Conserv Genet
DOI 10.1007/s10592-010-0095-7

RESEARCH ARTICLE

Conservation genetic inferences in the carnivorous pitcher plant *Sarracenia alata* (Sarraceniaceae)

Margaret M. Koopman · Bryan C. Carstens

Received: 16 September 2009 / Accepted: 2 May 2010

Abstract Conservation geneticists make inferences about their focal species from genetic data, and then use these inferences to inform conservation decisions. Since different biological processes can produce similar patterns of genetic diversity, we advocate an approach to data analysis that considers the full range of evolutionary forces and attempts to evaluate their relative contributions in an objective manner. Here we collect data from microsatellites and chloroplast loci and use these data to explore models of historical demography in the carnivorous Pitcher Plant, *Sarracenia alata*. Findings indicate that populations of *S. alata* exhibit high degrees of population genetic structure, likely caused by dispersal limitation, and that population sizes have decreased in western populations and increased in eastern populations. These results provide new insight to the management and conservation of plants restricted to small, declining populations isolated in increasingly scarce and highly threatened habitat, including other rare and endangered species of *Sarracenia*.

Keywords Information theory · Historical demography · Pine savannah · Phylogeography · *Sarracenia*

Introduction

Members of the carnivorous Pitcher Plant genus *Sarracenia* from the Southeastern United States are particularly vulnerable to increased competition following fire suppression (Schnell 1976; Weiss 1980; Folkerts 1982; Barker and Williamson 1988; Brewer 2005), habitat fragmentation and degradation as well as overcollecting (CITES May 2009, IUCN Red List). Among the most important habitats for *Sarracenia* in this region are the Longleaf pine savannahs, herb-dominated wetlands that were once extensive throughout the Gulf Coast of North America, but have been severely reduced in the last several hundred years. In Louisiana only 1–3% of the original Longleaf pine savannahs remain (Conservation Habitat Species Assessment 16, 36). *Sarracenia alata* is locally abundant (CITES: LR (nt)) throughout these habitats in Louisiana but remains vulnerable to all the pressures that jeopardize more rare and critically endangered Pitcher Plant species, like *S. oreophila* (CITES: CR (B1,2b,c)), throughout the Southeast. Here, we investigate the population genetic structure of *S. alata* and use this species as a model to understand how different evolutionary processes shape genetic diversity among isolated populations of Pitcher Plants as a means to further inform conservation management strategies for all taxa in this charismatic genus.

Sarracenia is a convenient system for genetic investigation because the plants occupy savannahs or bogs that are bordered by habitats inhospitable to the plants. Given these habitat requirements, it is reasonable to expect that population genetic structure evolves quickly within species. Life history characteristics related to gene flow support this expectation, for example, while the mechanism of seed dispersal in the genus is unknown, the seeds are minute (~2 mm (Ellison 2001), move a short distance from the parent plant (average 12.8 cm; (Ellison and Parker 2002)) and have no ornamentation indicative of successful or frequent long distance dispersal (Schnell 1976; Godt and Hamrick 1998; Ellison and Parker 2002). In *Sarracenia*

M. M. Koopman · B. C. Carstens (✉)
Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA
e-mail: carstens@lsu.edu

Published online: 19 May 2010

species where pollination observation has been conducted, flowers are primarily pollinated by bees (Schnell 1976; Slack 1979; O'Neil 1983), and their foraging range is generally estimated to be less than 2 km (Visscher and Seeley 1982; Breed et al. 1999; Kreyer et al. 2004; Osborne et al. 2008).

Sarracenia alata is restricted to pine savannahs and seepage bogs between extreme eastern Texas and southwestern Alabama (Wherry 1929; McDaniel 1971; Bayer et al. 1996) (Fig. 1), though the plants are much less common than implied by the distribution map because of habitat restrictions. *S. alata* is the only *Sarracenia* species west of the Mississippi river (MacRoberts and MacRoberts 1991) and most populations in Louisiana are monotypic for *S. alata*, greatly reducing the potential that interspecific hybridization will confound results in this study. *S. alata* is absent from the Mississippi River flood plains, due to of a lack of suitable habitat (Sheridan 1991), and this large river has been identified as a major barrier to gene flow in other organisms (Near et al. 2001; Al-Rabab'ah and Williams 2002; Soltis et al. 2006; Burbrink et al. 2008). While populations of *S. alata* are separated by up to 200 km, no morphological traits have been identified that distinguish eastern individuals from western individuals (Sheridan 1991).

Habitat fragmentation has dramatic implications for conservation (Haila 2002; Hobbs and Yates 2003; Foley et al. 2005; Fischer and Lindenmayer 2007), in large part due to the loss of genetic diversity and increased risk of local extinction. Fragmentation should produce populations of plants that are far more genetically structured through processes such as reduced gene flow and increased homozygosity via genetic drift. However, there could be other evolutionary forces that are important to this system. An important (yet unresolved) question for conservation genetics is how to make inferences in a manner that does not introduce bias into the results. Psychologists are aware of confirmation bias (Nickerson 1998), the tendency to interpret novel information in a manner that is consistent with preexisting beliefs, and we suggest that a related bias could occur while conducting genetic data analyses. For example, in the *S. alata* system we have *a priori* expectations that populations are genetically structured and that rates of gene flow are low, so it is reasonable to utilize programs that estimate these parameters. If other evolutionary processes have influenced genetic diversity (such as selection or demographic events like population size change), then limiting our analyses to the parameters that we suspect are important in *S. alata* might cause us to fail

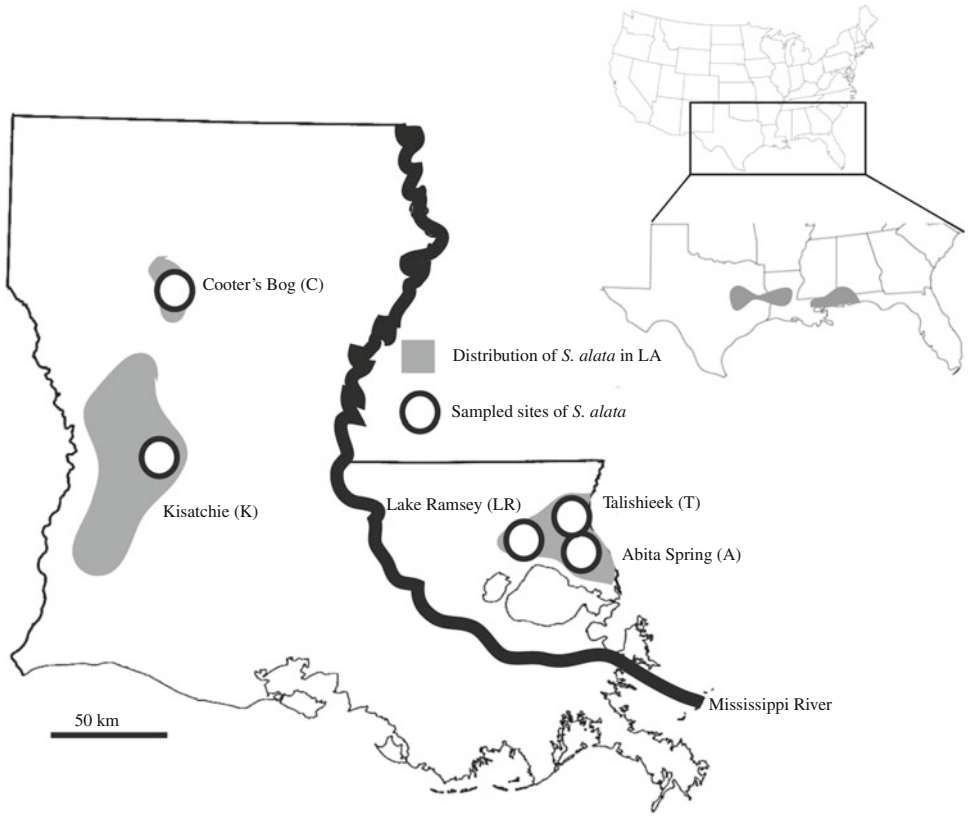

Fig. 1 Distribution of *Sarracenia alata* in the US (USDA PLANTS profile) and the distribution of the species in the state of Louisiana (MacRoberts and MacRoberts 1991; USDA PLANTS profile); population locales sampled in this study are also highlighted

not only to identify the important processes but also to ascertain the best conservation plan.

Methods

Genetic markers

We collected leaf tissue from 86 individual *Sarracenia alata* plants from five populations throughout Louisiana (Table 1), including three populations east of the Mississippi River, and two populations west of the river (Fig. 1). The number of individuals per population ranged from 10 to 27 (mean = 17). Sequence data and microsatellites were both utilized to sample intraspecific genetic variation. Microsatellites have become a favored marker for population level studies because of their high level of polymorphism (Kliman et al. 2000; Machado and Hey 2003). Maternally inherited chloroplast DNA, while less variable, will provide useful information regarding the historical structure of populations.

DNA was extracted from silica-dried leaf tissues using the DNeasy plant extraction kit (Quiagen, Valencia, CA). We amplified six regions of the chloroplast: *rps*16-*trn*K, *atp*I-*atp*H, *trn*C-*ycf*6, *ycf*6-*psb*M, *trn*H-*psb*A, and *rpl*16 following established protocols (Shaw et al. 2005, 2007) and assessed sequence variability using an initial screening set composed of two individuals from each of the five populations. Sequencing reactions were cleaned with an ethanol precipitation and visualized using an ABI PRISM® 3100, and sequences were aligned by hand. Eight microsatellite loci were amplified using primers and reaction conditions outlined previously (Koopman et al. 2009). Alleles were determined for each locus using GENEMAPPER version 4.0 (ABI, Foster City, CA).

Data analysis

Preliminary

Estimates of haplotype and nucleotide diversity from the chloroplast data were obtained using the program DNAsp (Librado and Rozas 2009). Models of molecular evolution for sequence data were selected using DT-ModSel (Minin et al. 2003). The gene tree for the *rps*16-*trn*K region (the only variable chloroplast region among those screened) was estimated with maximum likelihood (ML) as an optimality criterion using PAUP* 4.0 (Swofford 2001). A random-addition tree was used as the starting point, and tree space was explored heuristically using tree bisection and reconnection (with ten trees held at each step). Nodal support was assessed using 10,000 nonparametric bootstrapping replicates (Felsenstein 1985).

Microsatellite data were initially analyzed using Genepop (Raymond and Rousset 1995) to conduct several standard population genetic analyses, including the extent of genetic diversity, tests of locus and population level deviations from Hardy–Weinberg Equilibrium (HWE) as well as estimates of F_{IS} (Weir and Cockerham 1984) and F_{ST} (Wright 1969). Within each population the frequency of null alleles was determined and adjusted accordingly in Microchecker (Van Oosterhout et al. 2004).

Population genetic structure

Several methods were employed to identify population genetic structure with the microsatellite data. An analysis of molecular variance (AMOVA) within and among discrete populations was performed in Geno (Dyer 2009) for all individuals. Separate AMOVA's were performed under several models of population delineation (i.e., different

Table 1 Summary statistics for microsatellite data were calculated in GENEPOP: N_A number of alleles, H_O observed, and H_E expected heterozygosity under HWE, F_{IS} inbreeding coefficient

	LR				AS				T				C				K			
Locus	N_A	H_O	H_E	F_{IS}	N_A	H_O	H_E	F_{IS}	N_A	H_O	H_E	F_{IS}	N_A	H_O	H_E	F_{IS}	N_A	H_O	H_E	F_{IS}
5	7	0.54	0.70	0.23	5	0.46	0.68	0.33	3	0.10	0.47	0.80	4	0.46	0.49	0.07	4	0.57	0.72	0.21
7	5	0.39	0.45	0.14	3	0.55	0.62	−0.13	3	0.20	0.19	−0.03	2	0.00	0.17	1.00	2	0.07	0.19	0.64
18	5	0.62	0.75	0.18	4	0.70	0.62	0.12	4	0.60	0.58	−0.03	4	0.36	0.39	0.06	3	0.29	0.54	0.48
21	13	1.00	0.90	−0.11	8	0.54	0.68	0.21	8	0.80	0.71	−0.13	9	0.64	0.69	0.08	4	0.79	0.66	−0.19
27	2	0.00	0.36	1.00	1	–	–	–	1	–	–	–	3	0.14	0.13	−0.04	3	0.14	0.34	0.59
36	1	–	–	–	1	–	–	–	1	–	–	–	1	–	–	–	1	–	–	–
44	3	0.62	0.55	−0.13	2	0.46	0.34	−0.39	3	0.40	0.54	0.27	4	0.68	0.49	−0.42	3	0.64	0.45	−0.47
47	3	0.34	0.34	0.06	3	0.15	0.27	0.44	4	0.60	0.64	0.06	2	0.14	0.13	−0.08	4	0.57	0.60	0.05

Populations abbreviated as follows: Lake Ramsey (*LR*, n = 27), Abita Springs (*AS*, n = 13), Talisheek (*T*, n = 10), Cooter's Bog (*C*, n = 22), Kisatchie National Forest (*K*, n = 14)

values of K); including one that treated all populations as separate, one that divided samples into an eastern and western group, and one that collapsed the two most proximal eastern populations. We also used the Bayesian clustering method implemented in the program STRUCTURE 2.2 (Pritchard et al. 2000; Pritchard and Wen 2004) to infer patterns of population genetic structure. To calculate the most probable number of clusters, a model that partitioned the data in a range between one and six clusters (e.g., K = 1–6) was used, with each K iterated five times. All STRUCTURE analyses were conducted with 250,000 steps in the Markov-Chain and a burn-in of 100,000 steps. K was also estimated in STRUCTURAMA (Huelsenbeck et al. 2007) which clusters individuals into K populations such that HWE is maximized within populations and allows the placement of a prior distribution on K in order to permit the data to determine the most appropriate value. Settings were as follows: model Numpops = rv, ExpectedPriorNumPops = rv, shape = 2.5, scale = 0.5, mcmc ngen = 250000, nchains = 4, samplefreq = 25, printfreq = 500.

The life history characteristics of *S. alata* strongly suggest that population genetic structure will be present, but we are also interested in identifying the evolutionary forces that contributed to the formation of this structure. The most obvious explanation would simply be genetic drift, and this would predict that genetic differentiation accumulates as a function of geographic distance. Thus, isolation by distance (IBD) was calculated using the IBD Web Service (Jensen et al. 2005), and the significance of the regression of genetic distance on geographic distance between sample pairs was tested using a Mantel test (Mantel 1967) with 10,000 permutations. However, since we seek to avoid confirmation bias, and are open to the possibility that other forces have contributed to the evolution of *S. alata*, we also estimated a variety of other parameters.

Specific evolutionary forces

Selection

Microsatellite genetic differentiation among populations is thought to be the result of genetic drift, migration and mutation because the repeats are generally thought to be restricted to noncoding regions and it is assumed that selection has little or no effect on microsatellite differentiation. However, microsatellites may be genetically linked to genes under selection, and indeed can be used to conduct genomic scans for adaptive divergence (Storz 2005). Consequently, two tests were used to identify deviations from neutral evolution, the sign test for heterozygosity excess implemented in BOTTLENECK 1.2.02 (Cornuet and Luikart 1996) and the Lewtonin-Krakauer (LK) test (Lewontin and Krakauer 1973). Significance of the former was evaluated using simulations in BOTTLENECK for three sample partitioning schemes (as explored with STRUCTURE). This test was performed under the step-wise mutation model (SMM) and the two-phase mutation model (TPM) with one-step mutations accounting for 90% of the total because these models are considered to be realistic microsatellite mutation models (Ellegren 2000). This test compares expected heterozygosity estimated from allele frequencies with heterozygosity estimated from the sample size and number of alleles. In a population that has undergone recent population reduction the heterozygosity measured at a locus will exceed the heterozygosity computed from the number of sampled alleles (Watterson 1984) because alleles will be lost faster than heterozygosity. The LK test investigates allele frequency differences between populations; loci with different selectively favored alleles will exhibit larger allele frequency differences between populations than do loci with no selective pressures. Using simulations of allele frequency distributions, they suggest a metric to ascertain which loci have evidence of selection and which have been governed solely by drift. We restricted our tests to pairwise comparisons because allele frequencies could be correlated among populations and therefore lead to F_{ST} estimates with inflated variances (Baer 1999; Beaumont 2005).

Gene flow

Gene flow among sampled populations and genetic diversity ($\theta = 4N_e\mu$) from each population were estimated using MIGRATE-N, a method that implements a coalescent model of gene flow among multiple populations at equilibrium (Beerli and Felsenstein 2001; Beerli 2006). Parameters were estimated under a ladder model with 4 adaptively-heated chains (start temperatures of 1.0, 1.5, 3.0 and 6.0) and a swap-interval of ten. Results across three replicate runs were compared to ensure that the Markov Chains reached a stationary region of parameter space. Since the life history characteristics of *Sarracenia* suggest that seed dispersal is a limiting factor of dispersal, we also calculated the significance of migration-by-distance using an approach similar to that of the IBD (e.g., we replaced the genetic distance matrix from above with a migration matrix).

Population size change

Departures from neutrality and demographic processes, like population expansion or bottlenecks, may similarly produce an excess of alleles at low frequencies and tests for selection could be significant under either force (Fu 1997; Hahn et al. 2002). Given the history of the pine savannahs, we sought to explore population size change using LAMARC

(Kuhner 2006) by estimating gene flow and θ both with and without population size change (exponential growth parameter g). Maximum likelihood estimates of parameters were obtained by sampling every 20 trees over 10 chains of 20,000 generations each with the first 1,000 trees discarded as burn-in. Two replicate runs were performed with and without g to assess convergence of the MCMC chains.

Model selection

Each of the methods above allow users to estimate one or more summary statistics or parameters and make subsequent inferences regarding the importance of the biological process represented by the parameter on the basis of the magnitude of the estimate. However, it can be difficult to interpret these values in a qualitative manner, for example what value of estimated migration among any pair of populations would be sufficient to convince us that this process is biologically important? More directly, it may be difficult to infer *which* process contributed to the formation of a given pattern in the data. For example, both population expansion and natural selection can produce similar patterns of summary statistics such as Tajima's D (Hahn et al. 2002), and both gene flow and incomplete lineage sorting can result in polymorphism shared across populations (Slatkin 1981). As such we find it desirable, where possible, to differentiate among these processes using a formal statistical framework that estimates multiple parameters. For example, gene flow and incomplete sorting of ancestral polymorphism can be evaluated simultaneously using implementations of the isolation-with-migration model (Nielsen and Wakeley 2001; Hey and Nielsen 2004). Here we use IMa (Hey and Nielsen 2007) to estimate gene flow, $\theta = 4N_e\mu$, and population divergence for the eastern and western populations of *S. alata*. A significant advantage of IMa over its precursors is the capacity to conduct model selection that can reduce the full model (e.g., θ_A θ_1 θ_2 m_{12} m_{21} t) to one with a smaller set of parameters. We modified the model selection protocol in IMa using a previously described approach (Carstens et al. 2009). Briefly, this method uses information theory to quantify the probability of multiple models given the data, thereby identifying the subset of parameters that represent important biological processes. For example, if all shared polymorphism is best explained by incomplete lineage sorting, the selected models would not include migration parameters. We calculated AIC scores (Akaike 1973), AIC differences (Δ_i), Akaike weights (w_i) and evidence ratios (Anderson 2008). We explored a variety of run conditions in IMa, and chose a geometric heating scheme ($-g1 = 0.7$, $-g2 = 0.9$) with 20 coupled Markov chains, and four separate runs with a total run time of 480 h on a 3.0 GHz Intel Mac Pro. We also explored a variety of values for the parameter priors before settling on $\theta_A = 15$, $\theta_1 = 10$, $\theta_2 = 10$, $m_{12} = 5$, $m_{21} = 5$, and $t = 10$.

Results

One sampled locus in the chloroplast genome was variable in *Sarracenia alata*. The *trn*K-*rps*16 gene contained six variable sites in 527 base pairs; these sequences have been deposited in GenBank (HM159479–HM159551). We selected the F81 model of sequence evolution using DT-ModSel. Haplotypes ranged from 1 to 3 within populations; corresponding values of haplotype (H) (0–0.95) and nucleotide (π) (0–0.00058) diversity were small. The ML estimate of the genealogy is suggestive of population genetic structure; alleles sampled from populations to the east of the Mississippi form a monophyletic clade that is distinct from alleles sampled in the west (Fig. 2a). The remaining five chloroplast regions were without polymorphism and not used further in this study (single representatives/locus for the species have been submitted to GenBank, HM159474–HM159478).

Microsatellites were genotyped for eight loci from 86 individuals. The level of polymorphism was similar in eastern and western populations, and with 1–13 alleles (per locus) in the east and 1–9 in the west (Table 1). Observed heterozygosity ranged from 0 to 0.8 across both regions. Inbreeding coefficients (F_{IS}) were variable (−0.47–1.0), and differentiation estimates (F_{ST}: 0.11–0.36) were moderate (Table 2). Previous results had demonstrated that there is no evidence of deviation from HWE or linkage disequilibrium in the Lake Ramsey population (Koopman et al. 2009).

Population genetic structure

Analysis of molecular variance resulted in highly significant ($P < 0 = 0.001$) results regardless of how the data were partitioned (e.g., using the five sampled population clusters, the four population clusters indicated by Structure and Structurama, and dividing samples into eastern and western populations). Analyses suggest that the majority (71.5–83.1%) of the genetic diversity in *S. alata* is attributable to within-population differentiation among individuals (Table 3). Tests of isolation by distance indicate a significant positive correlation between geographic distance and genetic differentiation ($R^2 = 0.0560$, $P = 0.0001$).

Consistent with the AMOVA results, Bayesian simulations in Structure generally assigned individuals to clusters that corresponded to sampling sites (Fig. 2b). The exceptions are two clusters representing a pair of adjacent and closely spaced pine savannahs in the eastern distribution of the state (Abita Springs and Talisheek). The four clusters assigned by Structure are further supported by an

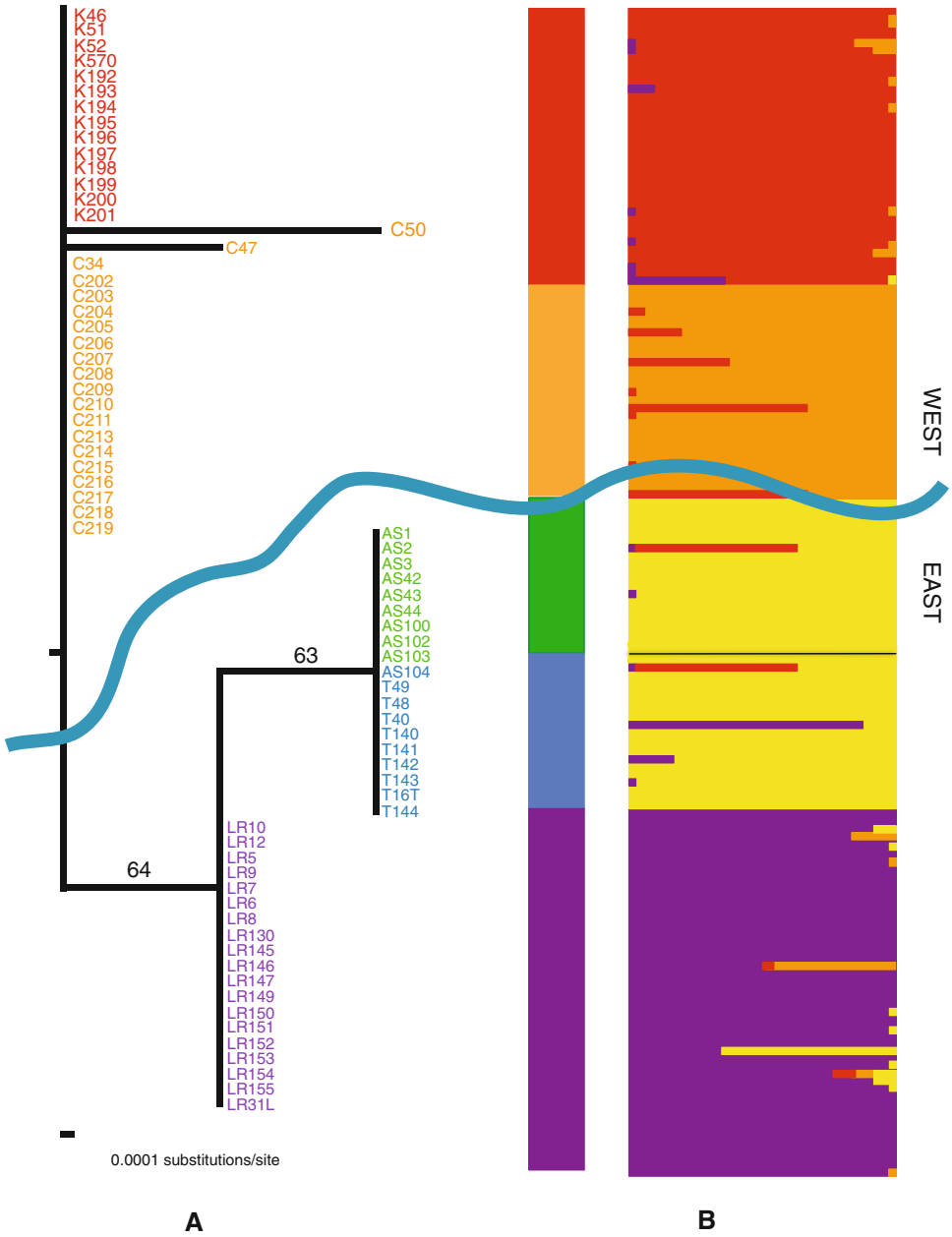

Fig. 2 **a** ML gene tree reconstructed from *rps*16-*trn*K sequence data. *Color* of individual is indicative of population origin. Values above branches are bootstrap values. **b** Bayesian cluster simulations performed in STRUCTURE. Each *vertical line* represents one individual and *color coding* shows the percentage of admixture that belongs to each genetic cluster (K). K = 2, across the Mississippi River, was determined when a ΔK statistic was used (Evanno et al. 2005). By calculating the value of K with the largest posterior probability, we found support for finer-scale clustering with K = 4, corresponding more closely to individual sample sites. For population abbreviations, see Table 1 (Color figure online)

Table 2 Estimates of *k* for pairwise genetic differentiation based on F_{ST} (Lewtonin–Krakauer test) are shown above the diagonal

	LR	AS	T	C	K
LR	–	2.83*	1.56	0.32	0.43
AS	0.3449	–	0.44	0.81	0.82
T	0.3129	0.1746	–	0.73	1.52
C	0.1931	0.1806	0.1397	–	0.66
K	0.1154	0.3555	0.2605	0.2471	–

Significance of *k* assessed by estimating expected value of 2 for neutral loci governed only by drift. Pairwise (average) estimates of genetic differentiation (F_{ST}), estimated using GENEPOP, are shown below the diagonal. Population abbreviations follow Table 1

Table 3 Analysis of molecular variance results under three models of population differentiation

	σA (among, %)	σB (within, %)	ΦST
k = 5	28.50	71.50	0.2847*
k = 4	27.30	72.70	0.2726*
k = 2	16.90	83.10	0.1694*

An *asterisk* denotes values that are significant at the $P \leq 0.001$ level

analysis in STRUCTURAMA which designates a K = 4 (Table 4). In contrast to the four clusters identified by STRUCTURAMA, the ΔK method of (Evanno et al. 2005)

Table 4 STRUCTURAMA results: the first column indicates the different values of K, the second indicates the posterior probability for each value of K, and the third column the prior probability for each value

K	Pr(K = i | X)	Pr(K = i)
1	0	0.0027
2	0.0002	0.0069
3	0.2867	0.0121
4	0.5112	0.0185
5	0.1665	0.0262
6	0.0301	0.0346
7	0.0048	0.0428
8	0.0004	0.0499

yielded K = 2 populations corresponding to populations east and west of the Mississippi River. Because Evanno's ΔK detects the highest level of genetic structure present (Evanno et al. 2005), a value of K = 2 does not invalidate the fine scale divergence reflected by STRUCTURAMA. Furthermore these results suggest the presence of hierarchical structure concordant to that observed in the plastid gene tree (Fig. 2a).

Selection

Statistics intended to detect selection were calculated. The sign test for heterozygote excess detected no significant departures from equilibrium for all five populations or for reduced population delineation (Abita Springs and Talisheek as one; not shown) nor for the eastern or western groups of populations (Table 5). Estimates of k for pairwise population comparisons using FST were generally low, with only one significant value, i.e., between Lake Ramsey and Abita Springs population, applying the $k > 2$ criterion of LK for inferring selection (Table 2). Furthermore, estimates of Tajima's D for the chloroplast sequence data were not significant (-0.1067; $P > 0.10$). Taken together, these results indicate that the genomic regions sampled in this investigation are not under strong selection.

Table 5 Sign tests for heterozygote excess conducted in BOTTLENECK (Cornuet & Luikart 1996)

	SMM		*TPM*	
Population	H_e/H_d	*P*	H_e/H_d	*P*
West	3/4	0.321	3/4	0.352
East	5/3	0.492	4/4	0.514

Number of loci with a significant heterozygosity excess and with a significant heterozygosity deficiency under two models of mutation. *SMM* stepwise mutation model, *TPM* two-phase mutation model

Gene flow

MIGRATE-N was used to estimate gene flow among the five populations. Maximum likelihood parameter estimates of relative migration rate for microsatellite data from MIGRATE-N ranged from $m = 0.903$, representing migrants shared across the Mississippi River, to $m = 26.47$ representing the migration rate between two populations in the east (Table 6). If migration was responsible for the population structure in *S. alata* then populations in close proximity should share more migrants than populations that are distantly spaced. For all populations there was no significant correlation between geographic distance and migration rate (MBD) based on the Mantel test ($R^2 = 0.103$, $P = 0.7492$). MIGRATE-N was also used to estimate gene flow among populations to the east and west of the Mississippi river. Measured migration rates are asymmetrical, rates of gene flow from eastern populations into the west are lower ($m = 8.835$) than estimates for migration in the other direction ($m = 20.203$) (Table 7).

Table 6 Estimates of two parameters from MIGRATE-N ANALYSIS

Comparison	CI_{low}	MLE	CI_{hi}
θ LR	0.334	0.422	0.534
θ AS	0.606	0.688	0.806
θ T	0.384	0.481	0.699
θ C	0.327	0.394	0.457
θ K	0.571	0.715	0.895
M AS-LR	22.732	26.477	30.691
M T-LR	12.773	15.639	19.047
M C-LR	7.344	9.619	12.210
M K-LR	4.052	5.699	7.827
M LR-AS	7.566	9.114	10.790
M T-AS	1.321	2.085	2.992
M C-AS	2.581	3.415	4.408
M K-AS	1.948	2.792	3.731
M LR-T	8.305	10.270	12.592
M AS-T	3.425	4.721	6.278
M C-T	1.009	1.528	2.476
M K-T	1.517	2.405	3.804
M LR-C	5.544	7.085	8.908
M AS-C	3.254	4.485	5.970
M T-C	0.428	0.903	1.737

For each of five populations, $\theta = 4N_e\mu$ is estimated. Pairwise estimates of gene flow among populations, averaged across two replicate MIGRATE-N runs, are shown. For each comparison, the 95% lower and upper confidence intervals are shown. Population abbreviations follow Table 1

Table 7 Estimates of θ and migration with and without growth from LAMARC and MIGRATE-N, as well as likelihood values reported in LAMARC from East to West pairwise comparisons

Analysis		0.005	MLE	0.995	$-lnL$
LAMARC, no growth	θ_{east}	0.048	0.052	0.056	
	θ_{west}	0.046	0.050	0.055	
	M_{east}–M_{west}	46.131	56.135	67.041	
	M_{west}–M_{east}	54.786	66.084	79.213	
					−50.81
LAMARC, growth	θ_{east}	0.006	0.018	0.046	
	θ_{west}	0.008	0.022	0.039	
	M_{east}–M_{west}	69.638	211.09	469.45	
	M_{west}–M_{east}	17.718	67.901	164.257	
	g_{east}	−51.995	44.122	121.909	
	g_{west}	−956.877	−230.816	−54.977	
					−20.65
MIGRATE-N	θ_{east}	0.852	0.959	1.056	
	θ_{west}	0.617	0.706	0.812	
	M_{east}–M_{west}	7.697	8.835	10.056	
	M_{west}–M_{east}	17.217	20.203	22.756	

Population size

The growth parameter, g, was estimated using LAMARC in order to assess whether there was evidence for historical population expansion or reduction. The growth estimate for the eastern populations is positive and indicates growth while the western population is negative and indicative of reduction in population size. While the likelihood values reported by LAMARC are approximate and can not be compared across runs (for example, with a likelihood ratio test (Kuhner 2006)), the magnitude of the growth parameter suggests that population size change has historically been an important demographic process in this system (Table 7, Fig. 3). Furthermore, the magnitude of change that might be expected in parameter estimates may increase as additional parameters (such as g) are introduced to the demographic model (Table 7).

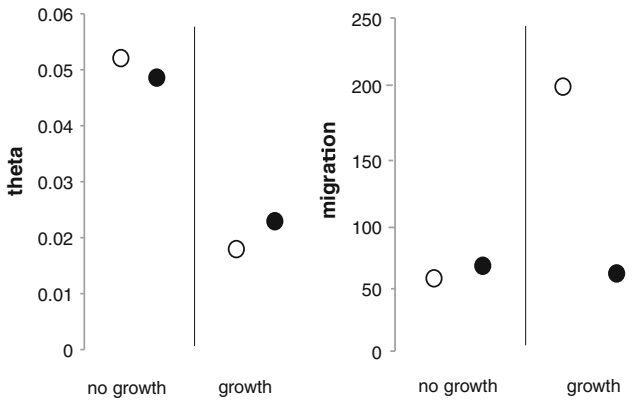

Fig. 3 Point estimates of theta and migration rate, with and without growth (g), from LAMARC averaged over two runs. *Open circles* indicate eastern populations while *closed circles* indicate western population estimates

Model selection

IMA was used to estimate θ, gene flow and population divergence. After experimenting with a range of priors, we choose values that resulted in posterior density functions that were both contained within the range of values specified by the priors and unimodal. All post-run metrics indicated that the Markov chains had reached stationarity and mixed well, in particular there was no trend to posterior plots of parameter estimates and effective sample sizes were large (e.g., >86). While we report values estimated from two runs that each lasted 20 days, we conducted several preliminary runs of shorter duration that resulted in similar estimates. Results from the IMA runs suggest that θ_E is approximately twice the value of

Table 8 Parameter estimation using IMA

	θ_E	θ_W	θ_A	m_{EW}	m_{WE}	τ
High point	4.6586	2.6456	2.7481	0.7975	0.0025	0.185
HPD90Lo	1.6227	1.5474	1.3348	0.0025	0.0025	0.095
HPD90Hi	9.0555	4.0432	145.4907	1.5275	0.5025	9.175

Shown are estimates of ancestral and descendant population θ (A ancestral, W western populations, E eastern populations), migrations rates from east to west and the reverse (m_{EW}, m_{WE}), and population divergence (τ). For each parameter, the high point of the posterior distribution is shown, as well as the boundaries of the 90% highest posterior density interval

Table 9 Results of model selection using IMA

Model	τ	θ_E	θ_W	θ_A	m_{EW}	m_{WE}	k	AIC	Δ_i	W_i
ABC00	0.4165	7.0147	2.8719	4.834	0.0001	0.0001	3	6.396	0	0.59852025
ABB00	0.2178	31.4389	2.5023	2.5023	0.0001	0.0001	2	7.5054	1.1094	0.197366096
ABCD0	0.4165	7.0022	2.8571	4.8296	0.000102	0.0001	4	8.3939	1.9979	0.081171188
ABCDD	0.4165	7.003	2.8704	4.8428	0.000103	0.000103	4	8.3954	1.9994	0.081049523
ABBDD	0.2178	31.4612	2.5036	2.5036	0.0001	0.0001	3	9.5057	3.1097	0.026702585
ABADD	0.571	6.9068	3.5045	6.9068	0.0113	0.0113	3	11.2175	4.8215	0.004820899
FULL	0.185	4.6586	2.6456	2.7481	0.7975	0.0025	5	11.4925	5.0965	0.00366182
ABBDE	0.2178	31.4791	2.5039	2.5039	0.000101	0.000104	4	11.5064	5.1104	0.003611273
ABC0D	0.2365	82.7498	2.9201	2.9425	0.0001	0.000107	4	12.2295	5.8335	0.001752355
ABADE	0.4515	5.0873	3.7935	5.0873	0.0689	0.000107	4	12.5145	6.1185	0.001317796
AACDE	1.0978	2.8049	2.8049	20.5143	0.85	0.000102	4	16.5126	10.1166	2.42E-05
ABA00	0.4165	6.5527	2.8632	6.5527	0.0001	0.0001	2	19.7581	13.3621	9.42E-07
AAADE	0.4373	4.2095	4.2095	4.2095	0.4106	0.1509	3	19.9051	13.5091	8.13E-07
AACDD	1.4988	2.6675	2.6675	18.7647	0.8913	0.8913	3	20.9774	14.5814	2.78E-07
AAADD	0.4271	4.2163	4.2163	4.2163	0.2917	0.2917	2	26.6608	20.2648	9.47E-10
AAA00	0.4515	4.6222	4.6222	4.6222	0.0001	0.0001	1	95.0127	88.6167	1.96E-39
AAC00	0.4515	4.5959	4.5959	4.8009	0.0001	0.0001	2	96.7981	90.4021	3.28E-40

The values from the short run are shown. High point values for each parameter are listed, as well as the marginal likelihoods of each model given the data, number of parameters (k), AIC scores, AIC differences (Δ_i), and model probabilites (w_j). Information theoretic statistics were computed following (Burnham and Anderson 1998). Estimates of $m_1 = 0.0001$ and $m_2 = 0.0001$ are indistinguishable from zero

ancestral or western populations (Table 8); these results are consistent with the population expansion in the eastern populations indicated by the LAMARC results (Fig. 3). Secondly, estimates of gene flow are low in this analysis, in contrast to the results from MIGRATE-N (Table 6), suggesting that incompletely sorted ancestral polymorphism likely accounts for most of the shared polymorphism.

Akaike weights (w_i) indicate that nearly 80% of the total model likelihood is contributed by models that do not include gene flow (Table 9), but that do allow θ to differ among lineages (ABC00 and ABB00). Results from the demographic model selection suggests that gene flow is not an important parameter to include in the demographic model, but that the eastern and western populations should be allowed to differ in populations size (assuming that mutation rates are equivalent).

Discussion

Population genetic structure and the Mississippi River

The newly designed microsatellite markers used in this study provide the highest estimates of heterozygosity (0.34–0.5) ever observed in *Sarracenia* (Schwaegerle and Schaal 1979; Wang et al. 2004). Further, the data reveal population genetic subdivisions in *S. alata* that are largely congruent with the Mississippi River (at a large scale) and habitat boundaries (at a smaller scale). These results are supported by all analyses (e.g., AMOVAs conducted at several hierarchical levels, STRUCTURE and STRUCTURAMA; Tables 3, 4, Fig. 2). Note that while Evanno's ΔK (Evanno et al. 2005) suggested that the data are best partitioned into two populations corresponding to plant distributions east and west of the Mississippi River, the pattern observed in the various analyses of both MSAT and cpDNA gene tree (Fig. 2) data suggest that population structure is hierarchical, with the Mississippi river at the deepest level and habitat boundaries at a finer scale acting to promote population genetic structure.

Historical demography

In addition to population structure, the other dominant signal in our genetic data is population expansion. LAMARC analyses indicate that the populations on either side of the Mississippi river have contrasting growth parameters; specifically estimates of g indicate that populations in the west are contracting while the populations to the east are expanding (Fig. 3). This result is apparently not caused by natural selection (Tables 2, 5), which can lead to false inferences of population size change (Hahn et al. 2002). The inference of population contraction in the west is consistent with what is known about the pine savannah habitat, and reductions in the population sizes of *Sarracenia* have been implicated in the literature (*S. oreophila*/*S. jonesii* Godt and Hammrick 1998). In contrast, the inference of expansion in the east appears at odds with the

fragmentation of the pine savannahs, as well as the documented slow rates of intrinsic population growth in *Sarracenia* (Schwaegerle 1983). However, all three populations in the east have undergone restoration management with regular burning treatments since at least 1998, are now deemed high quality habitat and are home to between 22 and 27 rare plant species (Keddy et al. 2006).

Gene flow, model selection and confirmation bias

Gene flow was estimated using three methods (MIGRATE-N, LAMARC, IMA) that implement coalescent models that include parameters for migration and genetic diversity (θ). The most dramatic differences among these methods are related to the other parameters that are also included in the model; essentially, each calculated the probability of some model given the data, however the models differ in their complexity. LAMARC incorporates a parameter for population growth (g), and IMA incorporates a parameter for temporal divergence (t). For the sake of simplicity, we restrict the discussion below to estimates made between eastern and western populations. Migration estimates with MIGRATE-N (Table 6) are higher than they are with LAMARC (Table 7) or IMA (Table 8); our interpretation is that this results from the inclusion of other parameters in the LAMARC or IMA model. For example, when population expansion of the eastern population or temporal divergence between populations is incorporated into the model (as they are in LAMARC and IMA, respectively), the shared polymorphism interpreted by MIGRATE-N as resulting from migration is attributed to other processes. Our skepticism of the relatively high migration estimates from MIGRATE-N is also related to the lack of a correlation between migration rates and geographic distance in the five-population analyses. These findings highlight the need for objective and assumption-free genetic data analyses.

We used information theory to compute the probability of multiple models encompassed by the full IMA model. Each submodel represents a hypothesis incorporating a set of historical processes, and by evaluating each of the models within IMA we were able to identify the set of parameters (historical processes) that are most probable given the data. The best models given the data are biologically similar and suggest three major inferences about recent historical demography in *S. alata*. We can conceptualize the total probability of the set of models given the data by the summed Akaike weights. For our data, nearly 80% of the total model probability is contributed by two models (ABC00 and ABB00; Table 9) that do not include migration, but that do allow populations sizes to differ among the east/west/ancestral lineages. Estimates of descendent θ are different to either side of the Mississippi River, supporting the inference that these populations have undergone fluctuations in size, however, the estimates of these parameters vary (Table 9). Estimates of population divergence times do not vary dramatically, and if we assume a mutation rate of 2.03×10^{-3}–4.96×10^{-5} mutations per generation (Thuillet et al. 2002; Vigouroux et al. 2002; Marriage et al. 2009), then divergence between the eastern and western populations occurred between 100 and 3700 generations ago, or well after the formation of the Mississippi River.

The lower Mississippi River valley was originally formed by the Mississippi Embayment, which filled the valley with sediment throughout the Jurassic period (Murray 1961). By the end of the Jurassic, the ancestral Mississippi Valley was well established and appeared as a narrow and shallow basin (Saucier 1994). More recently, Pleistocene glaciation and melting caused this valley to deepen and widen (Fisk 1944; Saucier 1994), and a phenomenon known as avulsion, whereby a river channel is abandoned for a new channel at a lower elevation, greatly impacted the course of the lower Mississippi River (Fisk 1947; Gomez et al. 1995). Avulsion has increased the width of the Mississippi River basin as bayous and swamps have formed. Today the Mississippi River, its abandoned distributaries (bayous) and the Atchafalaya swamp act to physically bisect the state of Louisiana (Saucier 1994).

So just how did pitcher plants, which are apparently poor dispersers, get across the river? One explanation is that the seeds are dispersed by floodwaters; however several facts contradict this idea. The seeds are hydrophobic (Ellison and Parker 2002), and *S. alata* does not occupy Mississippi river floodplains. Additionally, in *S. purpurea*, seeds are not found in the persistent seed bank in the bogs in which they live (unpublished data in (Ellison and Parker 2002)). A second explanation is that the meandering of the river, in conjunction with the formation of oxbow lakes, has allowed *S. alata* to become established on the opposite banks of the river. In this case seed and pollen would not disperse genetic material, instead the long-lived plants would find themselves inhabiting the opposite bank of the river via the process of oxbow lake formation.

Model selection and information theory in conservation genetics

Drastic changes in population size, population isolation and genetic variation can indicate that populations are vulnerable. However, these processes can be inferred using a variety of analytical programs, each of which incorporates parameters representing only a fraction of the evolutionary processes which may have influenced the focal taxa. We argue that conservation geneticists should (whenever possible) select the optimal demographic model using statistical approaches such as information theory. These

methods have the potential to transform genetics research and thus guide conservation programs and management strategies.

The information theoretic analysis performed in this study indicate that populations of *S. alata* on either side of the Mississippi river have contrasting growth parameters congruent with current and historic land use patterns in Louisiana's bogs and savannas; populations of *S. alata* in the east exist on highly controlled management areas (Keddy et al. 2006), are part of one of the largest complexes of pine savanna wetlands in the United States and appear to be increasing in size. In contrast, populations in the west are in decline, persist on multi-use Forest Service lands and are in poor condition (MacRoberts et al. 2001). Though the factors that maintain plant diversity in pine savannahs remain poorly understood (Keddy et al. 2006) our findings indicate that populations of *S. alata* are structured by changes in population size and that at least to some extent these changes are correlated with management strategies of pine savannah habitat in Louisiana. The aggressive management strategies in eastern Louisiana (including re-establishment of natural fire regimes) give insight for how to better manage *S. alata* habitat to the west of the Mississippi River but also could guide the future preservation of more rare and endangered species of *Sarracenia*.

Acknowledgments Funding was provided by the Louisiana State University Faculty Research Program and the Louisiana Board of Regents Research Competitiveness Subprogram (33825). We thank S. Hird for collecting *S. alata* material, V. Smith and M. Olinde for assistance with collection permits, and D. Fuselier and E. Gallagher for help in the lab. We thank L. Urbatch for helpful discussions related to this work.

References

Akaike H (1973) Information theory as an extension of the maximum likelihood principle. In Kotz S, Johnson NL (eds) Breakthroughs in statistics, vol 1. Springer-Verlag, London, pp 610–624

Al-Rabab'ah MA, Williams GC (2002) Population dynamics of *Pinus taeda* L. based on nuclear microsatellites. For Ecol Manag 163:263–271

Anderson DR (2008) Model based inference in the life sciences. Springer, New York

Baer CF (1999) Among-locus variation in Fst: fish, allozymes and the Lewontin–Krakauer test revisited. Genetics 152:653–659

Barker NG, Williamson GB (1988) Effects of a winter fire on *Sarracenia alata* and *S. psittacina*. Am J Bot 75:138–143

Bayer RJ, Hufford L, Soltis DE (1996) Phylogenetic relationships in Sarraceniaceae based on rbcL and ITS sequences. Syst Bot 21: 121–134

Beaumont MA (2005) Adaptation and speciation: what can Fst tell us? Trends Ecol Evol 20:435–440

Beerli P (2006) Comparison of Bayesian and maximum likelihood inference of population genetic parameters. Bioinformatics 22: 341–345

Beerli P, Felsenstein J (2001) Maximum likelihood estimation of a migration matrix and effective population sizes in n subpopulations using a coalescent approach. Proc Natl Acad Sci USA 98:4563–4568

Breed MD, McGlynn TP, Sanctuary MD, Stocker EM, Cruz R (1999) Distribution and abundance of colonies of selected meliponine species in a Costa Rican tropical wet forest. J Trop Ecol 15:765–777

Brewer JS (2005) The lack of favorable responses of an endangered pitcher plant to habitat restoration. Restor Ecol 13:710–717

Burbrink FT, Fontanella F, Pyron RA, Guiher TJ, Jimenez C (2008) Phylogeography across a continent: the evolutionary and demographic history of the North American Racer (*Serpentes Colubridae: Coluber constrictor*). Mol Phylogenet Evol 47:274–288

Burnham KP, Anderson DR (1998) Model Selection and inference: a practical information-theoretic approach. Springer, New York

Carstens BC, Stoute HN, Reid NM (2009) An information-theoretic approach to phylogeography. Mol Ecol 18:4270–4282

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144:2001–2014

Dyer R (2009) GeneticStudio: a suite of programs for spatial analysis of genetic-marker data. Mol Ecol Resour 9:110–113

Ellegren H (2000) Microsatellite mutations in the germline: implications for evolutionary inference. Trends Genet 16:551–558

Ellison AM, Gotelli NJ (2001) Evolutionary ecology of carnivorous plants. Trend Ecol Evol 16:623–629

Ellison AM, Parker JN (2002) Seed dispersal and seedling establishment of *Sarracenia purpurea* (Sarraceniaceae). Am J Bot 89: 1024–1026

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

Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39:783–791

Fischer J, Lindenmayer DB (2007) Landscape modification and habitat fragmentation: a synthesis. Glob Ecol Biogeogr 16:265–280

Fisk HN (1944) In: Engineers USACo (ed) Geological investigation of the alluvial valley of the lower Mississippi River. Mississippi River Commission, Vicksburg, MS

Fisk HN (1947) In: Engineers USACo (ed) Fine grained alluvial deposits and their effects on Mississippi River activity. Mississippi River Commission, Vickburg, MS

Foley JA, DeFries R, Asner GP, Barford C, Bonan G, Carpenter SR, Chapin FS, Coe MT, Daily GC, Gibbs HK, Helkowski JH, Holloway T, Howard EA, Kucharik CJ, Monfreda C, Patz JA, Prentice IC, Ramankutty N, Snyder PK (2005) Global consequences of land use. Science 309:570–574

Folkerts GW (1982) The Gulf Coast pitcher plant bogs. Am Sci 70:260–267

Fu YX (1997) Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147:915–925

Godt MJW, Hamrick JL (1998) Allozyme diversity in the endangered pitcher plant *Sarracenia rubra* ssp. *alabamensis* (Sarraceniaceae) and its close relative *S. rubra* ssp. *rubra*. Am J Bot 85:802–810

Gomez B, Mertes LAK, Phillips JD, Magilligan FJ, James LA (1995) Sediment characteristics of an extreme flood: 1993 upper Mississippi River valley. Geology 23:963–966

Hahn MW, Rausher MD, Cunningham CW (2002) Distinguishing between selection and population expansion in an experimental lineage of bacteriophage T7. Genetics 161:11–20

Haila Y (2002) A conceptual genealogy of fragmentation research: from island biogeography to landscape ecology. Ecol Appl 12: 321–334

Hey J, Nielsen R (2004) Multilocus methods for estimating population sizes, migration rates, and divergence times, with applications to the divergence of *Drosophila pseudoobscura* and *D. persimilis*. Genetics 167:747–760

Hey J, Nielsen R (2007) Integration within the Felsenstein equation or improved Markov chain Monte Carlo methods in population genetics. Proc Natl Acad Sci USA 104:2785–2790

Hobbs RJ, Yates CJ (2003) Impacts of ecosystem fragmentation on plant populations: generalising the idiosyncratic. Aust J Bot 51:471–488

Huelsenbeck JP, Huelsenbeck ET, Andolfatto P (2007) Structurama: Bayesian inference of population structure. Bioinformatics (in press)

Jensen JL, Bohonak AJ, Kelley STBGv (2005) Isolation by distance, web service. BMC Genet 6:13

Keddy PA, Smith L, Campbell DR, Clark M, Montz G (2006) Patterns of herbaceous plant diversity in southeastern Louisiana pine savannas. Appl Veg Sci 9:17–26

Kliman RM, Andolfatto P, Coyne JA, Depaulis F, Kreitman M, Berry AJ, McCarter J, Wakeley J, Hey J (2000) The population genetics of the origin and divergence of the *Drosophila simulans* complex species. Genetics 156:1913–1931

Koopman MM, Gallagher E, Carstens B, Consortium MERPD (2009) Isolation and characterization of nine microsatellite loci in the Pale Pitcher Plant *Sarracenia alata* (Sarraceniaceae). Mol Ecol Resour 9:1460–1466

Kreyer D, Oed AA, Walther-Hellwig K, Frankl R (2004) Are forests potential landscape barriers for foraging bublebees? Landscape scale experiments with *Bombus terrestris* agg. and *Bombus pascuaorum* (Hymenoptera, Apidae). Biol Conserv 116:111–118

Kuhner MK (2006) LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters. Bioinformatics 22:768–770

Lewontin RC, Krakauer J (1973) Distribution of gene frequency as a test of theory of selective neutrality of polymorphisms. Genetics 74:175–195

Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451–1452

Machado CA, Hey J (2003) The causes of phylogenetic conflict in a classic *Drosophila* species group. Proc R Soc Lond B 270:1193–1202

MacRoberts MH, MacRoberts BR (1991) The distribution of *Sarracenia* in Louisiana, with data on its abundance in the western part of the state. Phytologia 70:119–125

MacRoberts BR, MacRoberts MH, Stacey L (2001) Florisitcs and management of pitcher plant bogs in northern Natchitoches and Winn Parishes, Louisiana. Proc La Acad Sci 64:14–21

Mantel N (1967) The detection of disease clustering and a generalized regression approach. Cancer Res 27:209–220

Marriage TN, Hudman S, Mort EM, Orive ME, Shaw RG, Kelly JK (2009) Direct estimation of the mutation rate at dinucleotide microsatellite loci in *Arabidopsis thaliana* (Brassicaceae). Heredity 103:310–317

McDaniel S (1971) The genus *Sarracenia* (Sarraceniaceae). Bull Tall Timbers Res Stat 9:36

Minin V, Abdo Z, Joyce P, Sullivan J (2003) Performance-based selection of likelihood models for phylogeny estimation. Syst Biol 52:674–683

Murray GE (1961) Geology of the Atlantic and Gulf Coastal Province of North America. Harper & Brothers, New York, NY

Near NJ, Page LM, Mayden RL (2001) Intraspecific phylogeography of *Percina evides* (Percidae: Etheostomatinae): an additional test of the Central Highlands pre-Pleistocene vicariance hypothesis. Mol Ecol 10:2235–2240

Nickerson RS (1998) Confirmation bias: a ubiquitous phenomenon in many guises. Rev Gen Psychol 2:175–220

Nielsen R, Wakeley J (2001) Distinguishing migration from isolation: a markov chain monte carlo approach. Genetics 158:885–896

O'Neil W (1983) A preliminary reposrt on the pollination of *Sarracenia purpurea* in the forest-swale ecotone. Carniv Plant Newsl 12:60–62

Osborne JL, Martin AP, Carreck NL, Swain JL, Knight ME, Goulson D, Hale RJ, Sanderson RA (2008) Bumblebee flight distances in relation to the forage landscape. J Anim Ecol 77:406–415

Pritchard JK, Wen W (2004) Documentation for the STRUCTURE software Version 2. Chicago

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

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. J Hered 86:248–249

Saucier RT (1994) In: Engineers USACo (ed) Geomorphology and quaternary geologic history of the lower Mississippi Valley. Mississippi River Commission, Vicksburg, MS

Schnell DE (1976) Carnivorous plants of the United States and Canada, Second edn. Timber Press, Portland, Oregon

Schwaegerle KE (1983) Population growth of the pitcher plant, *Sarracenia purpurea* L., at Cranberry Bog, Licking County, Ohio. Ohio J Sci 1:19–22

Schwaegerle KE, Schaal BA (1979) Genetic variability and founder effect in the pitcher plant *Sarracenia purpurea*. Evolution 33:1210–1218

Shaw J, Lickey EB, Beck JT, Farmer SB, Liu W, Miller J, Siripun KC, Winder CT, Schilling EE, Small RL (2005) The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am J Bot 92:142–166

Shaw J, Lickey EB, Schilling EE, Small RL (2007) Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am J Bot 94:275–288

Sheridan PM (1991) What is the identity of the West Gulf Coastal pitcher plant, *Sarracenia alata?*. Carniv Plant Newsl 20:102–110

Slack A (1979) Carnivorous plants. MIT Press, Cambridge, MA

Slatkin M (1981) Estimating levels of gene flow in natural populations. Genetics 99:323–335

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

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. Mol Ecol 14:671–688

Swofford DL (2001) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts

Thuillet A, Bru D, David J, Roumet P, Santoni S, Sourdille P, Bataillon T (2002) Direct estimation of mutation rate for 10 microsatellite loci in Durum Wheat, *Triticum turgidum* (L.) Thell. ssp *durum*. Mol Biol Evol 19:122–125

Van Oosterhout C, HW F, WDP M, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in mirosatellite data. Mol Ecol Notes 4:535–538

Vigouroux Y, Jaqueth JS, Matsuoka Y, Smith OS, Beavis WD, Smith JSC, Doebley J (2002) Rate and Pattern of Mutation at Microsatellite Loci in Maize. Molecular Biology and Evolution 19:1251–1260

Visscher PK, Seeley TD (1982) Foraging strategy of honey bee colonies in a temperate deciduous forest. Ecology 63:1790–1801

Wang ZF, Hamrick JL, Godt MJW (2004) High genetic diversity in *Sarracenia leucophylla* (Sarraceniaceae), a carnivorous wetland herb. J Hered 95:234–243

Watterson GA (1984) Allele frequencies after a bottleneck. Theor Popul Biol 26:387–407

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370

Weiss TE (1980) The effects of fire and nutrient availability on the pitcher plant *Sarracenia flava* L. University of Georgia, Athens

Wherry ET (1929) Acidity relations of the *Sarracenias*. Wash Acad Sci 19:379–390

Wright S (1969) Evolution and the genetics of populations. University of Chicago Press, Chicago