Conserv Genet (2009) 10:1–14
DOI 10.1007/s10592-008-9509-1

RESEARCH ARTICLE

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 · Qiu-Yun (Jenny) Xiang · David T. Thomas · Jon Stucky · Noah K. Whiteman

Received: 31 January 2007 / Accepted: 4 January 2008 / Published online: 22 January 2008
© Springer Science+Business Media B.V. 2008

Abstract *Echinacea laevigata* (Boynton and Beadle) Blake is a federally endangered flowering plant species restricted to four states in the southeastern United States. 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 ($G_{ST} = 0.294$; $\Theta^{II} = 0.218$ from the Bayesian $f = 0$ model). Results from the AMOVA analysis suggest that a majority of the genetic variance is attributed to variation within populations (70.26%), which is also evident from the PCoA. However, 82% of individuals were assigned back to the original population based on the results of the assignment test. An isolation by distance analysis indicated that genetic differentiation among populations was a function of geographic distance, although long-distance gene dispersal between some populations was suggested from an analysis of relatedness between populations using the neighbor-joining method. An estimate of the outcrossing rate based on genotypes of progenies from six of the 22 populations using the multilocus method from the program MLTR ranged from 0.780 to 0.912, suggesting that the species is predominantly outcrossing. These results are encouraging for conservation, signifying that populations may persist due to continued genetic exchange sustained by the outcrossing mating system of the species.

Keywords *Echinacea laevigata* · Genetic structure · Outcrossing rate · AFLP

M. D. Peters · Qiu-Yun (Jenny)Xiang · D. T. Thomas · J. Stucky
Department of Plant Biology, North Carolina State University, 2115 Gardner Hall, Campus Box 7612, Raleigh, NC 27695-7612, USA

N. K. Whiteman
Museum of Comparative Zoology and Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

M. D. Peters (✉)
Harvard University Herbaria, 22 Divinity Ave, Cambridge, MA 02138, USA
e-mail: mpeters@oeb.harvard.edu

Introduction

The conservation of *Echinacea laevigata* (Boynton and Beadle) Blake (Asteraceae), smooth coneflower, began to take form when it was listed by the United States Fish and Wildlife service as federally endangered in 1992; the attention continued a few years later when managers realized the populations were in decline and set forth with a recovery plan (USFWS 1995). This plan provided a summary of the species and outlined reasonable actions that would allow this species to be "recovered and/or protected" (USFWS 1995). Although current management efforts are underway in all states and populations in federal and state lands are monitored, the self-sustainability of the known populations under the recovery/protection plan remains unclear. Key information on genetic structure and mating patterns of the species is lacking, preventing critical evaluation of the current management plan and prediction of the future status of the species.

Echinacea laevigata is described as a diploid ($2n = 22$) 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 flowering heads are usually solitary with purplish to light pink ray flowers (McGregor 1968). The species flowers from May through July; the fruits or achenes, which have a pappus, develop in late June and mature in September (Gaddy 1991). McGregor (1968) reported that all *Echinacea* species are self-incompatible in a taxonomic revision of the genus, based on his study including experiments on 500 bagged heads that did not produce any seeds. These data suggest that the species may be strictly outcrossing. However, the failure of producing seeds from the bagged heads could be due to ineffective self-pollination (or cross-pollination within heads). Thus, a recent study was conducted to describe flower phenology, compatibility, and the pollination biology of *E. laevigata* (Gadd 2006). This study involved bagging individual flowering heads and hand crosses among flowers of the bagged heads. Results of the study support that *E. laevigata* is self-incompatible and requires pollinators for effective pollination (Gadd 2006). The most effective pollinators reported by this study (Gadd 2006) are bees and butterflies (Order Hymenoptera and Lepidoptera, respectively). Given that the study of Gadd (2006) was performed on only 30 flowering heads from the largest population (PCPL), the mating system of the entire species remains uncharacterized.

Historically, *E. laevigata* had a range over 26 counties in eight states (Pennsylvania, Maryland, Virginia, North Carolina, South Carolina, Georgia, Alabama, and Arkansas) with 62 populations (USFWS 1995); currently, there are 24 reported extant populations in the southeastern United States of Virginia, North Carolina, South Carolina, and Georgia (USFWS 1995). Over time the decline of these populations has been attributed to 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). Monitoring efforts have been underway in some states since the 1980s. The census data reported here are based on a combination of counts from the states with the remaining populations. The data are from various organizations, managers and from personal observation at the field sites.

Habitat type also seems to play a role in the continued existence of this species. It grows in 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). Table 1 reports the habitat characteristics for each population in more detail, and gives approximate numbers of rosettes based on summaries from working groups in each state between the years 2000 and 2004, plus personal observations.

An earlier study by Apsit and Dixon (2001), examined 28 allozyme loci for 11 populations from Virginia, North Carolina, and South Carolina to evaluate the genetic diversity and population genetic structure of *E. laevigata*. The study found moderate levels of genetic diversity ($H_e = 0.178$), significant population structure ($G_{ST} = 0.109$), and significant, positive correlation between pairwise genetic and geographic distances among the 11 populations ($r = 0.38$; $P \leq 0.025$). This correlation suggests that gene flow in the species is restricted by geographic distance between populations. The hierarchical analysis of molecular variance (AMOVA) indicated that significant genetic variation was partitioned among states, among populations within states, and among populations ($\Phi_{CT} = 0.105$; $\Phi_{SC} = 0.127$; $\Phi_{ST} = 0.219$, respectively). Based on these results, Apsit and Dixon (2001) suggested that *E. laevigata* might be adapting to local environmental heterogeneity. They further indicated that analysis of mating system is needed to assess levels of inbreeding and outcrossing, which determines the future genetic diversity of the species.

There were two major goals of our study: (1) to conduct a thorough evaluation of the genetic diversity and population structure of *E. laevigata* by examining nearly all of the populations using amplified fragment length polymorphism (AFLP) markers, and (2) to estimate the outcrossing in present populations using AFLP genotyping of progeny grown from known mother plants. This study will provide data for more informed conservation management of the species.

Materials and methods

Plant material

Leaf material was collected from 22 of the 24 populations for genotyping using AFLP analysis. The two populations that are not included here are because the presence of the populations was uncertain at the time of collecting. A total of 420 individuals were chosen for the study randomly across populations. The number of individuals sampled per population ranged from 10 to 23 depending on population size (Tables 1, 2).

AFLP markers

AFLP genotyping have been widely used for population genetic analyses and have provided valuable insight into

Table 1 Description of *E. laevigata* populations and the habitat association for each population

Population	Code	Origin[a]	N^b	Habitat type
Picture Creek Diabase Barren	PCPL	NC, P	50,000	Power line right-of-way (PLROW); diabase, slightly basic soil
Picture Creek Woodlands	PCWL	NC, P	150	Oak-hickory forest; in adjacent woodlands to PCPL
Snow Hill Road	SH	NC, P	30	Roadside; adjacent to oak forest
Knap of Reeds Creek Diabase	KOR	NC, P	25	PLROW; diabase
Northside Diabase	NS	NC, P	20	Along an abandoned rail line bank in ditch; diabase
Shuffletown Prairie	SHP	NC, P	30	PLROW
Den Creek Woodland Preserve	DENC	VA, P	130	Ridge/valley; dolomite barrens (glades) and montane dry calcareous woodlands
Harrington Road	HR	VA, M	100	Roadside; calcareous shale and acidic woodlands
Johnson's Creek Natural Area Preserve	JC	VA, M	40	Forest; calcareous shale and acidic woodlands
Pedlar Hills Natural Area Preserve	PH	VA, M	570	Ridge/valley; dolomite barrens (glades) montane dry calcareous woodlands
Grassy Hill Natural Area Preserve	GH	VA, M	2000	PLROW; montane basic woodland and mafic rock
Difficult Creek Natural Area Preserve	DC	VA, M	350	PLROW; mafic rock and hardpan woodlands
Currahee Mountain	GA-006	GA, GA/SC	500	Forest service roadside; dry-mesic soil
Georgia Power ROW/GA HWY 184	GA-026	GA, GA/SC	15	PLROW; dry-mesic soil
Toccoa Creek Glades	GA-022	GA, GA/SC	500	Forest; xeric to dry-mesic soil
Yellowback Mountain	GA-001	GA, GA/SC	300	Roadside in forest; dry-mesic soil
Habersham County (Lon Lyons Rd)	HC	GA, GA/SC	200	PLROW; dry-mesic soil
Rich Mountain Road	RMR	SC, GA/SC	350	Roadside; high in Ca/ Mg, basic dry-mesic oak hickory associates
Cedar Creek Site	CCS	SC, GA/SC	30	Roadside; high in Ca/ Mg
US RT 76	RT 76-001	SC, GA/SC	25	PLROW; soil rich in Ca/ Mg, forests and woodlands
US RT 76	RT 76-002	SC, GA/SC	50	PLROW; soil rich in Ca/ Mg
Pine Mountain Site	PMS	SC, GA/SC	650	Forest; soil rich in Ca/ Mg

Origin[a] (State, Region)—NC, North Carolina; VA, Virginia; GA, Georgia; SC, South Carolina; P, Piedmont region; M, Mountain region; GA/SC, GA and SC region

N^b—The approximate number of rosettes in a given population based on census data from 2000 to 2004

rare species management (Palacios et al. 1999; Schmidt and Jensen 2000; Gaudeul et al. 2000; Zawko et al. 2001; Tero et al. 2003; Kim et al. 2005). The analysis of AFLP loci generates multi-locus phenotypic data for comparison across individuals and populations. This permits the survey of numerous DNA loci across the entire nuclear genome by using multiple PCR primers to allow the detection of genetic variation between closely related individuals (Vos et al. 1995). Although this method has many advantages, it is important to note that like any genotyping method precautions must be taken to minimize errors that can come from various sources. Major errors in AFLP genotyping occur from poor quality or quantity of DNA, contamination of DNA samples, laboratory artefacts and human error. However, these errors have not been found to affect the conclusions of other studies (Bonin et al. 2004; Pompanon et al. 2005). We took caution in the analyses to minimize these errors with the resources available (see below).

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). Healthy and clean leaf tissues were used for DNA extraction to avoid contamination of DNA from fungi or bacteria. A GeneQuant spectrophotometer (Amersham Pharmacia Biotech, Cambridge, England) was used to estimate the quality and quantity of DNA, and each sample was electrophoresed on a 1.0% agarose gel to compare band densities across individuals. Samples with 260/280 UV absorbance ratios of 1.8–1.9 and had consistent band densities were used for AFLP analysis.

AFLP reactions were performed as described by Vos et al. (1995) with the following modifications. The restriction digest and ligation steps were done as separate

Table 2 *N* is the number of *E. laevigata* individuals collected per population for this study *P %* is the percentage of polymorphic loci; and (*h*) is Nei's gene diversity (1978) estimates for all populations and overall for *E. laevigata*

Population	Code	*N*	*P %*	*h*
Picture Creek Diabase Barren	PCPL	19	44.76	0.1722
Picture Creek Woodlands	PCWL	20	42.38	0.1612
Snow Hill Road	SH	23	52.38	0.1907
Knap of Reeds Creek Diabase	KOR	20	47.14	0.1860
Northside Diabase	NS	20	47.14	0.1690
Shuffletown Prairie	SHP	20	39.05	0.1398
Den Creek Woodland Preserve	DENC	19	42.38	0.1474
Harrington Road	HR	20	56.67	0.2166
Johnson's Creek Natural Area Preserve	JC	20	55.24	0.2183
Pedlar Hills Natural Area Preserve	PH	20	70.00	0.2606
Grassy Hill Natural Area Preserve	GH	19	52.38	0.2059
Difficult Creek Natural Area Preserve	DC	20	65.71	0.2532
Currahee Mountain	GA-006	20	57.62	0.1963
Georgia Power ROW/GA HWY 184	GA-026	13	47.62	0.1800
Toccoa Creek Glades	GA-022	20	50.48	0.1885
Yellowback Mountain	GA-001	20	52.86	0.1866
Habersham County (Lon Lyons Rd)	HC	17	42.86	0.1644
Rich Mountain Road	RMR	20	46.19	0.1629
Cedar Creek Site	CCS	20	69.05	0.2371
US RT 76	RT 76-001	10	50.48	0.1801
US RT 76	RT 76-002	20	44.29	0.1547
Pine Mountain Site	PMS	20	39.52	0.1464
Total for all individuals		420	95.24	0.2611

reactions. For the digestion, approximately 500 ng of genomic DNA was incubated at 37°C for 3 h in a 10 µl volume reaction containing 1× 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 10× 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 h. After ligation the samples were diluted 10-fold with sterile deionized water (dH_2O). A pre-selective polymerase chain reaction (PCR), using a Robocycler thermalcycler (Stratagene, LaJolla, California, USA) was done using primer pairs with a single selective nucleotide extension. 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 10× Buffer B (Promega, Madison, Wisconsin, USA), 200 µM of each dNTP, 1.5 mM $MgCl_2$, 1 µg/ml BSA, 8.3 µM *Eco*RI primer, and 8.3 µM *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 performed with a final extension at 60°C for 10 min. Samples were diluted 10-fold with dH_2O and stored at −20°C.

Preliminary tests were performed to identify the most variable selective primer extensions, which included 12 primer combinations. Four primer combinations were selected (E-AGG/M-CAG, E-ACC/M-CAG, E-AGG/M-CTG, and E-ACC/M-CTT). Selective PCRs (total volume of 10 µl) included 3 µl of template from the pre-selective step, 4 U *Taq* DNA Polymerase, with 10× Buffer B, 200 µM of each dNTP, 1.5 mM $MgCl_2$, 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 bromophenol 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, 1,500 V) on a LI-COR 4300L automated DNA sequencer. Twelve of the same samples were included in all gels as a control in order to provide a check for genotyping error from electrophoresis. Gels were scored using AFLP Quantar (Keygene 2003), software designed for analysis of AFLP gels from

the LI-COR sequencer, which minimizes human error in scoring. 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 "?". Bands that were not consistent among gels in the 12 control samples were not scored for all samples. Overall, only fragments between 63 and 564 bp were scored.

Gene diversity, gene flow, and population structure

Linkage disequilibrium was calculated for each population using the output matrix in ARLEQUIN version 2.0 (Schneider et al. 2000) and the percentage of shared pairs across all populations was calculated using the Unix shell scripts and the AWK program written by Jordan Mueller (Mueller unpublished). 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 version 1.31 (Yeh et al. 1999). The percentage of polymorphic loci within and among populations was determined, a G_{ST} estimate was computed to estimate the genetic differentiation among populations, and gene diversity (Nei 1978) was calculated for each population based on unbiased estimates for the different loci. The Hickory software, version 1.1 (Holsinger et al. 2002), which incorporates a Bayesian model that corresponds to the traditional F_{ST}, was used to estimate parameters of genetic variation (Θ^{B}). This Bayesian model, which does not assume that genotypes are in Hardy–Weinberg equilibrium, also has the option of reporting Θ^{II} that is analogous to the G_{ST} value for genetic differentiation. We used the default parameters for burn-in (5,000), sampling (100,000), and thinning (20) in the four models: (i) a full model, (ii) a model which assumes no inbreeding ($f = 0$), (iii) a model that assumes no population structure ($\Theta^{B} = 0$), and (iv) the *f*- free model that allows the incorporation of uncertainty about *f* into the analysis. Since estimates of F_{IS} based on dominant markers must be taken with caution, we used the *f*-free analysis to calculate estimates of Θ^{II}. The models were evaluated based on the measures of deviance information criterion (DIC) to decide which model is preferred (smaller DIC values indicate a better fit) (Baus et al. 2005; Zhang et al. 2007).

To estimate the relationships among populations, Nei's unbiased pairwise genetic distance data (1978) was calculated by the program AFLP-SURV 1.0 (Vekemans 2002) and imported into PAUP* 4.0b10 (Swofford 2002) for tree construction using the neighbor-joining (NJ) method. The AFLP-SURV program was also used to calculate 1,000 bootstrap distance matrices, which were opened with the NEIGHBOR program in PHYLIP (Felsenstein 1993) in order to obtain a tree file and elucidate relationships among populations. The CONSENSE program from the PHYLIP software was used to compute a majority-rule consensus tree.

The geographical information software package ARC-VIEW Version 3.3 was used to plot the localities of the populations based on GPS coordinates from all 22 populations. A Mantel test was used to assess isolation by distance (Bohonak 2002) between geographic distance and Nei's genetic distance. To further examine the partitioning of genetic variation, ARLEQUIN, version 3.0 (Excoffier et al. 2005) 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). The AMOVA was first preformed by partitioning genetic variation among and within populations regardless of their geographic distribution. Populations were also grouped into defined geographic regions for further analyses to determine if genetic differentiation among the Mountain region, the Piedmont region, and the Georgia/South Carolina region exist (Table 6). The AMOVA analyses are based on pairwise squared Euclidean distances (Excoffier et al. 1992) and assume the mating system to be the same in all populations for dominant markers (Tero et al. 2003). To further asses genetic structure, a principal coordinate analysis (PCoA) also based on Euclidean distances between AFLP phenotypes was performed using GEN-ALEX version 6.0 (Peakall and Smouse 2006). Finally, to calculate the probability of which population each individual came from, an assignment test was carried out using the DOH software (Brzustowski 2002).

These analyses were also run by excluding individuals that had missing data for two or more loci, but the results are not presented because no significant difference was found. Only six individuals (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

Material was collected from six of the 22 populations to be used for outcrossing rate estimation (Table 3). Seven to eleven seed heads, each from a different randomly chosen individual (mother plant), were collected from each of these six populations. Fruits were soaked for 24 h in a 1 mM ethephon solution and stratified on wet blotters in Petri dishes at 4°C for 2 weeks (McKeown and Widrlechner, unpublished protocol modified from Sari et al. 1999). After stratification, the fruits were planted in small Styrofoam cups and placed in a misting bed in a

Table 3 Mating system estimates for the populations sampled in North Carolina (NC) and Virginia (VA) with standard deviations (SD)

Population (States)	A[a]	B[b]	t_m^c(SD)	t_s^d(SD)	$t_m - t_s^e$ (SD)
Picture Creek Diabase Barren (NC)	10	119	0.852 (0.544)	0.744 (0.476)	0.108 (0.076)
Picture Creek Woodlands (NC)	7	49	0.864 (0.538)	0.815 (0.507)	0.049 (0.031)
Snow Hill Road (NC)	11	92	0.912 (0.512)	0.830 (0.466)	0.082 (0.05)
Den Creek Woodland Preserve (VA)	8	74	0.852 (0.544)	0.811 (0.518)	0.041 (0.027)
Grassy Hill Natural Area Preserve (VA)	7	24	0.780 (0.572)	0.690 (0.506)	0.090 (0.072)
Difficult Creek Natural Area Preserve (VA)	9	73	0.912 (0.512)	0.792 (0.444)	0.120 (0.07)
Total	52	431			

[a] Number of mother plants from which progeny were sampled

[b] Number of progeny harvested

[c] Multilocus outcrossing rate

[d] Single-locus outcrossing rate

[e] Difference between multilocus and single-locus outcrossing rate

greenhouse for up to 3 weeks. Leaf tissue was collected for DNA extraction and the AFLP genotyping protocol, described above, was followed. To voucher the greenhouse collections, one plant from the six populations was collected and deposited in the North Carolina State University Herbarium (NCSC).

The MLTR (Multilocus Mating System Program) software Version 3.0 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), which ranges from 0 for complete selfing to 1 for complete outcrossing, using the Newton-Raphson (NR) numerical method with AFLP data from 431 progenies of six populations (Table 3) based on the model described by Ritland (2002). We have also included the calculations for the single-locus outcrossing rate (t_s) in order to determine the rate of biparental inbreeding, which is $t_m - t_s$ (Table 3). Since AFLPs are treated as dominant markers, the program encodes a diploid locus by a single allele rather than a pair of alleles (Ritland 2002). Standard errors were estimated based on 100 bootstraps between individuals within progeny arrays.

Results

Gene diversity, gene flow, and population structure

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 2). The percentage of polymorphic loci per population ranged from 39.05 to 70.00% (SHP and PH populations, respectively) as shown in Table 2. The overall gene diversity including all individuals was 0.261. Linkage disequilibrium was examined between all pairs of polymorphic loci across all populations. There were 4,737 unique combinations in the 22 populations. Of these, 1,217 combinations (25.7%) were significant at the 0.05 level.

The overall estimate of genetic structure (G_{ST}) was 0.294. 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 4). All of the genetic distances in the matrix were significantly different from zero at the $P < 0.05$ level except between populations GA-006/GA-001 and populations CCS/RT 76-001. As a comparison, the mean genetic differentiation between populations, based on the Bayesian $f = 0$ model with the smallest deviance information criteria (DIC = 12095.2), of Holsinger (Θ^{II}) was 0.218. Results of the four models used are presented in Table 5.

The neighbor-joining phylogram constructed for all populations based on Nei's unbiased pairwise genetic distance data is shown in Fig. 1. Populations clustered basically by geographic proximity (Fig. 2), with the exception of three populations, Snow Hill (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. A positive and significant (Mantel $r = 0.2769$; $P < 0.002$) correlation between genetic differentiation, Φ_{ST}, and geographic distance (Km), was detected (Fig. 3).

Results of all AMOVA analyses indicated that most of the genetic variance occurred within populations, although there was also significant variance among populations (29.74%) (Table 6). AMOVA analyses among the specified geographic regions showed the majority of genetic variance was within populations and that there was little variance among the three regions (Table 6). The first and second axes of the principal coordinate analysis (plotted against one another in Fig. 4) explained 29.5 and 21% of the genetic similarities among populations, respectively. Individuals within

Table 4 Pairwise genetic distance matrix based on genetic differentiation (Φ_{ST}) (upper diagonal) and geographic distance (km, lower diagonal) for 22 populations of *E. laevigata*

	PCPL	PCWL	SH	KOR	NS	SHP	HR	JC	DENC	PH	GH
PCPL	0	0.058	0.331	0.285	0.267	0.296	0.162	0.213	0.377	0.209	0.357
PCWL	0.348	0	0.325	0.319	0.296	0.34	0.155	0.212	0.376	0.215	0.337
SH	15.655	15.354	0	0.357	0.408	0.415	0.313	0.29	0.404	0.296	0.355
KOR	6.465	6.14896	9.24374	0	0.211	0.206	0.279	0.286	0.234	0.085	0.201
NS	10.483	10.3483	9.53013	7.14641	0	0.274	0.256	0.253	0.306	0.124	0.343
SHP	221.962	221.623	207.021	215.622	215.841	0	0.292	0.329	0.347	0.169	0.338
HR	215.961	215.846	219.4	216.621	223.752	292.827	0	0.107	0.287	0.165	0.332
JC	220.784	220.657	223.654	221.207	228.315	289.94	9.34031	0	0.296	0.199	0.316
DENC	182.964	182.737	180.686	181.238	187.733	212.531	81.079	77.4929	0	0.108	0.265
PH	182.916	182.689	180.635	181.189	187.684	212.482	81.1222	77.5401	0.058	0	0.168
GH	141.137	140.926	139.988	139.82	146.51	208.715	92.7447	93.5193	42.908	42.868	0
DC	65.858	65.9289	77.1366	69.8446	76.1419	256.644	162.981	169.374	151.963	151.93	109.753
GA-006	460.348	460.006	445.6	454.068	454.52	238.799	478.368	472.163	401.237	401.209	416.584
GA-026	460.015	459.673	445.223	453.722	454.113	238.322	479.798	473.646	402.461	402.433	417.441
GA-022	452.596	452.253	437.907	446.332	446.865	231.288	469.192	462.973	392.133	392.106	407.673
GA-001	448.726	448.383	434.077	442.474	443.059	227.603	464.304	458.07	387.316	387.289	403.015
HC	461.1	460.755	446.6	454.893	455.662	240.78	469.966	463.494	393.98	393.955	411.208
RMR	427.828	427.484	413.296	421.61	422.343	207.402	441.756	435.557	364.681	364.654	380.428
CCS	433.533	433.188	418.972	427.307	428.004	212.917	447.737	441.523	370.71	370.682	386.483
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
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
PMS	445.942	445.598	431.355	439.708	440.371	225.141	459.625	453.346	382.826	382.8	398.879

	DC	GA-006	GA-026	GA-022	GA-001	HC	RMR	CCS	RT 76-–001	RT 76-–002	PMS
PCPL	0.214	0.328	0.26	0.281	0.294	0.309	0.398	0.486	0.477	0.312	0.324
PCWL	0.228	0.321	0.261	0.285	0.299	0.326	0.432	0.504	0.486	0.335	0.353
SH	0.297	0.357	0.239	0.319	0.357	0.305	0.493	0.521	0.499	0.425	0.433
KOR	0.168	0.3	0.165	0.232	0.275	0.282	0.437	0.505	0.514	0.346	0.342
NS	0.199	0.311	0.223	0.23	0.3	0.288	0.418	0.486	0.496	0.372	0.332
SHP	0.188	0.322	0.21	0.206	0.273	0.255	0.344	0.455	0.468	0.323	0.251
HR	0.199	0.254	0.237	0.201	0.225	0.302	0.326	0.402	0.384	0.326	0.28
JC	0.203	0.116	0.187	0.172	0.119	0.231	0.303	0.332	0.26	0.252	0.305
DENC	0.204	0.318	0.215	0.238	0.296	0.343	0.429	0.492	0.49	0.366	0.363
PH	0.069	0.218	0.107	0.17	0.214	0.215	0.296	0.376	0.348	0.239	0.197
GH	0.166	0.321	0.217	0.239	0.288	0.245	0.451	0.516	0.529	0.385	0.336
DC	0	0.2	0.162	0.176	0.198	0.185	0.287	0.339	0.319	0.271	0.234
GA-006	489.13	0	0.135	0.157	**0.028***	0.208	0.309	0.323	0.274	0.225	0.298
GA-026	489.281	3.889	0	0.089	0.113	0.109	0.341	0.404	0.369	0.205	0.242
GA-022	480.809	9.228	11.545	0	0.081	0.122	0.262	0.35	0.3	0.248	0.214
GA-001	476.566	14.210	16.474	5.000	0	0.17	0.27	0.316	0.257	0.218	0.264
HC	487.33	20.233	24.113	17.278	17.085	0	0.327	0.435	0.42	0.275	0.262
RMR	454.697	36.613	38.187	27.458	22.646	33.378	0	0.171	0.11	0.281	0.214
CCS	460.623	30.642	32.351	21.455	16.607	27.979	6.055	0	**0.004***	0.414	0.388
RT 76-–001	465.861	25.787	27.762	16.560	11.609	23.118	11.256	5.241	0	0.383	0.379
RT 76-–002	465.825	25.856	27.838	16.628	11.675	23.121	11.209	5.203	0.089	0	0.169
PMS	473.184	19.410	21.928	10.384	5.545	16.647	18.554	12.565	7.324	7.363	0

(All genetic distances are significantly different from zero ($P < 0.05$) except the two bolded and marked *)

Table 5 Results of the mean genetic differentiation based on the Bayesian approach models

Θ^{II}					
Models	Mean	SD[a]	2.5%	97.5%	DIC[b]
Full	0.210	0.006	0.198	0.222	12099.3
$f = 0$	0.207	0.006	0.196	0.218	12095.2
$\Theta^B = 0$	–	–	–	–	24636.3
f = free	0.260	0.009	0.244	0.280	12690.7

[a] SD is standard deviation

[b] DIC is deviance information criterion

populations mostly overlap, with the exception of individuals from CCS, RT 76-001, and RMR. These three populations are the furthest from other individuals of other populations, but still overlap together. The assignment test supported the overall general pattern of genetic structure for individuals based on the other results from the study. The vast majority of individuals, 82% overall, were re-allocated to their populations of origin (Table 7).

Outcrossing rate estimates

The multilocus outcrossing rate (t_m) of the species was estimated based on six open-pollinated populations including 431 progeny. The estimates for the multilocus outcrossing rates (t_m), for the six populations included, ranged from 0.780 to 0.912 (GH to DC/SH; Table 3). Even though the result for multilocus outcrossing rates is high, there may be some genetic substructuring that is leading to some degree of biparental inbreeding (Sun and Ritland 1998). There is significance in the estimates for biparental inbreeding, which have a range from 0.041 (DENC) to 0.12 (DC; Table 3).

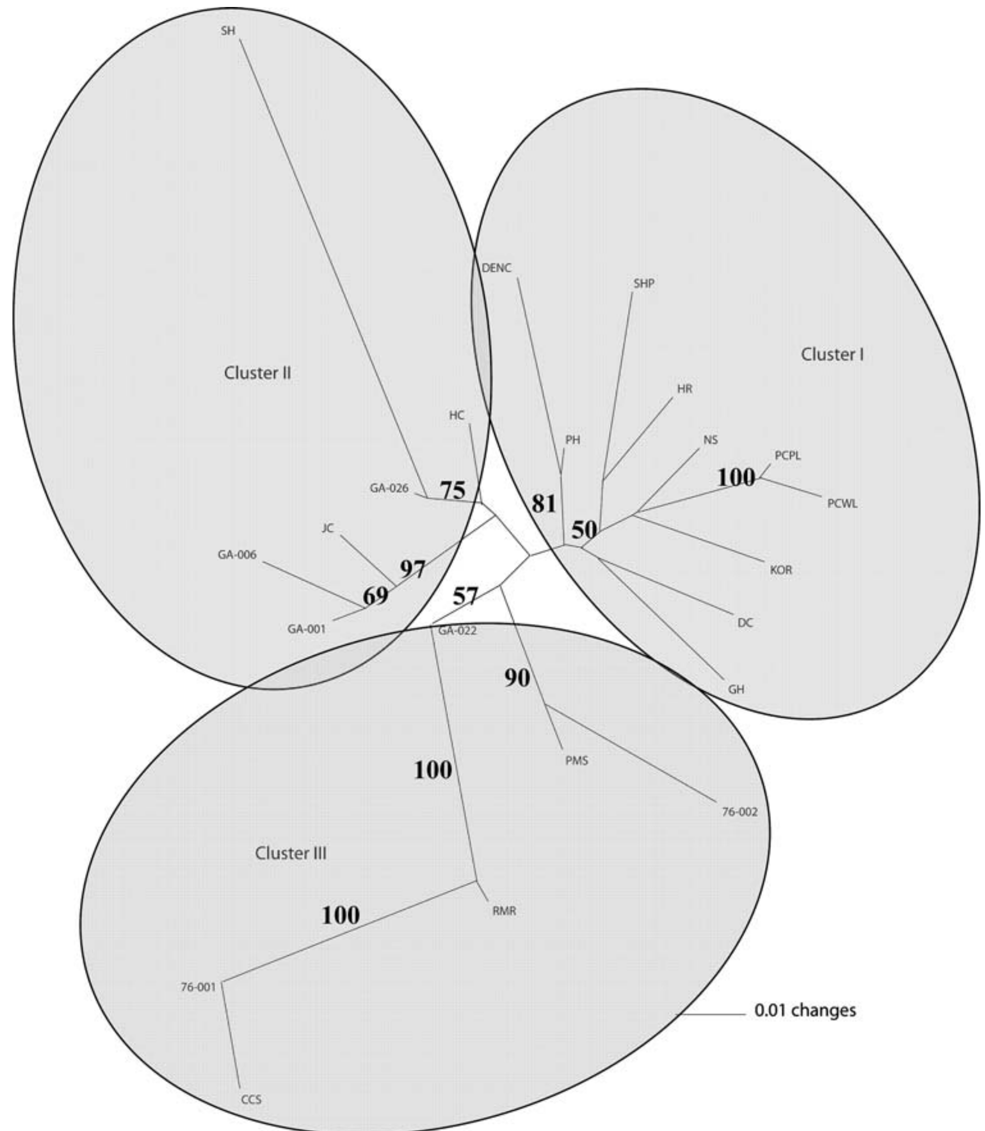

Fig. 1 Neighbor-joining phylogram based on Nei's genetic distance (Lynch and Milligan 1994) data with bootstrap values equal to or higher than 50% shown

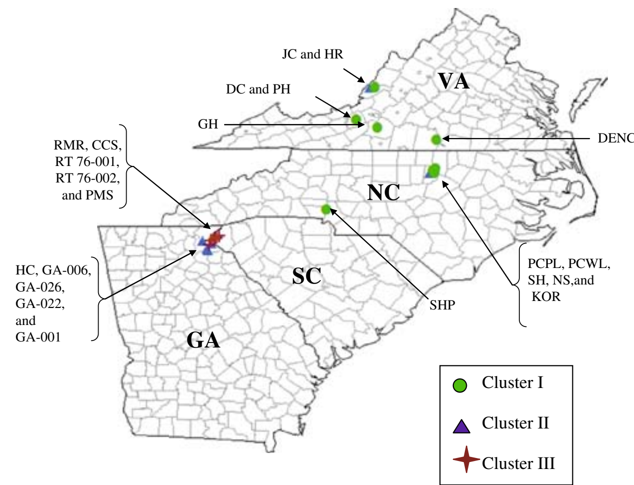

Fig. 2 Map of *E. laevigata* populations (Full names, Table 1) and how they correspond geographically to the clusters from the neighbor-joining phylogram

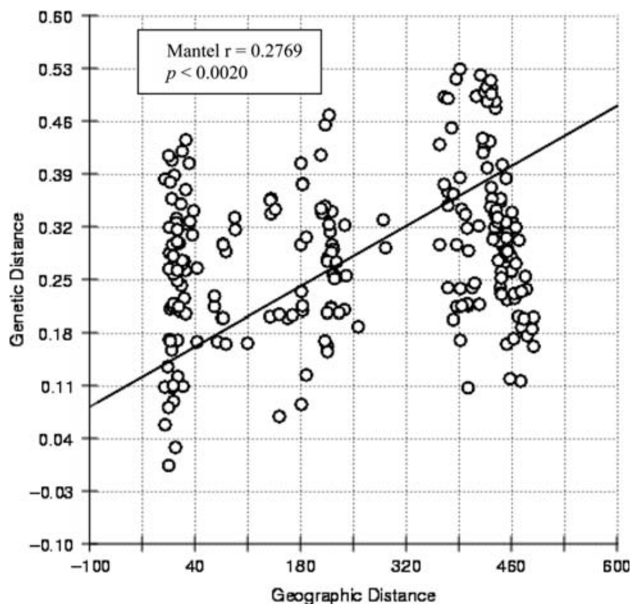

Fig. 3 Relationship between genetic distance (Φ_{ST}) and geographic distance (km) among 22 populations of *E. laevigata*

Discussion

Genetic variation and population structure

Our data show that genetic diversity in *E. laevigata* is not genetically depleted. Overall genetic diversity (0.261) and percentage of polymorphic loci (95%) are higher than the estimates from the allozyme study (0.178 and 63.6% respectively; Apsit and Dixon 2001). Other studies on species of concern with predominantly outcrossing systems have also reported comparable high percentages of polymorphic loci from AFLP analyses (*Euterpe edulis* Mart. 92.07% in Cardoso et al. 2000; *Pedicularis palustris* L. 64% in Schmidt and Jensen 2000; *Eryngium alpinum* L. 54% in Gaudeul et al. 2000; and *Leucopogon obtectus* Benth. 89% in Zawko et al. 2001). The high levels of diversity, which is unexpected for a rare species, can be attributed to a number of factors including population size, geographic distance, breeding system, gene flow, along with processes of selection, genetic drift, mutation, and migration (Hamrick and Godt 1996; Frankham et al. 2004; Luan et al. 2006). A relatively high proportion of the polymorphic loci pairs (25.7%; $P < 0.05$), were found to be at linkage disequilibrium. Linkage disequilibrium, or non-random association of alleles, can be caused by population bottlenecks, recent mixing of different populations, and selection (Frankham et al. 2004). Tero et al. (2003) indicated that new subpopulations in *Silene tatarica* Pers., were established by a few individuals and this founder effect is the probable reason for observed linkage disequilibrium within subpopulations.

The population census data based over the past few years cover a very large range from the largest population of ~50,000 rosettes to the smallest population being ~14 rosettes (Table 1). In theory, it is expected that genetic variation is higher in larger populations; however, we did not observe this trend in our study (Tables 1, 2), which suggests that the small populations of *E. laevigata* have not suffered substantial loss of genetic diversity due to genetic drift. This phenomenon of comparative levels of genetic diversity in small and large populations was also found in studies of other species that sampled from a range of small to somewhat larger populations (Cardoso et al. 2000; Schmidt and Jensen 2000; Gaudeul et al. 2000; Zawko et al. 2001; Tero et al. 2003; Kim et al. 2005). It has been reported that even small amounts of gene flow into small populations can counteract erosive effects of genetic drift (Wright 1969). If one knows more about the mechanism of gene flow between populations one can understand more about what this species may require to persist. Fortunately Gadd (2006) reported that the pollinators of *E. laevigata* are mostly bees (Insecta, Apoidea) and butterflies (Insecta, Rhopalocera) that can travel significant distances, thus potentially facilitating gene flow between the populations.

The AFLP data suggest that genetic differentiation ($G_{ST} = 0.294$; $\Theta^{II} = 0.218$) among populations of *E. laevigata* is substantial. The genetic differentiation for *E. laevigata* falls in the range reported from previous studies of predominately outcrossing species using AFLP data (*Euterpe edulis* Mart. $F_{ST} = 0.426$ in Cardoso et al. 2000; *Pedicularis palustris* L. $G_{ST} = 0.27$-0.89 in Schmidt and Jensen 2000; and *Sonchus gandogeri* Pitard $G_{ST} = 0.149$ in Kim et al. 2005). In an outcrossing, endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), Gaudeul et al. (2000) reported a G_{ST} of 0.42 among populations. Based on

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			
Populations separated by region				
M + P + GA/SC[a] among regions	0.66186	7.16	<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			

[a] M represents the Mountain region, P the Piedmont region, and GA/SC the Georgia and South Carolina region of populations

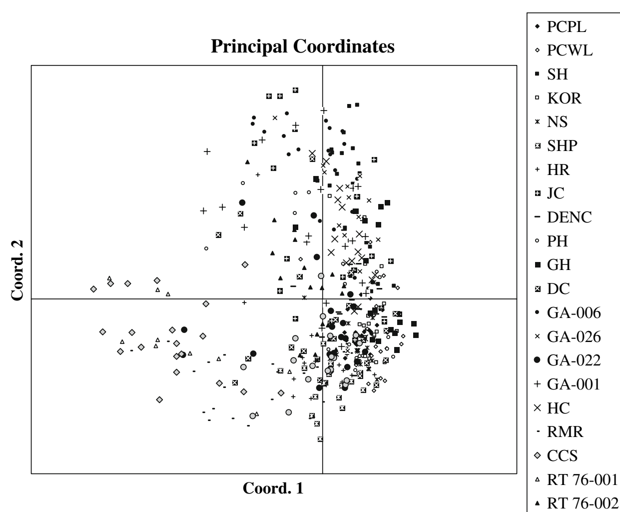

Fig. 4 Principal coordinates analysis (PCoA) of 420 *Echinacea laevigata* individuals from 22 populations, based on a genetic distance matrix of 210 AFLP loci. The first two coordinates explain 50.5% of the variance

current literature (mostly isozyme data), the average genetic differentiation (F_{ST}) for outbreeding species is around 0.2 and is 0.5 for inbreeding or selfing species (Loveless and Hamrick 1984; Hamrick and Godt 1990).

Results from the AMOVA indicated that even though there is significant ($P < 0.0001$; 1,023 permutations) variation among populations (29.74%), most of the molecular variance resides within populations (70.26%; $\Phi_{ST} = 0.297$) (Table 6). This result is similar to the finding from the allozyme study (Apsit and Dixon 2001). Our results of AMOVA by partitioning populations among the three geographical regions (Mountains, Piedmont, and GA/SC) similarly revealed that most genetic variation occurs within populations (68.64%) and little (7.1%, Table 6) occurs among the three regions. This pattern was also reported in several previous studies using dominant markers on rare plant species known to be predominately outcrossing (*Eryngium alpinum*, Gaudeul et al. 2000; *Silene tatarica*, Tero et al. 2003; *Pedicularis palustris*, Schmidt and Jensen 2000; and in *Leucopogon obtectus*, Zawko et al. 2001). The high level of within population diversity in *E. laevigata* and these species could be largely attributed to the predominantly outcrossing breeding system that permits gene flow among populations.

The observed high level of diversity and significant population differentiation in *E. laevigata* could be explained by considering metapopulation models of population structure. Historically, more populations were known to persist and over time many of them have been extirpated (USFWS 1995). Habitat loss due to anthropogenic factors has been reported to be the demise of these populations (USFWS 1995; Apsit and Dixon 2001). It is known that *E. laevigata* prefers a soil rich in magnesium and calcium, and populations with similar substrates group together. Thus gene flow in this species probably operates through two types of metapopulation models. One model can be referred to as a 'source-sink' model, where a source population is the exporter of migrants for surrounding subpopulations; the other is a 'classical' metapopulation model where subpopulations would go extinct without migration from surrounding source populations (Tero et al. 2003). The application of these models to smooth coneflower is further supported by the observed and significant ($P < 0.05$) correlation between geographic and genetic distance (Fig. 2) in the species. It is also supported by the idea of long-distance dispersal because the achenes have a pappus, which aid in dispersal.

While the present study suggests that all populations could have historically comprised a large metapopulation, there is also a degree of structure in the populations and some exhibit isolation by distance. The populations in exception are from the South Carolina region (RT 76-001, RMR, and CCS) (Figs. 3, 4). These three populations are genetically more diverse and are more genetically divergent from their surrounding populations. The assignment

Table 7 Results of the assignment test based on the DOH calculator

	PCPL	PCWL	SH	KOR	NS	SHP	HR	JC	DENC	PH	GH	DC	GA-006	GA-026	GA-022	GA-001	HC	RMR	CCS	76-001	76-002	PMS	%
PCPL	16	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	84
PCWL	0	19	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	95
SH	0	0	23	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
KOR	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
NS	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	95
SHP	0	0	0	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	100
HR	1	0	0	0	0	0	18	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	90
JC	0	2	0	1	0	0	0	11	0	0	0	0	3	1	0	2	0	0	0	0	0	0	55
DENC	0	0	0	0	0	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	10
PH	0	0	1	0	1	0	0	0	1	13	1	0	0	1	0	1	0	0	1	0	0	0	65
GH	0	0	1	0	0	0	0	0	0	0	16	1	1	0	0	0	0	0	0	0	0	0	84
DC	1	0	0	0	1	0	1	0	0	2	1	12	2	0	0	0	0	0	0	0	0	0	60
GA-006	0	0	0	0	0	0	0	0	0	0	1	0	14	1	0	4	0	0	0	0	0	0	70
GA-026	0	0	0	0	0	0	0	0	0	0	1	0	0	9	1	1	1	0	0	0	0	0	69
GA-022	0	0	0	0	0	0	1	0	0	0	0	0	0	0	15	2	0	0	2	0	0	0	75
GA-001	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	14	0	0	0	0	0	0	70
HC	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	16	0	0	0	0	0	94
RMR	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	14	2	0	0	4	70
CCS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	16	1	0	1	80
76–001	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	2	3	4	0	0	40
76–002	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	19	0	95
PMS	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0	18	90

The percentage (%) of individuals assigned correctly back to their respective populations is also reported

test results, based on the number of individuals correctly assigned back to their population, and percentage of polymorphic loci from these populations were significant but only to a moderate to high degree (assignment test: RT 76–001 = 40%, RMR = 70%, and CCS = 80%; percentage of polymorphic loci: RT 76–001 = 50%, RMR = 46%, and CCS = 69%). Significant differentiation coupled with recurrent gene flow could be explained by pollinator movement (pollen flow) and future studies may examine maternal vs. paternal gene flow among these populations. Evidence from the PCoA and the high percentages of individuals that were correctly assigned back to their source population, suggest that some populations are clustering together to form subpopulations within the overall metapopulation. Over time, if habitats are continuously lost gene flow may be limited among subpopulations, which will continue to differentiate. The NJ tree also shows that SH from North Carolina is on a particularly long branch (in Cluster II of Fig. 3), indicating that it is genetically highly divergent. This population also had 100% of the individuals correctly assigned back, from the assignment test, and 52% polymorphic loci. However, populations SH from North Carolina and JC from Virginia did not group with populations from their geographic proximity, but grouped with populations further south (northeastern Georgia and adjacent South Carolina). An explanation for this phenomenon could be long distance gene flow via pollen or seed dispersal across the range of the species when more intermediate populations were present to increase the gene pool. Alternatively, these two populations, SH and JC, could represent relics of an ancestral large population, or metapopulation, which once continuously stretched from the south to the north (NC and VA). Lastly, this difference in genetic structure could be due to human transport of seeds or from convergence due to common selection forces in geographically disparate populations acting on key loci in linkage disequilibrium with the markers used in this study (Bonin et al. 2006). The last hypothesis could be tested by conducting an outlier detection analysis.

Outcrossing rates

Our MLTR analysis suggested that *E. laevigata* is a predominately outcrossing species [t_m 0.780 to 0.912] congruent with the results from the flower bagging treatments and field observations of selected populations from McGregor (1968) and Gadd (2006). The standard deviations of the estimated t_m were notably high (Table 3) indicating that the range of the estimates for outcrossing rates from the 100 bootstrap replicates is large. This likely resulted from the relatively small number of sampled families and small number of progenies sampled for some families, which could cause the bootstrapping to be skewed for both higher and lower estimates of the outcrossing rate. Although up to 30 seed heads were collected from each population for the outcrossing analysis, only the heads with viable seeds that germinated can be included in the analysis, which led to the variation in sampling. Precautions were taken to minimize errors in the AFLP genotyping, but error may still be present (Bonin et al. 2004; Pompanon et al. 2005). This could have caused a slight overestimate of the outcrossing rate. Nonetheless, the overall estimates of t_m are high enough to suggest predominantly outcrossing in the species, compared to the t_m values based on AFLP analyses in other outcrossing species (*Eriogonum ovalifolium* var. *vineum* jeps. = 0.88, Neel et al. 2001; *Euterpe edulis* Pursh. = 0.94, Gaiotto et al. 2003; and *Arbutus menziesii* Pursh. = 0.97, Beland et al. 2005). We have also included an estimate of biparental inbreeding, or mating between close relatives, which could also explain why the outcrossing rate was not equal to one.

The mating system of a species is closely correlated with the level and pattern of gene flow among populations (Barrett 2003), thus permitting the prediction of these variables in a species if the mating system is known. Outcrossing species are typically found to have higher levels of gene flow among populations than selfing species. In a recent study of toadflaxes (*Linaria*), the authors compare the differences in diversity levels depending on whether the species is self-compatible or self-incompatible (Segarra-Moragues and Mateu-Andrés 2007). They found that in the self-incompatible species of *Linaria*, the mean number of alleles per locus, total genetic diversity and genetic diversity within populations were higher when compared to the selfing species (Segarra-Moragues and Mateu-Andrés 2007). It has also been reported that outcrossing species maintain a higher level of genetic diversity within species and populations, and show lower genetic differentiation among populations (Zawko et al. 2001). It is likely that the outcrossing system in *E. laevigata* plays a critical role in maintaining the high genetic variation in the species. This implies that although lower genetic diversity is now not evident in small populations of *E. laevigata*, loss of genetic variation as a result of reduction in outcrossing due to random loss of self-incompatibility alleles may occur in the future if the size of these populations is further reduced by any reason. Therefore, it is critical to maintain and increase population sizes across the smooth coneflower's range to ensure conservation of the species.

Conclusion

The extant populations of *E. laevigata* have significant levels of population diversity, and exhibit substantial

population genetic differentiation. It is encouraging that the species is outcrossing and gene flow is occurring to maintain the genetic diversity because the long-term survival of this species will depend on sufficient genetic diversity. Since *E. laevigata* is a prairie species, and favors disturbance in order to persist, it is not surprising that the populations in the power line rights-of-way are doing extremely well in population size and genetic diversity. Rights-of-way are kept free of large trees and so doing a periodic disturbance regime may facilitate persistence of relict prairie species such as *E. laevigata*. We suggest that the rarity of this species resulted from anthropogenic landscape alteration.

Ongoing efforts to help reintroduce and manage populations of *E. laevigata* are underway. For example, a study by Alley and Affolter (2004) focused on the requirements for reintroducing the species to areas based in Georgia. It is important for conservation agencies to be aware of the genetic structure and isolation by distance exhibited by these populations. Future management efforts should work to ensure continued levels of gene flow and high levels of diversity across the range of this species. Managers can use the information about which populations clustered together when planning for reintroductions or management of the current populations. Future loss of habitats that create geographic isolation can pose potential risks to species by limiting gene flow among populations. Population size can affect future diversity levels if populations continue to decline. A suggestion for management efforts in smaller populations would be to clear any encroachment and if possible perform prescribed burns or mowing regimes. It is evident from the populations in the powerline rights-of-way that the previously mentioned types of disturbance are beneficial to this species. Maintenance of habitats 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 should be set in order to meet the goals described by the recovery plan, which may be more informed as a result of this study.

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; and the lab of Dr. Fred Gould. A special thanks to Dr. Edward Vargo, NC State Entomology Department, for letting me use lab equipment and reading over our reviews. Others who aided in some aspect of this study were Jeff Essic (NCSU), GIS; Kathy McKeown, seed stratification; Jordan Mueller, Unix and AWK programming; Kermit Ritland, outcrossing rate program. The authors would also like to thank the anonymous reviewers for their time and helpful comments to improve this manuscript. This study was partially supported by a grant funded by the VA Nature Conservancy.

References

Alley H, Affolter JM (2004) Experimental comparison of reintroduction methods for the endangered *Echinacea laevigata* (Boynton and Beadle) Blake. Nat Areas J 24:345–350

Apsit VJ, Dixon PM (2001) Genetic diversity and population structure in *Echinacea laevigata* (Boynton and Beadle) Blake, and endangered plant species. Nat Areas J 21:71–77

Barrett SCH (2003) Mating strategies in flowering plants: the outcrossing-selfing paradigm and beyond. Philos Trans R Soc Lond B Biol Sci 358:991–1004

Baus E, Darrock J, Bruford MW (2005) Gene-flow patterns in Atlantic and Mediterranean populations of the Lusitanian sea star *Asterina gibbosa*. Mol Ecol 14:3373–3382

Beland JD, Krakowski J, Ritland CE, Ritland K, El-Kassaby YA (2005) Genetic structure and mating system of northern *Arbutus menziesii* (Ericaceae) populations. Can J Bot 83:1581–1589

Bohonak AJ (2002) IBD (Isolation By Distance): a program for analyses of isolation by distance. J Hered 93:153–154

Bonin A, Bellemain E, Eidesen PB, Pompanon F, Brochmann C, Taberlet P (2004) How to track and assess genotyping errors in population genetics studies. Mol Ecol 13:3261–3273

Bonin A, Taberlet P, Miaud C, Pompanon F (2006) Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). Mol Biol Evol 23:773–783

Brzustowski J (2002) DOH assignment test calculator [Online.] Available via www2.biology.ualberta.ca/jbrzusto/Doh.php. Cited 12 Sept 2007

Cardoso SRS, Eloy NB, Provan J, Cardoso MA, Ferreira PCG (2000) Genetic differentiation of *Euterpe edulis* Mart. Populations estimated by AFLP analysis. Mol Ecol 9:1753–1760

Caujapé-Castells J, Baccarani-Rosas M (2005) Transformer-3b.01: a program for the analysis of molecular population genetic data. Exegen software

Cullings KW (1992) Design and testing of plant-specific PCR primer for ecological and evolutionary studies. Mol Ecol 1:233–240

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure from small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

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

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

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

Frankham R, Ballou JD, Briscoe DA (2004) Introduction to conservation genetics. New York, USA

Gadd LE (2006) Pollination biology of the federally endangered *Echinacea laevigata* (Boynton and Beadle) Blake, smooth coneflower, in small isolated populations. Dissertation, North Carolina State University

Gaddy LL (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, Taberlet P, Till-Bottraud I (2000) Genetic diversity in an endangered alpine plant, *Eryngium alpinum* L. (Apiaceae), inferred from amplified fragment length polymorphism markers. Mol Ecol 9:1625–1637

Gaiotto FA, Grattapaglia D, Vencovsky R (2003) Genetic structure, mating system, and long-distance gene flow in heart of palm (*Euterpe edulis* Mart.). J Hered 94:399–406

Hamrick JL, Godt MJW (1990) Allozyme diversity in plant species. In: Brown AHD, Clegg MT, Kahler AL, Weir BS (eds) Plant population genetics, breeding and genetic resources. Sinauer Sunderland pp 43–63

Hamrick JL, Godt MJW (1996) Conservation genetics of endemic plant species. In: Avise JC, Hamrick JL (eds) Conservation genetics: case histories from nature. New York

Holsinger KE, Lewis PO, Dey DK (2002) A Bayesian approaches to inferring population structure from dominant markers. Mol Ecol 11:1157–1164

KeyGene NV (2003) AFLP Quantar. KeyGene NV. Wageningen, The Netherlands

Kim S-C, Lee C, Santos-Guerra A (2005) Genetic analysis and conservation of the endangered Canary Island woody sow-thistle, *Sonchus gandogeri* (Asteraceae). J Plant Res 118:147–153

Loveless MD, Hamrick JL (1984) Ecological determinants of genetic structure in plant populations. Annu Rev Ecol Syst 15:65–95

Luan S, Chiang TY, Gong X (2006) High genetic diversity vs. low genetic differentiation in *Nouelia insignis* (Asteraceae), a narrowly distributed and endemic species in China, revealed by ISSR fingerprinting. Ann Bot 98:583–589

McGregor RL (1968) The taxonomy of the genus *Echinacea* (Compositae). Univ Kans Sci Bull 48:113–142

Neel MC, Ross-Ibarra J, Ellstrand NC (2001) Implications of mating patterns for conservation of the endangered plant *Eriogonum ovalifolium* var. *vineum* (Polygonaceae). Am J Bot 88:1214–1222

Nei M (1978) Estimation of average heterozygosity and genetic distance from a small number of individuals. Genetics 89:583–590

Palacios C, Kresovich S, González-Candelas F (1999) A population genetic study of the endangered plant species *Limonium dufourii* (*Plumbaginaceae*) based on amplified fragment length polymorphism (AFLP). Mol Ecol 8:645–657

Peakall R, Smouse PE (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for research and teaching. Mol Ecol Notes 6:288–295

Pompanon F, Bonin A, Bellemain E, Taberlet P (2005) Genotyping errors: causes, consequences and solutions. Nat Rev Genet 6:847–859

Ritland K, Jain SK (1981) A model for the estimation of outcrossing rate and gene frequencies using n independent loci. Heredity 47:35–52 (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

Sari AO, Morales MR, Simon JE (1999) *Echinacea angustifolia*: an emerging medicinal. In: Janick J (ed) Perspectives on new crops and new uses. Alexandria, VA pp 490–493

Schmidt K, Jensen K (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. Am J Bot 87:678–689

Schneider S, Roessli D, Excoffier L (2000) Arlequin: a software for population genetics data analysis, version 2.000. Genetics and Biometry Lab, Dept. of Anthropology, University of Geneva

Segarra-Moragues JG, Mateu-Andrés I (2007) Levels of allozyme diversity in closely related toadflaxes (*Linaria*, Plantaginaceae) and their correspondence with the breeding systems of the species. Conserv Genet 8:373–383

Swofford DL (2002) PAUP*: phylogenetic analysis using parsimony (*and other methods). Version 4.0b10. Sinauer Associates, Sunderland, Massachusetts, USA

Sun M, Ritland K (1998) Mating system of yellow starthistle (*Centaurea solstitialis*), a successful colonizer in North America. Heredity 80:225–232

Tero N, Aspi J, Siikamäki P, Jäkäläniemi A, Tuomi J (2003) Genetic structure and gene flow in a metapopulation of an endangered plant species, *Silene tatarica*. Mol Ecol 12:2073–2085

USFWS (U.S. Fish and Wildlife Service) (1995) Smooth coneflower recovery plan. Atlanta, GA, pp 31

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

Vos P, Hogers R, Bleeker M, Reijans M, De Lee TV, Hornes M, Fritjers A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

Wright S (1969) Evolution and the genetics of populations, vol. 2, The theory of gene frequencies. Chicago, IL

Yeh FC, Yang RC, and Boyle T (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, Drauss SL, Dixon KW, Sivasithamparam K (2001) Conservation genetics of the rare and endangered *Leucopogon obtectus* (Ericaceae). Mol Ecol 10:2389–2396

Zhang Defang, Chen Shilong, Chen Shengyun, Zhang Dejun, Gao Q (2007) Patterns of genetic variation in *Swertia przewalskii*, and endangered endemic species of the Qinghai-Tibet Plateau. Biochem Genet 45:33–50