Molecular Genetic Variability of *Spigelia marilandica* and *S. gentianoides*

Amanda J. Hershberger

Department of Horticulture, University of Georgia, 1109 Experiment Street, Griffin, GA 30223

Tracie M. Jenkins

Department of Entomology, University of Georgia, 1109 Experiment Street, Griffin, GA 30223

Carol Robacker¹

Department of Horticulture, University of Georgia, 1109 Experiment Street, Griffin, GA 30223

ADDITIONAL INDEX WORDS. native plants, endangered species, population structure, within-population diversity, among-population diversity, phenogram

ABSTRACT. Despite the ecologic and ornamental potential of southeastern U.S. native *Spigelia*, little is known about the intraspecific or the interpopulation genetic variation. The southeastern U.S. native *Spigelia* habitat is becoming more and more fragmented as a result of human activity, making it imperative to gain an understanding of natural genetic variation among and within species and populations for the purpose of obtaining variability for plant breeding and preserve the genetic variability in *Spigelia*. Therefore, the objective of this study was to use amplified fragment length polymorphism analysis to determine interspecific and intraspecific genetic variation and to evaluate gene flow. Thirteen populations of two species of native *Spigelia*, *S. marilandica* (SM), *S. gentianoides* var. *gentianoides* (SGG), and *S. gentianoides* var. *alabamensis* (SGA), were analyzed using four primer pairs that amplified a total of 269 bands. Based on analysis of molecular variance and estimates of Nei's coefficients of gene diversity (percentage of polymorphic loci, average genetic diversity within populations, average genetic diversity within species, and proportion of species genetic diversity attributed to among population variation), the majority of variation found in *Spigelia* occurs within populations. Both among-species and among-population variation was low, likely the effect of common ancestry as well as relatively frequent introgression among individuals (and populations) of *Spigelia*. When all individuals were evaluated using Nei's unbiased genetic distances and viewed as a unweighted pair group method with arithmetic mean phenogram, three main groups were shown, one with two samples of SGG from one population, one with 13 individuals from both SGG populations used in this study, and one with all of the SM, SGA, and remaining SGG individuals. Further evaluation using STRUCTURE software showed introgression between populations and species, although all allele clusters have not entirely introgressed into all populations. The significance of these results is discussed in relation to breeding in *Spigelia*.

Spigelia (Loganiaceae) is a genus of ≈ 50 species that ranges from the southeastern United States to Central America and south to temperate areas of South America. Only five species are considered endemic to the United States and two of these are indigenous to the southeastern United States: *S. marilandica* and the federally endangered *Spigelia gentianoides* (Gould, 1997). To aid recovery of *S. gentianoides*, research on the genetic variability of the natural populations of this species has been encouraged (Negron-Ortiz, 2012). Furthermore, these species are of interest to plant breeders as a result of their ornamental potential. Therefore, for purposes of obtaining variability for plant breeding and to preserve the genetic variability in *Spigelia*, evaluation of the genetic variability within and among the species is essential. Observation of physical traits of *Spigelia* species has shown morphological variation among species, populations within a species, and between plants of the same population. Although these morphological differences have been observed, the degree

of genetic difference in these two species has yet to be determined.

Spigelia gentianoides (SG) is an upright, perennial herbaceous species that has two disjunct botanical varieties, and both are considered endangered (Gould, 1997). Threats to this species include loss or alteration of habitat, lack of natural disturbance regimes, unprotected populations on private lands, and competition from invasive species (Negron-Ortiz, 2012). *Spigelia gentianoides* var. *alabamensis* has pink corollas from 36 to 50 mm long and their lobes reflex (open) at anthesis. Flowering occurs from May to June. This botanical variety is endemic to 17 glades that have developed over an ancient rock formation known as Ketona Dolomite in Bibb County, AL (Negron-Ortiz, 2012). The glades vary in size from 0.1 to 5 ha. *Spigelia gentianoides* var. *gentianoides* has pink flowers that are barely opening and not reflexed at anthesis (Table 1). Flowering is from May to July. SGA differs from SGG in that it has longer corollas with broader throats, longer lobes as well as longer sepals (Table 1) (Gould, 1997). SGG is restricted to five locations within three counties in the Florida Panhandle and southeastern Alabama.

Gould labeled SGA as a botanical variety as opposed to a species designation (1997) despite listing numerous and distinct trait differences between the two botanical varieties. Subsequent assessment of SG and its two botanical varieties suggests that the two SG botanical varieties warrant specific

Received for publication 26 June 2014. Accepted for publication 18 Dec. 2014. This study is a portion of a dissertation submitted by Amanda Hershberger in partial fulfillment of the requirements for a PhD degree.

We thank Drs. Noelle Barkley and Zhenbang Chen and Mr. Tyler Eaton for assistance with amplified fragment length polymorphism methodology and data analysis.

¹Corresponding author. E-mail: crobback@uga.edu.

rank (Weakley et al., 2011), although additional molecular and morphological studies were suggested.

Spigelia marilandica is an upright perennial herbaceous plant growing from 30 to 60 cm high. Corollas are bright red outside and yellow to greenish yellow inside (Table 1). In Kentucky, the flowering period starts in late May through June; occasionally scattered blooms will occur in the fall (Dunwell, 2003). It is thought to be pollinated by ruby-throated hummingbird [*Archilochus colubris* (Cullina, 2000)] as well as by insects (Affolter, 2005; Rogers, 1988). Distribution is from eastern South Carolina to north-central and northwestern Florida, westward to southeastern Oklahoma and far eastern Texas, and north to Kentucky (Duncan and Duncan, 2005; Gould, 1997).

Limited research has been conducted to determine genetic diversity in *Spigelia*. Gould (1997) used internal transcribed spacer (ITS) sequencing to assess relationships of 14 species of *Spigelia*, whereas Affolter (2005) used allozymes to evaluate genetic variability in both botanical varieties of SG as well as in SM.

A technique that is especially useful to assess genetic diversity within and among species when no sequence information is available, as is the case in *Spigelia*, is amplified fragment length polymorphism (AFLP) analysis. AFLP technology has been used to characterize genetic diversity within many genera of plants including horticultural crops such as *Chrysanthemum* sp. (Zhang et al., 2010), sweetpotato [*Ipomoea batatas* (Cervantes-Flores et al., 2008)], pecan [*Carya illinoensis* (Beedanagari et al., 2005)], *Rhododendron* sp. (Chappell et al., 2008), barberry [*Berberis thunbergii* (Lubell et al., 2009)], and *Lactuca* sp. (Koopman et al., 2001). AFLP markers have a very high diversity index (a measure of evaluation efficiency that combines the effective number of alleles identified per locus

and the number of polymorphic bands in each assay) (Philips and Vasil, 2001; Russell et al., 1997) and is therefore appropriate for use in the evaluation of genetic variability in *Spigelia*. We used AFLP technology to determine genetic differences within and among populations of two *Spigelia* species indigenous to the southeastern United States.

Materials and Methods

PLANT MATERIAL. Ten wild populations of SM, one of SGA, and two of SGG were used in genetic analysis (Table 2; Fig. 1). Before collection of plant material, the area was surveyed to define the population area. Plants were collected evenly throughout the population and locations of individual plants collected in each population were recorded using global positioning system coordinates. The populations' size ranged from 14 to 800 and the number of plants evaluated per population ranged from 4 to 11 (Table 2). Approximately 100 mg of immature leaves of each sample were collected on-site, placed on ice in a cooler, and transported to the laboratory for immediate extraction of DNA.

DNA EXTRACTION PROTOCOL AND QUANTIFICATION. DNA extraction was carried out using the E.Z.N.A. plant DNA kit (Omega Bio-Tek, Norcross, GA) with slight modifications including the addition of 26 µL β-mercaptoethanol to Step 1, an increase of the incubation time at 65 °C for 30 min in Step 2, the use of isopropanol stored at -20 °C in Step 4, and the use of 0.9 µL carrier RNA in the initial column step. DNA was tested for quantity and quality (shearing) using a standard agarose gel with the Low DNA Mass Ladder (Invitrogen™, Carlsbad, CA). Subsequently, genomic DNA was stored at 4 °C until AFLP analysis was performed.

Table 1. Phenotypic differences among *Spigelia gentianoides* var. *gentianoides* (SGG), *S. gentianoides* var. *alabamensis* (SGA), and *S. marilandica* (SM).

	Ht (cm)	Leaf shape	Flowers (no./inflorescence)	Corolla length (mm)	Corolla lobes at anthesis	Flower color
SGG	10 to 30	Broadly ovate	3 to 8	25 to 30	Not reflexed	Pink
SGA	10 to 30	Lanceolate to elliptic	2 to 4	36 to 50	Reflexed	Pink
SM	30 to 60	Lanceolate to elliptic	4 to 17	25 to 40	Reflexed	Red outside, green inside

Table 2. Locations of *Spigelia marilandica* (SM), *S. gentianoides* var. *alabamensis* (SGA), and *S. gentianoides* var. *gentianoides* (SGG) populations collected and grouped by species.^z

Population no.	Species	Location	Sample size	Approximate population size	Polymorphic loci (%)
1	SM	Hamilton County, TN	10	40	55.76
2	SM	DeKalb County, GA	10	14	50.93
3	SM	Marianna County, FL	9	500	44.98
4	SM	Edgefield County, SC	7	40	35.32
5	SM	Oconee County, SC	10	40	49.44
6	SM	Houston County, GA	5	200	34.57
7	SM	Houston County, GA	10	50	52.79
8	SM	Murray County, GA	10	800	55.39
9	SM	Charleston County, SC	4	400	31.97
10	SM	Wakulla County, FL	10	200	44.24
11	SGA	Bibb County, AL	10	400	50.19
12	SGG	Geneva County, AL	10	400	47.69
13	SGG	Jackson County, FL	11	500	42.75

^zEach population is numbered in succession and includes four to 11 individual plant samples.

Fig. 1. *Spigelia* collection sites: *S. marilandica* (triangles), *S. gentianoides* var. *gentianoides* (circles), and *S. gentianoides* var. *alabamensis* (squares). Numbers within symbols indicate assigned population numbers described in Table 2.

AFLP PROCEDURE. Generation of restriction fragments was accomplished using two restriction endonucleases, *EcoRI* and *MseI*, to fragment the genome, following the protocol described by Vos et al. (1995). Restriction-digestion, ligation, and pre-selective amplification of genomic DNA were carried out using the IRDye® AFLP Template Preparation Kit (LI-COR Biosciences, Lincoln, NE). All polymerase chain reactions (PCRs) were carried out in a PCR system (Model 9600 Thermal Cycler®; PerkinElmer, Waltham, MA). Thirty-two primer combinations containing both 700 and 800 wavelength IRDye® labeled *EcoRI* selective primers were screened over four individuals, two samples from separate SM populations, one sample from a SGG population, and one sample from a SGA population. After initial screening of the four samples, nine primer combinations were selected on the basis of number of polymorphic bands visualized on a gel. These were used to screen nine samples: five from separate SM populations, two from the same SGG population, one from a separate SGG population, and one from a SGA population. This screening of nine primer combinations included the same four samples from the initial screening. Based on the number of polymorphic bands and their repeatability on these nine samples, four primer pairs were selected for this study: *MseI*-CTC/*EcoRI*-ACC, *MseI*-CTC/*EcoRI*-ACG, *MseI*-CTT/*EcoRI*-ACT, and *MseI*-CTG/*EcoRI*-ACG. Selective amplification was

carried out on all individuals with each of the four selected primer pairs. In a 1.5-mL microcentrifuge tube (Fisher Scientific, Waltham, MA), the following was combined: 2.15 µL sterile deionized water, 2.0 µL MgCl₂ (Promega, Madison, WI), 0.05 µL GOTaq® DNA polymerase (Promega), 2 µL 5× GOTaq buffer (Promega), 0.8 µL 100 mM dNTP (Promega), 0.5 µL *MseI* primer, and 0.5 µL *EcoRI* primer. In each well of a 96-well PCR plate (Fisherbrand; Fisher Scientific), 8 µL of the aforementioned mix was combined with 4 µL of template DNA from the preselective amplification stage. PCR conditions for selective amplification were set based on the LI-COR Biosciences AFLP protocol. After completion of the selective amplification PCR program, 5 µL of Blue Stop Solution® (LI-COR Biosciences) was added to each well and samples denatured at 94 °C for 4 min. Samples were then cooled to 4 °C using the thermocycler.

Gels were cast using LI-COR Biosciences 25-cm glass plates with 0.25-mm spacers. Twenty milliliters of 6.5% KB Plus acrylamide gel solution was combined with 150 µL APS (Fisher Scientific) and 15 µL TEMED (Fisher Scientific). DNA was loaded at a volume of 0.5 µL per well. Each gel included individuals representing both species in this study. Each gel also included three standards, four test samples, and one blank lane to enable efficient and reliable gel comparison in the

scoring process. The first standard to be used was a 50:50 mixture of LI-COR Biosciences IRDye700® and IRDye800® 50- to 700-bp size standards placed on the outside two lanes of each gel and in the middle. The second standard was four test samples representing two populations of SM, one population of SGG, and one from a SGA population placed in the same locations in each gel. Extraction and analysis was repeated in 10% of the individuals to ensure repeatability of banding patterns.

Gels were run on a LI-COR Biosciences Model 4300S DNA Analysis System using the SagaLite® software package (LI-COR Biosciences) with the laser focus adjusted on a run-by-run basis to optimize performance. Run length was set to 4.5 h with KBplus standard electrophoresis conditions. Standard power and temperature settings were used with the exception of voltage that was reduced to 1000 to allow for low-bp band separation. Gel images produced by SagaLite® were graphically adjusted within the program and exported to GelBuddy (Zerr and Henikoff, 2005). Image files were then exported to Photoshop® CS2 (Adobe Systems, San Jose, CA) and all gel images from a single primer pair merged into a single graphics file. Individual gel images were aligned using two standards: the LI-COR Biosciences IRDye® 50- to 700-bp size standard and four test samples. The resulting single graphics file was used in the scoring of gels.

DATA ANALYSIS. Bands were manually scored in binary format as present (1) or absent (0) and values were recorded in Excel (Microsoft, Redmond, WA). PopGene Version 1.32 (Yeh et al., 1997) was used to calculate Nei's genetic diversity (Nei, 1987) and percentage of polymorphic loci. Settings for analysis included a significance level of $P \leq 0.05$, three groups (one for each species and botanical variety) when comparing species and botanical varieties, 13 groups (one for each population) when evaluating all populations, and 10,000 simulations. A matrix of Nei's unbiased genetic distances (Nei, 1978) was calculated with PopGene Version 1.32 using all markers and used to construct an unrooted unweighted pair group method with arithmetic mean (UPGMA) phenogram with PHYLIP Version 3.69 software (Felsenstein, 2009). Within PHYLIP, the programs used were seqboot for bootstrap analysis, restdist, and neighbor to create 1000 UPGMA phenograms, and consense to create one single UPGMA consensus tree using the majority rule function of the 1000 trees created in neighbor. TreeView 1.6.6 (Page, 1996) was used to view phenograms. Bootstrap support of less than 50% for nodes was considered collapsed in the phenogram and colored red. Analysis of molecular variance (AMOVA) was calculated among species using Arlequin Version 3.5.1.3 (Excoffier and Lischer, 2010) to determine the hierarchical partitioning of genetic variability among all species, populations within a single species, and within each population. Population structure was evaluated using STRUCTURE software Version 2.3.2.1 (Falush et al., 2003; Pritchard et al., 2000). This method uses a Markov Chain Monte Carlo algorithm to cluster individuals into populations on the basis of multilocus genotype data. Parameters of STRUCTURE were set to use the admixture model and correlated alleles frequencies model as is recommended in cases of subtle population structure (Falush et al., 2003). The degree of admixture (α) was inferred from the data. The number of population clusters (K) was estimated by performing 10 independent runs for each K (from 1 to 15). Each run was performed using 5000 replicates for burn-in and

5000 during the analysis. The highest likelihood was used to select K after 10 runs at each K . On this basis, $K = 6$ in this study. Ten independent runs were performed at $K = 6$ to assess convergence of the data to verify that estimates were consistent across runs (J.K. Pritchard, personal communication).

Results and Discussion

LEVEL OF POLYMORPHISM. AFLPs among and within populations of each species were analyzed to determine the genetic differences. The same AFLP data set obtained in the analysis of among species differences was used to evaluate among-population differences and within-population differences. The ability to use the same data set for multiple analyses is facilitated by scoring each individual plant separately. Four AFLP primer combinations amplified a total of 269 scorable bands. The average repeatability of AFLP fragments across two replications was 97.9% (data not shown). The numbers of amplified bands used per primer pair are as follows: *MseI*-CTC/*EcoRI*-ACC amplified 60 bands, *MseI*-CTC/*EcoRI*-ACG amplified 50 bands, *MseI*-CTG/*EcoRI*-ACG amplified 66 bands, and *MseI*-CTT/*EcoRI*-ACT amplified 93 bands.

The percentage of polymorphic loci across all species was 97.8% (Table 3). This is similar to that observed in azalea (*Rhododendron* sp. section *Pentanthera*), where the percentage of polymorphic loci across seven species was 89.7% (Chappell et al., 2008). Within species, polymorphic band percentages ranged from 59.1% in SGG to 88.8% in SM. The high degree of polymorphism is attributable primarily to SM; however, although combined values for both SG botanical varieties were 30% lower than SM, the polymorphisms found were still moderate in this species. When evaluated by individual populations, the polymorphic loci percentage ranged from 35.3% in population 4 to 55.8% in population 1, both SM populations (Table 2). Individual populations of SGG had 42.8% and 47.7% polymorphic loci in populations 13 and 12, respectively. The percentage of polymorphic loci found within population 11 (SGA) was 50.2%. There was no relationship between population size and polymorphic loci present in individual populations. In azalea, polymorphic band percentages within species ranged from 86.8% to 91.8% (Chappell et al., 2008). The high degree of polymorphism in azalea is not the result of a single species, but rather polymorphic loci are spread evenly across all species and individual populations.

DIVERSITY WITHIN AND AMONG SPECIES. Actual level of diversity and how diversity is proportioned among species and populations were measured. AMOVA and G_{ST} values

Table 3. Percentage of polymorphic loci, average genetic diversity within populations (H_S), average genetic diversity within species (H_T), and proportion of species genetic diversity attributed to among population variation (G_{ST}) for *Spigelia marilandica* (SM), *S. gentianoides* var. *alabamensis* (SGA), and *S. gentianoides* var. *gentianoides* (SGG).^z

	Polymorphic loci (%)	H_S	H_T	G_{ST}
Overall	97.77	0.144	0.186	0.226
SM	88.85	0.140	0.175	0.199
SGA and SGG	72.86	0.156	0.191	0.186
SGG	59.11	0.152	0.173	0.125

^zData compiled using PopGene Version 1.32 software (Yeh et al., 1997).

correspond to the proportion (percentage) of genetic variation partitioned among species, among populations, and/or within populations. H_S and H_T values, conversely, are a direct measure of diversity within populations and within species, respectively (Falk et al., 2001). Genetic diversity within species (H_T) was low, ranging from 0.17 in SGG to 0.19 when both SG botanical varieties' data were combined (Table 3). Because SM grows throughout a large area of the southeastern United States, it was expected that SM would have greater genetic variability than SGG and SGA, because they grow in very narrow geographic areas. However, each species/botanical variety occupies a fairly narrow ecological niche. Typically SM grows in calcareous wooded areas rich in organic matter, whereas SGG grows in the sandy loam of upland oak–pine woods of north–central Florida that are typically moist yet well drained. SGA occurs in an area of dolomitic limestone in Bibb County, AL. The narrow range of ecological niches within species may explain low within-species diversity estimates. Population genetic theory predicts that less variable environments will result in a more narrow range of genetic variation within species (Chesson, 1985; Cohen, 1966; Tilman, 1999; Tuljapurkar, 1989). Deciduous azaleas are distributed throughout the southeastern United States in more diverse environments than are SM, SGA, and SGG. H_T values for within the seven species of azalea ranged from 0.34 to 0.41, indicating a greater amount of diversity in these species (Chappell et al., 2008).

AMOVA results (Table 4) indicated that most of the variation was found within species/botanical variety (84.7%) rather than among species/botanical variety (15.3%). This low variability contrasts with the considerable phenotypic differences between SM and SG (Table 1). Similar results were found by Affolter (2005) using allozymes to compare genetic variability between an ex situ population of SGG at Bok Tower Gardens (Lake Wales, FL) to natural populations of SGA from the Alabama Ketona glades. On the basis of the loci studied, the Bok Tower sample was composed of a relatively narrow subset of the genetic variation observed in the SGA populations (Affolter, 2005). According to ITS sequence data, SG is the sister species of SM, which overlaps its range in northern Florida (Gould, 1997). The ITS region is known to be informative at the interspecific level in many groups of angiosperms (Baldwin et al., 1995; Kim and Jansen, 1994). The agreement between ITS data and AFLP data from this study indicates that members of *Spigelia* species used in this study are highly related and possibly derived from a common ancestor. Further research such as chloroplast DNA studies in addition to previous research conducted on three other southeastern U.S. native *Spigelia* species (*S. texana*, *S. loganioides*, and *S. hedyotideae*) (Gould and Jansen, 1999) is needed to conclusively determine the ancestry of SM and SG.

DIVERSITY WITHIN AND AMONG POPULATIONS. The average genetic diversity as measured by H_S was low within species, ranging from 0.14 in SM to 0.156 in SG (Table 3). These findings deviated from the expectation that diversity within populations of SG would be lower than in SM as a result of the comparatively narrower geographic area inhabited by SG. The proportion of species genetic diversity attributed to among-population variation overall species and loci as measured by G_{ST} ($G_{ST} = 1 - H_S/H_T$) was 0.23 and among populations within each species ranged from 0.12 in SGG to 0.20 in SM (Table 3). The relatively low overall G_{ST} indicates that a low proportion of diversity is observed among populations as opposed to a high level of diversity observed within populations using AMOVA (80.6%) (Table 4). Low G_{ST} values also indicate a high level of gene flow among populations, which tends to homogenize a species' genetic structure. SM, SGA, and SGG are pollinated by insects, including bees, moths, butterflies, ants, and beetles as well as hummingbirds (Affolter, 2005; Cullina, 2000; Rogers, 1988). Similar results have been observed in other outcrossing species. Shrestha et al. (2005) evaluated AFLP markers in nine populations of *Tectona*. Their AMOVA results showed that 57% of the total genetic variance occurred within populations, and 43% occurred between populations. In the outcrossing species hardinggrass (*Phalaris aquatica*), Mian et al. (2005) evaluated genetic diversity of 22 populations and found that variance within populations accounted for 74.1% of the total genetic diversity. AFLP studies of 15 populations of wild banana (*Musa balbisiana*) revealed that 72.9% of the variation was the result of within-population diversity (Wang et al., 2007). Similar results were found among populations of deciduous azalea; 71% of the diversity was within the populations (Chappell et al., 2008).

Insect pollination leads to populations with a high level of genetic variation, whereas individuals within the population share a similar complement of alleles in similar frequencies (Falk et al., 2001; Hamrick and Godt, 1996). The relatively low G_{ST} value combined with low proportion of variation among populations from AMOVA further indicates that individuals within populations are likely to be genetically distinct; however, each population contains a similar complement of alleles in similar frequencies. These results suggest that plant breeders seeking increased genetic diversity for their program may not obtain substantially different alleles from geographically isolated populations.

EVIDENCE OF INTROGRESSION BASED ON GENETIC DISTANCE COMPARISONS. Nei's unbiased genetic distances (Nei, 1978) among populations within species were similar to among species (Table 5), indicating that all populations evaluated are highly related and likely the result of high levels of gene flow. We used AFLP marker technology that provides markers without the need of sequencing. However, as a result of its

Table 4. Analysis of molecular variation for *Spigelia marilandica*, *S. gentianoides* var. *alabamensis*, and *S. gentianoides* var. *gentianoides* populations used in this study.

Source of variation	df	Sum of squares	Variance components	Variance (%)	P
Among all populations	12	957.77	6.16	19.43	<0.05
Within all populations	102	2,607.04	25.56	80.57	<0.05
Total	114	3,564.80	31.72		
Among species within botanical variety	2	315.31	5.26	15.34	<0.05
Within species within botanical variety	112	3,250.56	29.02	84.66	<0.05
Total	114	3,656.88	34.28		

Table 5. Nei's unbiased measures of genetic distance (Nei, 1978) below diagonal and geographic distance (kilometers) above diagonal for *Spigelia marilandica* (SM), *S. gentianoides* var. *alabamensis* (SGA), and *S. gentianoides* var. *gentianoides* (SGG) populations collected.

	1 ^z	2	3	4	5	6	7	8	9	10	11	12	13
1 ^z	***	227	536	406	259	373	364	88	676	615	303	504	546
2	0.052	***	381	225	182	156	145	166	483	422	325	394	380
3	0.020	0.053	***	504	562	288	290	505	697	129	341	116	36
4	0.035	0.051	0.035	***	179	240	232	323	285	496	547	557	494
5	0.031	0.052	0.031	0.035	***	303	291	192	441	576	475	575	560
6	0.030	0.048	0.032	0.018	0.037	***	13	320	452	291	388	343	282
7	0.035	0.030	0.029	0.043	0.034	0.039	***	309	449	295	380	341	282
8	0.029	0.051	0.023	0.039	0.028	0.044	0.033	***	594	571	333	501	515
9	0.053	0.083	0.068	0.082	0.050	0.067	0.078	0.072	***	639	805	768	676
10	0.044	0.072	0.044	0.076	0.043	0.071	0.050	0.039	0.068	***	460	235	101
11	0.052	0.073	0.048	0.087	0.063	0.074	0.063	0.049	0.076	0.061	***	267	364
12	0.06	0.084	0.059	0.079	0.066	0.081	0.065	0.047	0.104	0.062	0.076	***	142
13	0.056	0.060	0.043	0.076	0.057	0.068	0.050	0.057	0.091	0.066	0.068	0.052	***

^zPopulation numbers correspond to those listed in Table 2.

anonymous character, it cannot be unambiguously determined if the presence or absence of a band shows orthologous fragments of different genotypes. Although this approach has been used to provide a good estimate of genetic distance successfully (Chappell et al., 2008; Coulibaly et al., 2002; Wang et al., 2005), it cannot be excluded that closer or more distant relationships are caused by size homoplasy. Genetic distances were very low for all populations compared. Values ranged from 0.02 when comparing population 1 with population 3 (SM populations) to 0.10 when comparing population 9 with population 12 (SM and SGA populations).

When genetic distances were visualized as a UPGMA phenogram (Fig. 2), three groups were shown. Two individuals of population 12 (SGG) were placed within a group (hereafter, Group 1) at the base of the tree 100% of the time. Thirteen individuals from both populations 12 and 13 (SGG) were separated (hereafter, Group 2) from all of the SM and SGA samples as well as from six of the other SGG samples (hereafter, Group 3) with 91.8% confidence. Nodes with bootstrap values below 50% were collapsed and yielded polytomies (Perea et al., 2010). Bootstrap values of the nodes directly after the separation of Group 2 and Group 3 were 32.2% and 19.7%, respectively. These nodes, therefore, were collapsed and colored red indicating unresolved branches. These data suggest high levels of interspecific gene flow; however, many of the external branches have high levels of branch support. For example, although the node following Group 3 had only 32.2% bootstrap support, six individuals from population 12 were clustered together with 92.3% confidence. High levels of bootstrap values were found on some of the outermost branches of the tree (all of the following values are separate locations within the tree) including two individuals from population 4 (91.0%), two samples of population 8 (75.4%), three samples of population 2 (73.6%), and two samples of population 3 (98.9%). Although much of SGG makes up two separate groups, there are six SGG samples within Group 3, where all of the SM and SGA samples are located, indicating the occurrence of gene flow among all botanical varieties. However, SGG individuals appear to be diverging from SGA and SM.

POPULATION STRUCTURE ANALYSIS. Further illustration of gene flow among these populations is shown in Figure 3 from

STRUCTURE analysis. Six clusters (K) were chosen for analysis of individuals on the basis of highest likelihood values and convergence of alpha. These values show individual allelic contributions similar to that in PHYLIP, whereas previous PopGene analyses depict variability levels in accordance to populations as a whole.

Cluster 1 contained samples from populations 5, 8, and 10. Within this cluster, two samples contained 4% and 13% of their alleles from cluster 6. Both of these samples are from population 10 (SM) located in Wakulla Springs in the Florida Panhandle (Fig. 1). Cluster 6 contains only SGG samples from the Florida Panhandle and southeastern Alabama. These data suggest introgression from SGG populations into samples from population 10. Additionally, cluster 1 also contains alleles from clusters 2, 3, 4, and 5.

Cluster 2 contained samples from populations 1, 3, 4, 5, 6, 7, 8, 9, 11, 12, and 13, which further indicates gene flow among all botanical varieties used in this study. This cluster additionally contained alleles from clusters 1, 3, 4, 5, and 6.

Cluster 3 contained samples from populations 2 and 7 (SM populations) indicating particularly high levels of gene flow between those two populations. This cluster additionally contained alleles from clusters 1, 2, 4, 5, and 6.

Cluster 4 contained samples from populations 1, 3, 5, 7, 8, 11, and 12. Eight of 10 samples of population 11 (SGA) were placed in this cluster, all but one of which contained greater than 81.5% of its alleles from cluster 4. The other sample contained 62.7%. None of the SGA samples contained any alleles from cluster 6, a cluster that contains only SGG samples (populations 12 and 13). Furthermore, only one individual from population 12 and no individuals from population 13 were found in this cluster. This supports taxonomic separation of SGA from SGG. Samples from populations 1, 3, and 8 (SM populations) had an allelic distribution similar to many of the samples from population 11 (SGA), indicating that gene flow among these populations is exceptionally high. This cluster additionally contained alleles from clusters 1, 2, 3, 5, and 12.

Cluster 5 contained samples from populations 1, 3, 5, 6, 7, 9, 10, and 11. This cluster additionally contained alleles from clusters 1, 2, 3, 4, and 6. It should be noted that, although cluster 6 was represented, only one sample from the aforementioned

Fig. 2. The unrooted unweighted pair group method with arithmetic mean (UPGMA) phenogram of Nei's unbiased genetic distance matrix (Nei, 1978) overall 13 populations of *Spigelia* surveyed with an indication of bootstrap values of 1000. Numbers listed at the end of each branch indicate population sites; populations 1 to 10 = *S. marilandica*, population 11 = *S. gentianoides* var. *alabamensis*, populations 12 and 13 = *S. gentianoides* var. *gentianoides*. Letters designate individuals collected within each population site. Branches colored red indicate bootstrap values below 50%.

population 10 (SM; Wakulla Springs, FL) contained any alleles from this cluster.

Cluster 6 contained samples from populations 12 and 13, both SGG populations. No samples from SM or SGA were located in this cluster, indicating the unique alleles of SGG. These data provide support for taxonomic separation of SGG from SGA and SM (Weakley et al., 2011). Nine of 15 samples

in this cluster had greater than 89.8% of allelic contribution from cluster 6. The other individuals ranged from 40% to 72% allelic contribution from cluster 6.

In summary, AFLP markers exhibited a high level of efficiency in detecting genetic variation within *Spigelia*. Although AFLP analysis evaluates multilocus fragments, the sequences and locations of genes responsible for such distinctive phenotypic

Fig. 3. Inferred *Spigelia* population structure based on 116 individuals and 269 markers, assuming correlations among allele frequencies across clusters. Individuals are arranged by allelic distribution in each cluster ($K = 6$). Each vertical line represents one individual and the colors represent the membership coefficients to the K clusters. $K = 6$ clusters. Cluster 1 = red; cluster 2 = green; cluster 3 = blue; cluster 4 = yellow; cluster 5 = pink; 6 = aqua. Numbers preceding letters = population location described in Table 2. Letters following numbers indicate individual plants collected within each population; populations 1 to 10 = *S. marilandica*, population 11 = *S. gentianoides* var. *alabamensis*, population 12 and 13 = *S. gentianoides* var. *gentianoides*.

differences among these species as flower color, flower number, or leaf shape in *Spigelia* are not yet known. The results of this study indicate that members of *Spigelia* species, SM and SG, are genetically similar both among and within species, despite occupying differing ecological environments. Assessing the actual (not proportional) variation of individual species demonstrates that within-population and among-population variation is relatively low. The proportion of variation among species and among populations of each species is also low with the greatest proportion of genetic variation residing within individual populations. Therefore, for purposes of obtaining variability for plant breeding or to preserve the genetic variability in *Spigelia*, collection or preservation of plant material from multiple geographically dispersed populations may not be critical. However, population structure analysis revealed that all allele clusters have not entirely introgressed into all populations, and some populations have more allelic diversity than others such that breeding success in *Spigelia* may be impacted by individual collections.

Evaluation of population structure revealed subtle population differences and delineated how gene flow occurs through the populations. Some samples in this study were shown to be almost entirely composed of alleles from specific clusters. For example, every sample from SGG contained alleles from cluster 6 with some samples having nearly 90% of their allelic makeup from this cluster. No samples from SGA and only three populations of SM had alleles from cluster 6, indicating considerable genetic differences between SGG and both SGA and SM. The levels of genetic diversity, coupled with its entomophilous outcrossing nature, suggest that gene flow both

among populations and among species (introgression) is important in maintaining the heterozygosity of populations and/or species. Fragmentation or isolation of populations resulting from human activity could have a significant and abrupt negative effect on the adaptability and success of individual populations and/or species.

Literature Cited

- Affolter, J.M. 2005. Conservation biology of *Spigelia gentianoides* and *S. marilandica*: Genetic variation, reproduction biology, and propagation. Final Project Rpt. Georgia Coop. Fish Wildlife Res. Unit, Univ. Georgia, Athens, GA.
- Baldwin, B.G., M.J. Sanderson, J.M. Porter, M.F. Wojciechowski, C.S. Campbell, and M.J. Donoghue. 1995. The ITS region of nuclear ribosomal DNA: A valuable source of evidence on angiosperm phylogeny. *Ann. Mo. Bot. Gard.* 82:247–277.
- Beedanagari, S.R., S.K. Dove, B.W. Wood, and P.J. Conner. 2005. A first linkage map of pecan cultivars based on RAPD and AFLP markers. *Theor. Appl. Genet.* 110:1127–1137.
- Cervantes-Flores, J.C., G.C. Yencho, A. Krieger, K. Pecota, M. Faulk, R. Mwanga, and B. Sosinski. 2008. Development of a genetic linkage map and identification of homologous linkage groups in sweetpotato using multiple-dose AFLP markers. *Mol. Breed.* 21:511–532.
- Chappell, M., C. Robacker, and T.M. Jenkins. 2008. Genetic diversity of seven deciduous azalea species (*Rhododendron* spp. Section *Pentanthera*) native to the eastern United States. *J. Amer. Soc. Hort. Sci.* 133:374–382.
- Chesson, P.L. 1985. Coexistence of competitors in spatially and temporally varying environments: A look at the combined effects of different sorts of variability. *Theor. Popul. Biol.* 28:263–287.
- Cohen, D. 1966. Optimising reproduction in a randomly varying environment. *J. Theor. Biol.* 12:119–129.

- Coulibaly, S., R.S. Pasquet, R. Papa, and P. Gepts. 2002. AFLP analysis of the phenetic organization and genetic diversity of *Vigna unguiculata* L. Walp. reveals extensive gene flow between wild and domesticated types. *Theor. Appl. Genet.* 104:358–366.
- Cullina, W. 2000. The New England Wild Flower Society guide to growing and propagating wildflowers of the United States and Canada: A guide to growing and propagating native flowers of North America. Houghton Mifflin Harcourt, Boston, MA.
- Duncan, W.H. and M.B. Duncan. 2005. Wildflowers of the eastern United States. Univ. Georgia Press, Athens, GA.
- Dunwell, W. 2003. *Spigelia marilandica* propagation: A review. *Intl. Plant Prop. Soc. Comb. Proc.* 53:510–512.
- Excoffier, L. and H.E.L. Lischer. 2010. Arlequin suite version 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resources* 10:564–567.
- Falk, D.A., E.E. Knapp, and E.O. Guerrant. 2001. An introduction to restoration genetics. *Soc. Ecol. Restoration*, Washington, DC.
- Falush, D., M. Stephens, and J.K. Pritchard. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.
- Felsenstein, J. 2009. PHYLIP Version 3.69. Univ. Washington, Seattle, WA.
- Gould, K. 1997. Systematic studies in *Spigelia*. Univ. Texas, Austin, TX.
- Gould, K.R. and R.K. Jansen. 1999. Taxonomy and phylogeny of a Gulf Coast disjunct group of *Spigelia* (Loganiaceae sensu lato). *Lundellia* (Austin, Tex.) 2:1–13.
- Hamrick, J. and M. Godt. 1996. Conservation genetics of endemic plant species, p. 281–304. In: Avise, J.C. and J.L. Hamrick (eds.). *Conservation genetics: Case histories from nature*. Chapman and Hall, New York, NY.
- Kim, K.J. and R.K. Jansen. 1994. Comparisons of phylogenetic hypotheses among different data sets in dwarf dandelions (*Krigia*, Asteraceae): Additional information from internal transcribed spacer sequences of nuclear ribosomal DNA. *Plant Syst. Evol.* 190:157–185.
- Koopman, W.J.M., M.J. Zevenbergen, and R.G. Van Den Berg. 2001. Species relationships in *Lactuca* S.L. (Lactuceae, Asteraceae) inferred from AFLP fingerprints. *Amer. J. Bot.* 88:1881–1887.
- Lubell, J.D., M.H. Brand, J.M. Lehrer, and K.E. Holsinger. 2009. Amplified fragment length polymorphism and parentage analysis of a feral barberry (*Berberis thunbergii* DC.) population to determine the contribution of an ornamental landscape genotype. *HortScience* 44:392–395.
- Mian, M.A.R., J.C. Zwonitzer, Y. Chen, M.C. Saha, and A.A. Hopkins. 2005. AFLP diversity within and among hardinggrass populations. *Crop Sci.* 5:2591–2597.
- Negron-Ortiz, V. 2012. Recovery plan for *Spigelia gentianoides* (gentian pinkroot). 21 Nov. 2014. <<http://www.fws.gov/panamacity/resources/Spigelia%20gentianoides%20Recovery%20Plan.pdf>>.
- Nei, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89:583–590.
- Nei, M. 1987. *Molecular evolutionary genetics*. Columbia Univ. Press, New York, NY.
- Page, R.D.M. 1996. TREEVIEW: An application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12:357–358.
- Perea, S., M. Bohme, P. Zupancic, J. Freyhof, R. Sanda, M. Ozulug, A. Abdoli, and I. Doadrio. 2010. Phylogenetic relationships and biogeographical patterns in Circum-Mediterranean subfamily Leuciscinae (Teleostei, Cyprinidae) inferred from both mitochondrial and nuclear data. *BMC Evol. Biol.* 10:265.
- Philips, R.L. and I.K. Vasil, (eds.). 2001. DNA-based markers in plants. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Pritchard, J.K., M. Stephens, and P.J. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.
- Rogers, G.K. 1988. *Spigelia gentianoides*—A species on the brink of extinction. *Plant Conservation* 3:1–8.
- Russell, J.R., J.D. Fuller, M. Macaulay, B.G. Hatz, A. Jahoor, W. Powell, and R. Waugh. 1997. Direct comparison of levels of genetic variation among barley accessions detected by RFLPs, AFLPs, SSRs and RAPDs. *Theor. Appl. Genet.* 95:714–722.
- Shrestha, M.K., H. Volkaert, and D. Van Der Straeten. 2005. Assessment of genetic diversity in *Tectona grandis* using amplified fragment length polymorphism markers. *Can. J. For. Res.* 35:1017–1022.
- Tilman, D. 1999. The ecological consequences of changes in biodiversity: A search for general principles. *Ecology* 80:1455–1474.
- Tuljapourkar, S. 1989. An uncertain life: Demography in random environments. *Theor. Popul. Biol.* 35:227–294.
- Vos, P., R. Hogers, M. Bleeker, M. Reijans, T. van de Lee, M. Hornes, A. Frijters, J. Pot, J. Peleman, M. Kuiper, and M. Zabeau. 1995. AFLP: A new technique for DNA fingerprinting. *Nucleic Acids Res.* 23:4407–4414.
- Wang, X.L., T.Y. Chiang, N. Roux, G. Hao, and X.J. Ge. 2007. Genetic diversity of wild banana (*Musa balbisiana* Colla) in China as revealed by AFLP markers. *Genet. Resources Crop Evol.* 54:1125–1132.
- Wang, Y., G.L. Reighard, D.R. Layne, A.B. Abbott, and H. Huang. 2005. Inheritance of AFLP markers and their use for genetic diversity analysis in wild and domesticated pawpaw [*Asimina triloba* (L.) Dunal]. *J. Amer. Soc. Hort. Sci.* 130:561–568.
- Weakley, A.S., B.A. Sorrie, R.J. LeBlond, B.A. Sorrie, C.T. Witsell, L.D. Estes, K. Gandhi, K.G. Mathews, and A. Ebihara. 2011. New combinations, rank changes, and nomenclatural and taxonomic comments in the vascular flora of the southeastern United States. *J. Bot. Res. Inst. Texas* 5:437–455.
- Yeh, F.C., T.B.J. Boyle, Z.-H. Ye, and J.X. Mao. 1997. POPGENE, the user-friendly shareware for population genetic analysis. *Mol. Biol. Biotechnol. Ctr.*, Univ. Alberta, Edmonton, Alberta, Canada.
- Zerr, T. and S. Henikoff. 2005. Automated band mapping in electrophoretic gel images using background information. *Nucleic Acids Res.* 33:2806–2812.
- Zhang, F., S. Chen, F. Chen, W. Fang, and F. Li. 2010. A preliminary genetic linkage map of chrysanthemum (*Chrysanthemum morifolium*) cultivars using RAPD, ISSR and AFLP markers. *Sci. Hort.* 125:422–428.