*Euterpe edulis* Mart.). *Journal of Heredity* 94(5): 399-406.

HAMRICK, J.L. AND M.J.W. GODT. 1990. Allozyme diversity in plant species. In: *Plant population genetics, breeding and genetic resources*. Eds. A.H.D. Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir. Sinauer Sunderland, 43-63.

HAMRICK, J.L. AND M.J.W. GODT. 1996. Conservation Genetics of Endemic Plant Species. In: Conservation genetics: case histories from nature. Eds. J.C. Avise and J.L. Hamrick. New York: Chapman & Hall.

HILL, W.G., AND A. ROBERTSON. 1968. Linkage disequilibrium in finite populations. *Theoretical Applied Genetics* 38(6): 226-331.

KEPHART, S.R. 2004. Inbreeding and reintroduction: Progeny success in rare *Silene* populations of varied density. *Conservation Genetics* 5: 49-61.

KJØLNER, S., S.M. SÅSTAD, P. TABERLET, AND C. BROCHMANN. 2004. Amplified fragment length polymorphism versus random amplified polymorphic DNA markers: clonal diversity in *Saxifraga cernua*. *Molecular Ecology*, 13: 81-86.

LOVELESS AND HAMRICK. 1984. Ecological determinants of genetic structure in plant populations. *Annual Review of Ecology and Systematics* 15: 65-95.

LYNCH M. AND B.G. MILLIGAN. 1994. Analysis of population genetic structure with RAPD markers. *Molecular Ecology* 3: 91-99.

MCGREGOR, R.L. 1968. The taxonomy of the genus *Echinacea* (Compositae). *University of Kansas Bulletin*. 48: 113-142.

MCKEOWN, K. AND M. WIDRLECHNER. 2003. *Echinacea* seed stratification procedure. Unpublished protocol, modified from Sari et al. 1999. *Echinacea angustifolia,* an emerging new medicinal plant: Techniques to overcome germinations problems. In: Janick, J. (ed.), New crops and new uses. ASHS Press, Alexandria, VA.

NEEL, M.C., J. ROSS-IBARRA, AND N.C. ELLSTRAND. 2001. Implications of Mating Patterns for Conservation of the Endangered Plant *Eriogonum ovalifolium* var. *vineum* (Polygonaceae). *American Journal of Botany*, 88(7): 1214-1222.

NEEL, M.C., AND N.C. ELLSTRAND. 2003. Conservation of genetic diversity in the endangered plant *Eriogonum ovalifolium* var. *vineum* (Polygonaceae). *Conservation Genetics* 4: 337-352.

NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

PALACIOS, C., S. KRESOVICH, AND F. GONZÁLEZ-CANDELAS. 1999. A population genetic study of the endangered plant species *Limonium dufourii* (*Plumbaginaceae*) based on amplified fragment length polymorphism (AFLP). *Molecular Ecology*, 8: 645-657.

RIBEIRO, M.M., S. MARIETTE, G.G. VENDRAMIN, A.E. SZMIDT, C. PLOMION, AND A. KREMER. 2002. Comparison of genetic diversity estimates within and among populations of maritime pine using chloroplast simple-sequence repeat and amplified fragment length polymorphism data. *Molecular Ecology* 11: 869-877.

RITLAND, K. AND S.K. JAIN. 1981. A model for the estimation of outcrossing rate and gene frequencies using n independent loci. *Heredity* 47: 35-5A (Source of programs: http://www.genetics.forestry.ubc.ca/ritland/programs.html.)

RITLAND, K. 2002. Extensions of models for the estimation of mating systems using *n* independent loci. *Heredity*. 88: 221-228.

SAGA AFLP Quantar Software. 2003. Available online: <http://www.keygene-products.com/html/index_products.htm>

SCHMIDT, K., AND K. JENSEN. 2000. Genetic Structure and AFLP Variation of Remnant Populations in the Rare Plant *Pedicularis palustris* (Scrophulariaceae) and its Relation to Population Size and Reproductive Components. 2000. *American Journal of Botany* 87(5): 678-689.

SCHNEIDER, S., D. ROESSLI, AND L. EXCOFFIER. 2000. Arlequin: A software for population genetics data analysis, version 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva.

SHRESTHA, M.K., H. VOLKAERT, AND D. VAN DER STRAETEN. 2005. Assessment of genetic diversity in *Tectona grandis* using amplified fragment length polymorphism markers. *Canandian Journal of Forest Research* 35: 1017-1022.

SLATKIN, M. 1994. Linkage Disequilibrium in Growing and Stable Populations. *Genetics* 137: 331-336.

SWOFFORD, D.L. 2002. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts, USA.

TAKAMI, Y., C. KOSHIO, M. ISHII, H. FUJII, T. HIDAKA, AND I. SHIMIZU. 2004. Genetic diversity and structure of urban populations of *Pieris* butterflies assessed using amplified fragment length polymorphism. *Molecular Ecology* 13: 245-258.

TERO, N., J. ASPI, P. SIIKAMÄKI, A. JÄKÄLÄNIEMI, AND J. TUOMI. 2003. Genetic structure and gene flow in a metapopulation of an endangered plant species, *Silene tatarica*. *Molecular Ecology*, 12:2073-2085.

U.S. FISH AND WILDLIFE SERVICE. 1995. Smooth Coneflower Recovery Plan. Atlanta, GA. 31pp.

U.S. FISH AND WILDLIFE SERVICE. 2000. Policy Regarding Controlled Propagation of Species Listed Under the Endangered Species Act. 65 Fed. Reg., 56916 (September 20, 2000).

VEKEMANS, X. 2002. AFLP-SURV version 1.0. Distributed by the author. Laboratoire de Génétique et Ecologie Végétale, Université Libre de Bruxelles, Belgium.

VEKEMANS X., T. BEAUWENS, M. LEMAIRE AND I. ROLDAN-RUIZ. 2002. Data from amplified fragment length polymorphism (AFLP) markers show indication of size homoplasy and of a relationship between degree of homoplasy and fragment size. *Molecular Ecology* 11: 139-151.

VOS, P., R. HOGERS, M.BLEEKER, M. REIJANS, T.V. DE LEE, M. HORNES, A. FRITJERS, J. POT, J. PELEMAN, M. KUIPER, AND M. ZABEAU. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Research*. 23:4407-4414.

WEIR, B.S. 1990. Genetic Data Analysis. Sinauer Associates, Sunderland, MA.

WRIGHT, S. 1969. *Evolution and the Genetics of Populations*, vol. 2, *The Theory of Gene Frequencies*. University of Chicago Press, Chicago, IL.

YEH, F.C., R.C. YANG, AND T. BOYLE. 1999. POPGENE, version 1.31. Microsoft window-based freeware for population genetic analysis. Quick User-Guide (Francis Yeh, University of Alberta, Canada). <www.ualberta.ca/~fyeh>

ZAWKO, G., S.L. DRAUSS, K.W. DIXON, AND K. SIVASITHAMPARAM. 2001. Conservation genetics of the rare and endangered *Leucopogon obtectus* (Ericaceae). *Molecular Ecology*, 10:2389-2396.

Table 1. Description of populations of *Echinacea laevigata*. N is the number of individuals collected per population and P % is the percentage of polymorphic loci.

Population	**Abbr.**	**Origin**	**N**	**P %**
Picture Creek Diabase Barren	PCPL	North Carolina	19	44.76
Picture Creek Woodlands	PCWL	North Carolina	20	42.38
Snow Hill Road	SH	North Carolina	23	52.38
Knap of Reeds Creek Diabase	KOR	North Carolina	20	47.14
Northside Diabase	NS	North Carolina	20	47.14
Shuffletown Prairie	SHP	North Carolina	20	39.05
Harrington Road	HR	Virginia	20	56.67
Johnson's Creek Natural Area Preserve (NAP)	JC	Virginia	20	55.24
Den Creek Woodland Preserve	DENC	Virginia	19	42.38
Pedlar Hills Natural Area Preserve	PH	Virginia	20	70.00
Grassy Hill Natural Area Preserve	GH	Virginia	19	52.38
Difficult Creek Natural Area Preserve	DC	Virginia	20	65.71
Currahee Mountain	GA-006	Georgia	20	57.62
Georgia Power ROW/ GA HWY 184	GA-026	Georgia	13	47.62
Toccoa Creek Glades	GA-022	Georgia	20	50.48
Yellowback Mountain	GA-001	Georgia	20	52.86
Habersham County (Lon Lyons Rd)	HC	Georgia	17	42.86
Rich Mountain Road	RMR	South Carolina	20	46.19
Cedar Creek Site	CCS	South Carolina	20	69.05
US RT 76	RT 76-001	South Carolina	10	50.48
US RT 76	RT 76-002	South Carolina	20	44.29
Pine Mountain Site	PMS	South Carolina	20	39.52
Total			**420**	**95.24**

Table 2. Populations sampled for both mother plants and their progeny. '**A'** represents the number of mother plants from which seed heads were collected, '**B'** represents the number of progeny harvested, and '**C'** represents the number of mother plants from which progeny were sampled. The difference between '**A'** and '**C'** was caused by failure of seed germination from some of the heads.

Population (States)	**A**	**B**	**C**
Picture Creek Diabase Barren(NC)	30	119	10
Picture Creek Woodlands (NC)	20	49	7
Snow Hill Road (NC)	12	92	11
Grassy Hill Natural Area Preserve (VA)	10	24	7
Den Creek Woodland Preserve (VA)	10	74	8
Difficult Creek Natural Area Preserve (VA)	10	73	9
Total	**92**	**431**	**52**

Table 3. Oligonucleotide adapters and primers used for AFLP analysis of *Echinacea laevigata*.

Adapters	**Sequence**
*Eco*RI adapters	5'-CTCGTAGACTGCGTACC-3'
	3'CATCTGACGCATGGTTAA-5'
*Mse*I adapters	5'-GACGATGAGTCCTGAG-3'
	3'-TACTCAGGACTCAT-5'
Pre-selective Primers	
*Eco*RI +1	GACTGCGTACCAATTC**A**
*Mse*I +1	GATGAGTCCTGAGTAA**C**
Selective Primers	
*Eco*RI +3	GACTGCGTACCAATTC**AGG**
	GACTGCGTACCAATTC**ACC**
*Mse*I +3	GATGAGTCCTGAGTAA**CAG**
	GATGAGTCCTGAGTAA**CTG**
	GATGAGTCCTGAGTAA**CTT**
Selective primer combinations	
1	E-AGG/M-CAG
2	E-ACC/M-CAG
3	E-AGG-M-CTG
4	E-ACC/M-CTT

Table 4. Nei's Gene diversity (**h**) (1978) estimates for all populations and overall for *E. laevigata.*

Population	**Abbreviation**	**h**
Picture Creek Diabase Barren	PCPL	0.1722
Picture Creek Woodlands	PCWL	0.1612
Snow Hill Road	SH	0.1907
Knap of Reeds Creek Diabase	KOR	0.1860
Northside Diabase	NS	0.1690
Shuffletown Prairie	SHP	0.1398
Harrington Road	HR	0.2166
Johnson's Creek Natural Area Preserve (NAP)	JC	0.2183
Den Creek Woodland Preserve	DENC	0.1474
Pedlar Hills Natural Area Preserve	PH	0.2606
Grassy Hill Natural Area Preserve	GH	0.2059
Difficult Creek Natural Area Preserve	DC	0.2532
Currahee Mountain	GA 006	0.1963
Georgia Power ROW/GA HWY 184	GA 026	0.1800
Toccoa Creek Glades	GA 022	0.1885
Yellowback Mountain	GA 001	0.1866
Habersham County (Lon Lyons Rd)	GA HC	0.1644
Rich Mountain Road	RMR	0.1629
Cedar Creek Site	CCS	0.2371
US RT 76	RT 76-001	0.1801
US RT 76	RT 76-002	0.1547
Pine Mountain Site	PMS	0.1464
Total		**0.2611**

Table 5. Pairwise genetic distance matrix based on genetic differentiation (Φ_{ST}) (upper diagonal) and geographic distance (km, lower diagonal) for 22 populations of *E. laevigata*. (All genetic distances are significantly different from zero ($p < 0.05$) except the two bolded and marked *)

	PCPL	PCWL	SH	KOR	NS	SHP	HR	JC	DENC	PH	GH	DC	GA-006	GA-026	GA-022	GA-001	HC	RMR	CCS	RT 76-001	RT 76-002	PMS
PCPL	0	0.058	0.331	0.285	0.267	0.296	0.162	0.213	0.377	0.209	0.357	0.214	0.328	0.26	0.281	0.294	0.309	0.398	0.486	0.477	0.312	0.324
PCWL	0.348	0	0.325	0.319	0.296	0.34	0.155	0.212	0.376	0.215	0.337	0.228	0.321	0.261	0.285	0.299	0.326	0.432	0.504	0.486	0.335	0.353
SH	15.655	15.354	0	0.357	0.408	0.415	0.313	0.29	0.404	0.296	0.355	0.297	0.357	0.239	0.319	0.357	0.305	0.493	0.521	0.499	0.425	0.433
KOR	6.465	6.14896	9.24374	0	0.211	0.206	0.279	0.286	0.234	0.085	0.201	0.168	0.3	0.165	0.232	0.275	0.282	0.437	0.505	0.514	0.346	0.342
NS	10.483	10.3483	9.53013	7.14641	0	0.274	0.256	0.253	0.306	0.124	0.343	0.199	0.311	0.223	0.23	0.3	0.288	0.418	0.486	0.496	0.372	0.332
SHP	221.962	221.623	207.021	215.622	215.841	0	0.292	0.329	0.347	0.169	0.338	0.188	0.322	0.21	0.206	0.273	0.255	0.344	0.455	0.468	0.323	0.251
HR	215.961	215.846	219.4	216.621	223.752	292.827	0	0.107	0.287	0.165	0.332	0.199	0.254	0.237	0.201	0.225	0.302	0.326	0.402	0.384	0.326	0.28
JC	220.784	220.657	223.654	221.207	228.315	289.94	9.34031	0	0.296	0.199	0.316	0.203	0.116	0.187	0.172	0.119	0.231	0.303	0.332	0.26	0.252	0.305
DENC	182.964	182.737	180.686	181.238	187.733	212.531	81.079	77.4929	0	0.108	0.265	0.204	0.318	0.215	0.238	0.296	0.343	0.429	0.492	0.49	0.366	0.363
PH	182.916	182.689	180.635	181.189	187.684	212.482	81.1222	77.5401	0.058	0	0.168	0.069	0.218	0.107	0.17	0.214	0.215	0.296	0.376	0.348	0.239	0.197
GH	141.137	140.926	139.988	139.82	146.51	208.715	92.7447	93.5193	42.908	42.868	0	0.166	0.321	0.217	0.239	0.288	0.245	0.451	0.516	0.529	0.385	0.336
DC	65.858	65.9289	77.1366	69.8446	76.1419	256.644	162.981	169.374	151.963	151.93	109.753	0	0.2	0.162	0.176	0.198	0.185	0.287	0.339	0.319	0.271	0.234
GA-006	460.348	460.006	445.6	454.068	454.52	238.799	478.368	472.163	401.237	401.209	416.584	489.13	0	0.135	0.157	**0.028***	0.208	0.309	0.323	0.274	0.225	0.298
GA-026	460.015	459.673	445.223	453.722	454.113	238.322	479.798	473.646	402.461	402.433	417.441	489.281	3.889	0	0.089	0.113	0.109	0.341	0.404	0.369	0.205	0.242
GA-022	452.596	452.253	437.907	446.332	446.865	231.288	469.192	462.973	392.133	392.106	407.673	480.809	9.228	11.545	0	0.081	0.122	0.262	0.35	0.3	0.248	0.214
GA-001	448.726	448.383	434.077	442.474	443.059	227.603	464.304	458.07	387.316	387.289	403.015	476.566	14.210	16.474	5.000	0	0.17	0.27	0.316	0.257	0.218	0.264
HC	461.1	460.755	446.6	454.893	455.662	240.78	469.966	463.494	393.98	393.955	411.208	487.33	20.233	24.113	17.278	17.085	0	0.327	0.435	0.42	0.275	0.262
RMR	427.828	427.484	413.296	421.61	422.343	207.402	441.756	435.557	364.681	364.654	380.428	454.697	36.613	38.187	27.458	22.646	33.378	0	0.171	0.11	0.281	0.214
CCS	433.533	433.188	418.972	427.307	428.004	212.917	447.737	441.523	370.71	370.682	386.483	460.623	30.642	32.351	21.455	16.607	27.979	6.055	0	**0.004***	0.414	0.388
RT 76-001	438.7	438.356	424.128	432.471	433.152	217.999	452.707	446.467	375.773	375.746	391.66	465.861	25.787	27.762	16.560	11.609	23.118	11.256	5.241	0	0.383	0.379
RT 76-002	438.674	438.33	424.103	432.445	433.129	217.981	452.645	446.403	375.715	375.688	391.609	465.825	25.856	27.838	16.628	11.675	23.121	11.209	5.203	0.089	0	0.169
PMS	445.942	445.598	431.355	439.708	440.371	225.141	459.625	453.346	382.826	382.8	398.879	473.184	19.410	21.928	10.384	5.545	16.647	18.554	12.565	7.324	7.363	0

Table 6. Results of the three hierarchical analyses of molecular variance (AMOVA).

Source of Variation	Variance components	% of total variance	P-value	Φ Statistics
Populations Pooled				
Among populations	2.68486	29.74	<0.0001	
Within populations	6.34146	70.26	<0.0001	$\Phi_{ST} = 0.297$
Total	9.02631			
Among states	1.64408	7.94	<0.0001	$\Phi_{CT} = 0.079$
Among populations within states	4.46693	21.56	<0.0001	$\Phi_{SC} = 0.234$
Within populations	14.60825	70.51	<0.0001	$\Phi_{ST} = 0.295$
Total	20.71925			
Populations Separated by Region				
M + P + GA/SC[1]				
Among regions	0.66186	7.1	<0.0001	$\Phi_{CT} = 0.072$
Among populations within regions	2.23544	24.20	<0.0001	$\Phi_{SC} = 0.261$
Within populations	6.34146	68.64	<0.0001	$\Phi_{ST} = 0.314$
Total	9.23876			

[1] M represents the Mountain region, P the Piedmont region, and GA/SC the GA and SC region of populations.

Table 7. Estimates of outcrossing rates for the three populations each from Virginia and North Carolina. (standard deviation)

Populations	**Outcrossing rate** t_m
VA Pops	
DENC	0.844 (0.548)
GH	0.833 (0.553)
DC	0.876 (0.532)
NC Pops	
PCPL	1.2 (0.545)
PCWL	1.2 (0.548)
SH	1.2 (0.550)

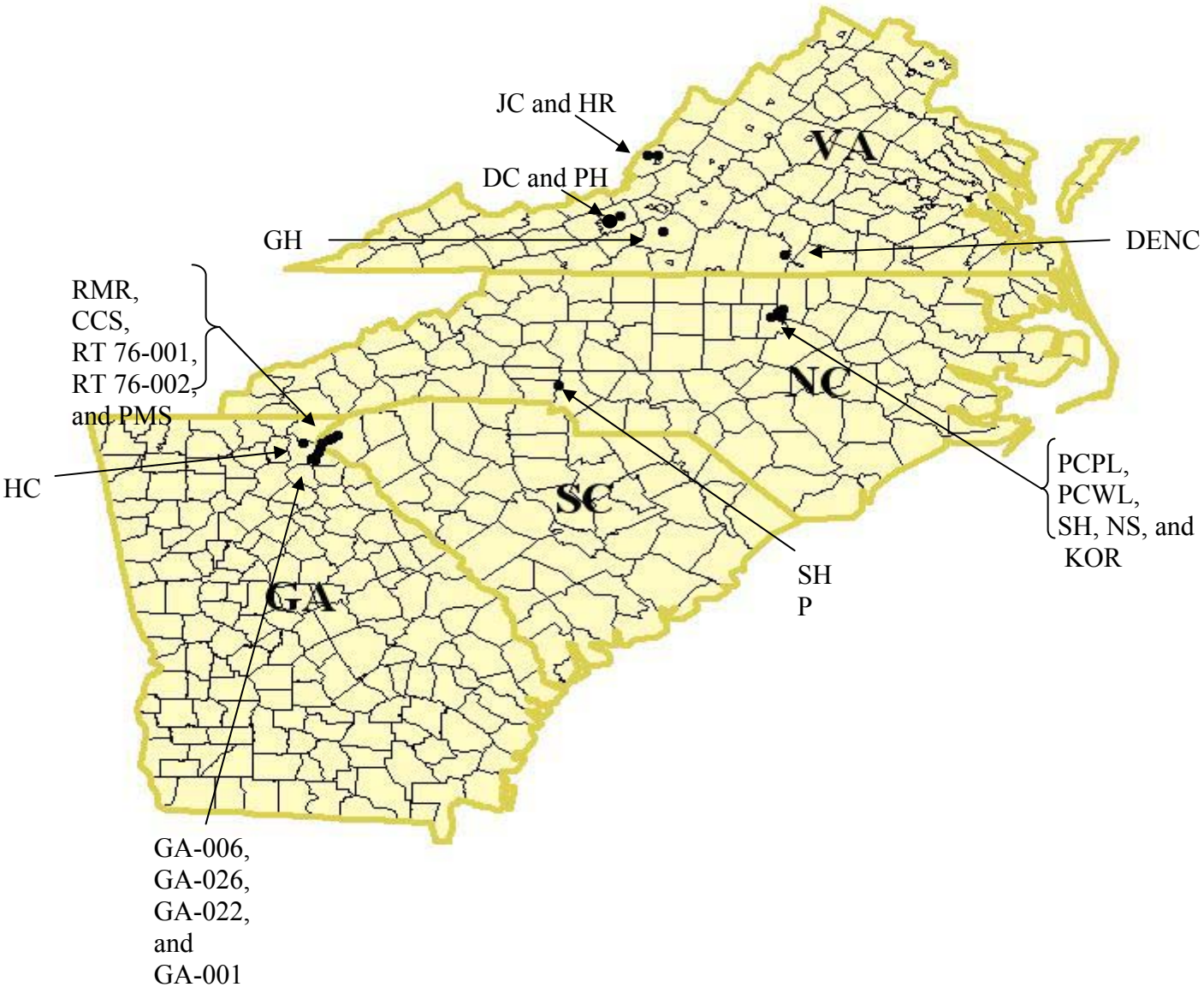

Figure 1. Geographic locations of the 22 *E. laevigata* populations included in this study. (Full population names can be found on Table 1.)

Figure 2. Photo of *E. laevigata* from South Carolina population (RT 76-001).

Figure 3. Picture of seedlings grown from seed in the NCSU Phytotron.

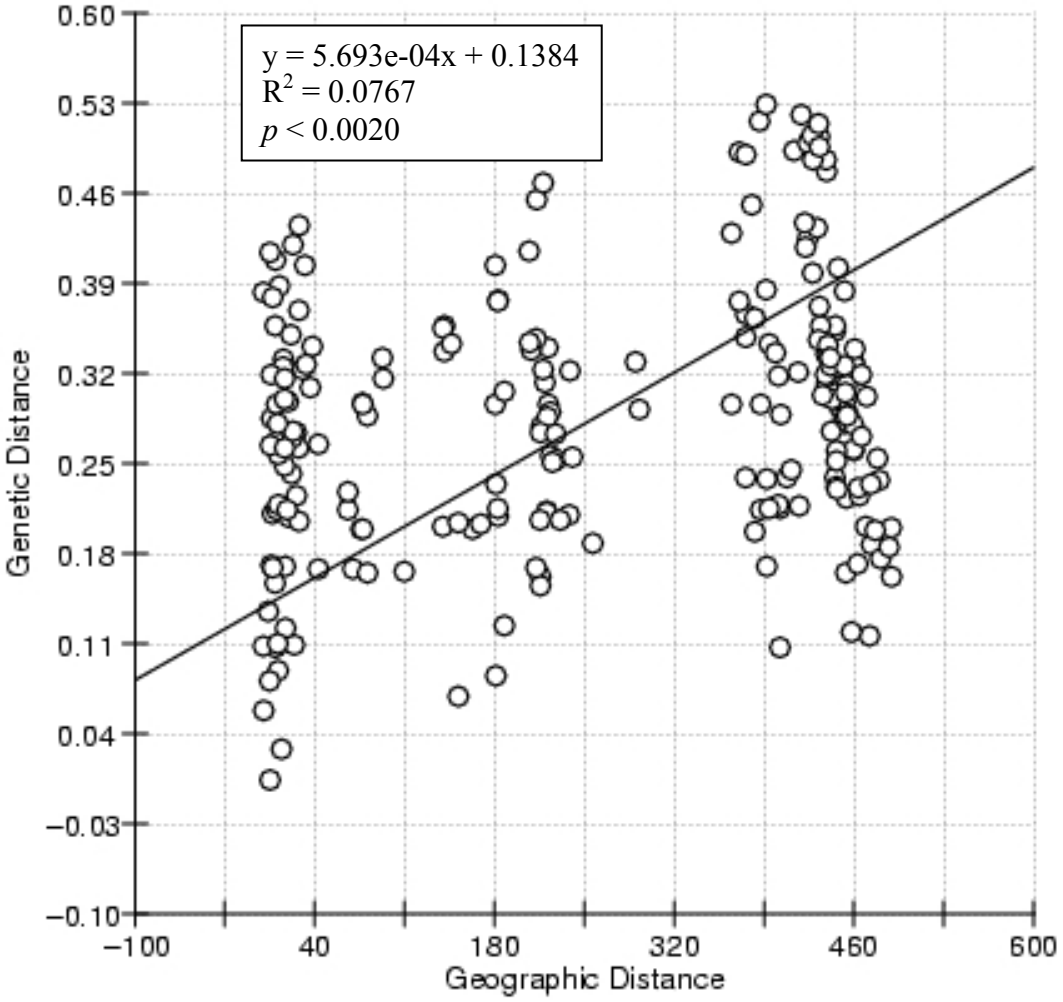

Figure 4. Relationship between genetic distance (Φ_{ST}) and geographic distance (Km) among 22 populations of *E. laevigata*.

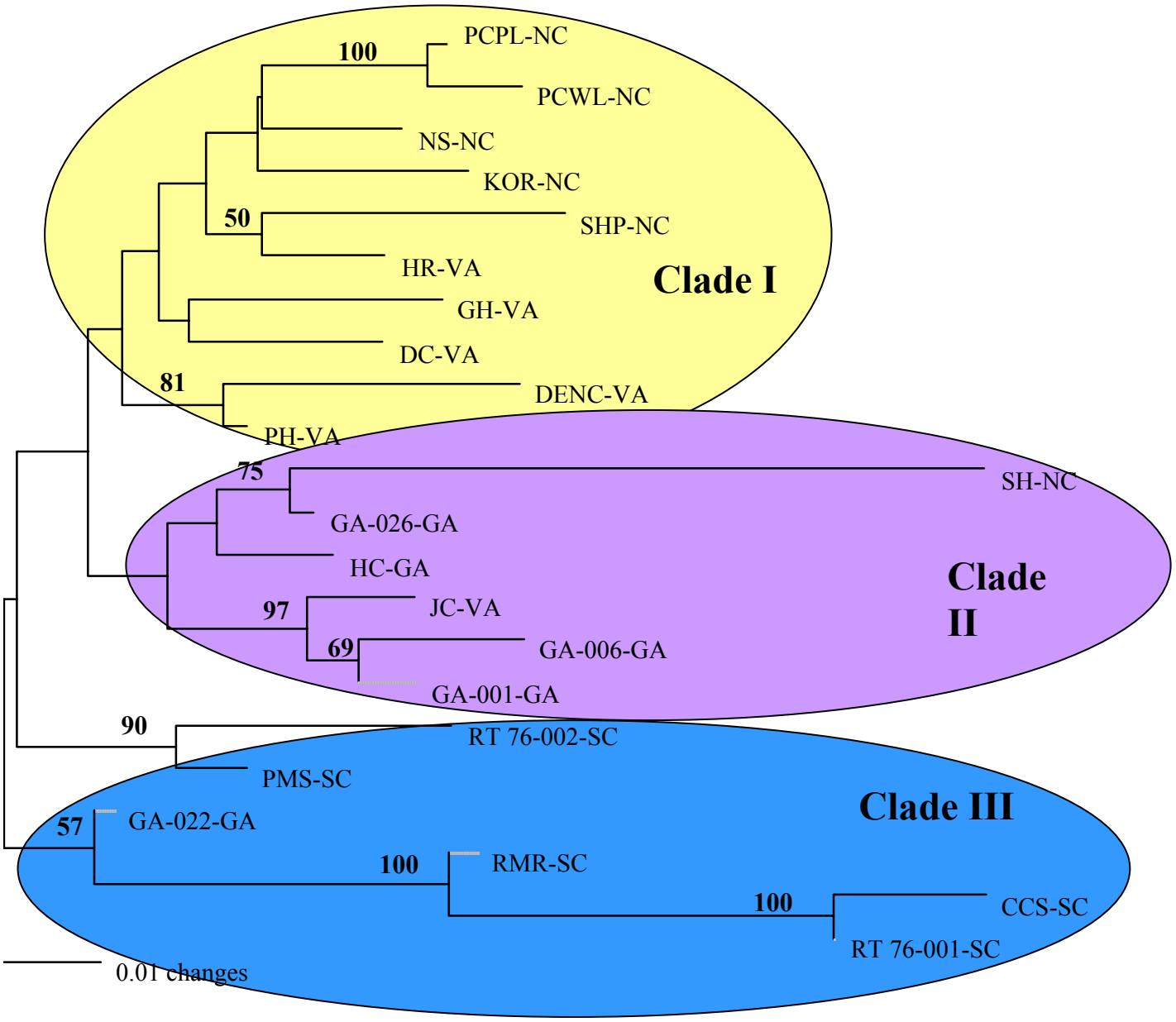

Figure 5. Neighbor-joining phylogram based on Nei's genetic distance (Lynch and Milligan, 1994) data with bootstrap values equal to or higher than 50% shown (Populations – State Abbreviations).

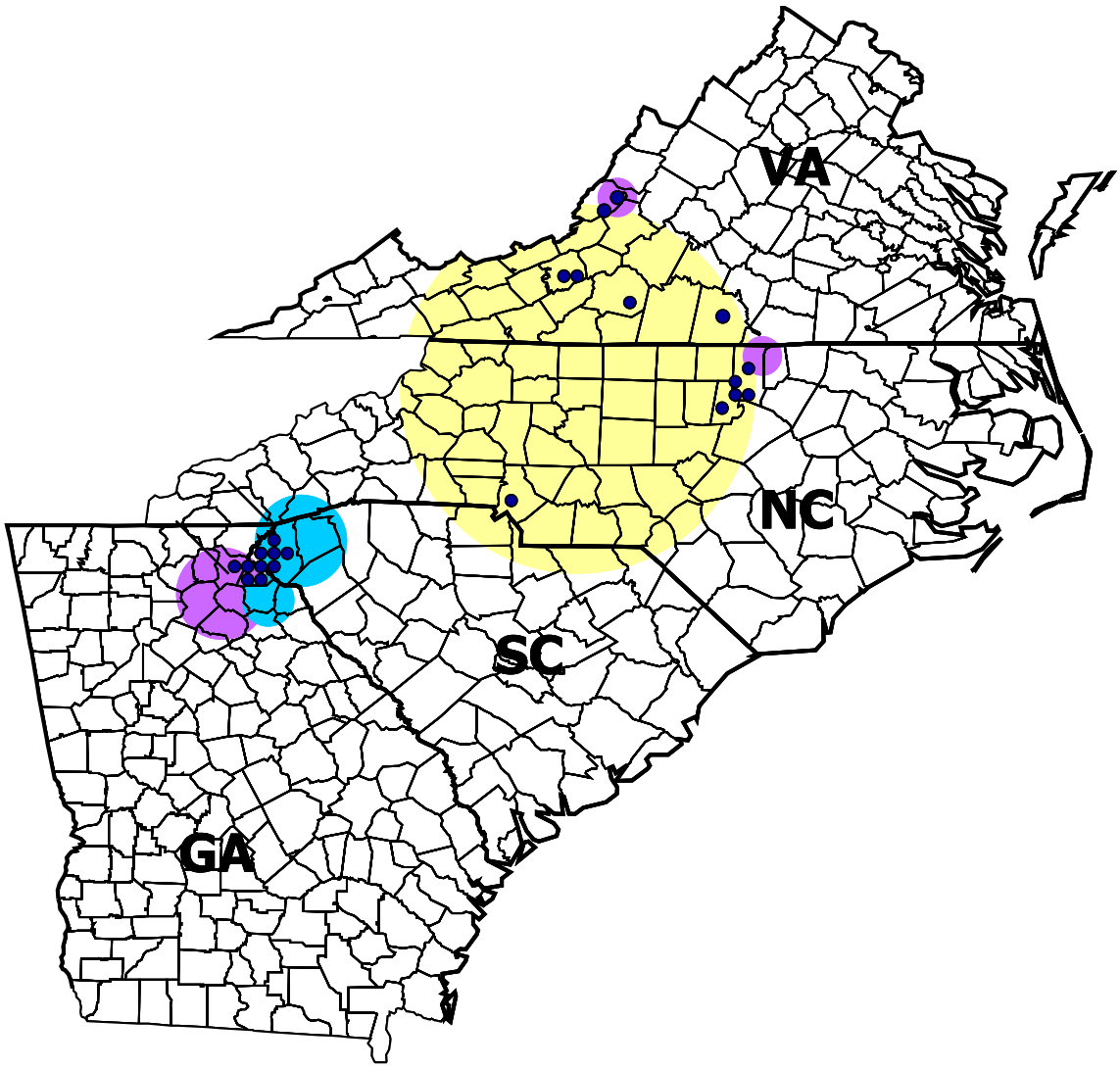

Figure 6. Map of *E. laevigata* populations with clades from neighbor-joining phylogram highlighted. (Yellow grouping represents clade I, purple grouping represents clade II, and blue grouping represents clade III; dark blue dots represent the 22 populations included in this study).