POPULATION GENETICS OF THE RARE CEDAR GLADE ENDEMIC *ASTRAGALUS BIBULLATUS* (FABACEAE) USING NUCLEAR MICROSATELLITES

by

Kevin Trostel

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Master of Science in Biology

Middle Tennessee State University
August 2017

Thesis Committee:

Dr. Ashley B. Morris, Chair

Dr. Sarah E. Bergemann

Dr. Jeffrey L. Walck

ACKNOWLEDGEMENTS

This research was fully funded by U.S. Fish and Wildlife Service (USFWS) contract F13AC00084 to Dr. Ashley B. Morris. I thank my committee members Dr. Ashley Morris, Dr. Sarah Bergemann, and Dr. Jeffrey Walck. I thank Geoff Call (USFWS), Matthew Albrecht (Missouri Botanical Garden), and Quinn Long (Missouri Botanical Garden) for their input and guidance on this project.

ABSTRACT

Astragalus bibullatus is a federally endangered legume endemic to the limestone cedar glades of Rutherford County, Tennessee. Previous molecular studies have suggested that population structure is limited, and that the soil seed bank harbors more genetic diversity than does the standing vegetation. Current restoration efforts are underway, but success of reintroductions has varied across sites. The purpose of the current study was to use nuclear microsatellite loci to genotype individuals of *A. bibullatus* associated with long-term demographic data in an effort to better understand possible factors underlying reintroduction successes and failures. These results are consistent with previous studies in that most genetic variation is within individuals, not among sites. Future efforts will include detailed analyses of genetic data in the context of the demographic data to better inform reintroductions moving forward.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

Cedar glades are a rare ecosystem in the southeastern United States. Soil depth is shallow and varies across cedar glades, usually being deepest at the ecotone between the tree-less glade and surrounding forest. Cedar glades also contain pavements, fragments, or flagstones of limestone or dolomite rocks, scattered over the surface (Baskin and Baskin, 1999, 2003). Dominant canopy tree species in the forests surrounding cedar glades primarily include *Juniperus virginiana* L. (eastern red cedar; Cupressaceae) and various species of oaks and hickories (Baskin et al., 2007).

Astragalus bibullatus Barneby & Bridges (Fabaceae) is a perennial legume endemic to the limestone glades of Middle Tennessee. The legume family consists of more than 900 genera and nearly 24,000 recognized species (www.theplantlist.org). Of these 900 genera, *Astragalus* is the largest genus with over 3,000 species (Frodin, 2004). *Astragalus bibullatus* was first described in 1987 (Barneby and Bridges, 1987), and it was designated as a federally endangered species in 1991 (U.S. Fish and Wildlife Service, 1991). *Astragalus bibullatus* is a shade-intolerant species that occurs in open, rocky, or shallow soils (Gladeville and Talbot soil types) along forest edges of cedar glades (Baskauf and Snapp, 1998). The depth of the rooting zone is typically only a few centimeters, but can reach as deep as 30 cm. The plant has pod-like fruits that ripen to a papery consistency and drop immediately below the maternal plant. Pollination data are lacking; however, casual observations suggest that bees are the primary pollinators (Morris et al., 2002).

All known extant sites of *A. bibullatus* are located in Rutherford County, Tennessee (U.S. Fish and Wildlife Service, 2009). Residential expansion of Rutherford County has led to increased urgency for reintroduction efforts of this critically endangered plant (U.S. Fish and Wildlife Service, 2009). According to data reported by the US Census Bureau, as of July 1, 2015, Rutherford County had a population of 298,612 (www.census.gov). This reflects an increase of 13.7% from the previous census data reported April 1, 2010. Based on these current trends, the occurrences of *A. bibullatus* that are not on protected lands are at high risk of extirpation.

Prior studies revealed limited genetic diversity in *A. bibullatus*, suggesting that gene flow has decreased over time among populations (Baskauf and Snapp, 1998; Morris et al., 2002; Baskauf and Burke, 2009). Allozymes showed low levels of polymorphism across all populations and low estimates of genetic variation among sites (Baskauf and Snapp, 1998). Amplified fragment length polymorphisms (AFLPs) showed higher variation than previous allozyme studies as would be expected of these markers, but they also exhibited relatively low estimates of genetic diversity when compared to other plant species found in the cedar glades (Baskauf and Burke, 2009). An analysis of the soil seed bank using allozymes indicated more variation with increasing depth, suggesting that genetic variation in *A. bibullatus* has decreased over time (Morris et al., 2002).

Other endangered *Astragalus* species follow similar trends in terms of reduced genetic variability when fragmentation of populations is present (Neel, 2008). *Astragalus ampullarioides* (Fabaceae), a critically endangered species found in Utah with only seven known occurrences, was analyzed using 217 polymorphic AFLP loci (Breinholt et al.,

2009). Breinholt et al. (2009) concluded that *A. ampullarioides* historically experienced more gene flow, but habitat destruction by humans has resulted in fragmented populations with limited gene flow.

Astragalus albens is an endangered annual plant species occurring in southern California that has 91 known occurrences distributed in a 537 hectare area along a single network of carbonate substrate spanning a distance of 21,000 meters (Neel, 2008). These occurrences tend to be dense clusters of individuals located on rocky outcrops but no evidence of reduced genetic diversity was apparent from analyses of allelic diversity. No population structure was evident in populations as they exhibited extensive levels of gene flow. In contrast, populations of *A. nitidiflorus*, an endangered species endemic to southeast Spain, exhibited lower levels of genetic variation, likely as a result of a genetic bottleneck event caused by fragmentation of populations due to human disturbance (Vicente et al., 2011).

Astragalus bibullatus seeds are reportedly spread by gravity and water (Albrecht and Penagos, 2012). Seed dormancy and seed coat permeability play a role in the ability to replenish and establish seedling populations. Prolonged seed dormancy leads directly to slower recovery when the habitat undergoes major climate shifts or extreme weather patterns. Albrecht and Penagos (2012) conducted seed germination studies with *A. bibullatus* seeds collected from the soil seed bank. The results of their study indicated that germination rates were maximized when seeds went through scarification followed by incubation in dark, warm temperatures. The authors concluded that physical dormancy

is a major limitation to germination, and many years may be needed for seeds to break dormancy.

As summarized by Duncan et al. (2007), disturbance plays an important role in maintenance of biodiversity and ecosystem processes, and it is therefore a key consideration for conservation planning and restoration (Levey, 1988; Myers and van Lear, 1998). In cedar glades, one possible contributing factor to the decline in populations of *A. bibullatus* may be a lack of disturbance by fire. Duncan et al. (2007) performed the first known controlled fire study in the Ketona dolomite glades in Alabama, which are ecologically similar the limestone glades of Middle Tennessee where *A. bibullatus* occurs. While the study was short-term in nature, lasting only two years, the authors concluded that controlled fire occurrences had no negative impact on endemic herbaceous species. They also noted that invasive species and tree saplings declined in prevalence, although the decline was minimal. This could be due to the low overall temperatures present in the controlled study and the time of the applied burnings. The overall conclusions suggested that fire plays an important role in controlling young tree saplings along rocky outcrops of glade systems. Encroachment by woody species is expected to prevent growth of shade-intolerant herbaceous species if not otherwise controlled through fire or active management (Duncan et al. 2008).

Natural forest encroachment is thought to be a limiting factor controlling available habitat for *A. bibullatus* (Albrecht et al., 2016). Although sufficient studies have not been

conducted, controlled stressors such as burns should be considered in formulating reintroduction strategies for *A. bibullatus*.

The goal of the present study is to assess levels of genetic diversity and gene flow within and among naturally occurring populations of *A. bibullatus* using nuclear microsatellite loci. Microsatellite markers tend to be useful in studies documenting evolutionarily recent gene flow because of their high rates of polymorphism. These rates of polymorphism occur by replication slippage during DNA replication. This phenomenon leads to insertions or deletions of nucleotides. The resulting new base-pair sizes can be visualized, and levels of heterozygosity can be determined within and among populations.

Based on previous analyses of *A. bibullatus* and comparisons with other endangered *Astragalus* species, genetic diversity within populations was expected to be low due to limited gene flow among populations and low effective population sizes within populations. Any structure among populations may be a result of genetic drift. Demographic studies (reproductive and morphological) have been conducted over the last ten years (Albrecht et al., 2016), but such studies have yet to be integrated with analyses of the genetic diversity within these populations. The results from the current work, coupled with long-term demographic data for this species, should provide insight into developing management strategies for conservation of this species. A better understanding of reproductive patterns can be obtained with more detailed data on genetic diversity and population differentiation of the monitored plants within each population.

CHAPTER 2. MATERIALS AND METHODS

2.1. Sample collection

Astragalus bibullatus is known from 10 element occurrences (EOs) as classified by the Tennessee Department of Environment and Conservation (TDEC). These 10 EOs include two sites in which reintroductions have occurred (Couchville State Natural Area and Stones River National Battlefield) and eight naturally occurring sites, almost all of which are located to the east and southeast of Murfreesboro within a five-mile radius of each other. One exception to this cluster is an EO located northwest of Rockvale, approximately 15 miles west of the remaining naturally occurring populations.

Only plants that were permanently tagged as part of a long-term demographic study (Albrecht et al., 2016) were sampled for analyses of genetic diversity and population differentiation. A total of five populations were sampled, all of which occur within the cluster of populations east of Murfreesboro. Leaves were collected from two populations (Davenport East, $n = 30$; Davenport West, $n = 37$) in May 2013 and from three additional populations in April 2014 (Airport, $n = 51$; Manus, $n = 52$, Davis, $n = 23$). Each sample was placed in a silica gel packet labeled with the permanent tag number from Albrecht et al. (2016) and stored at room temperature in the Morris Lab in the Department of Biology at Middle Tennessee State University.

2.2. Microsatellite primer development

Microsatellite loci were developed for *A. bibullatus* using a next-generation sequencing approach (Morris et al., 2016). Genomic DNA of *A. bibullatus* was extracted

and shipped to Steve Bogdonawicz at the Evolutionary Core Genetics Facility (EGCF) at Cornell University (Ithaca, New York, USA). The DNA from one plant was used to develop and sequence microsatellite-enriched genomic libraries using 2 x 250 paired end sequencing on an Illumina MiSeq (San Diego, California, USA). Bogdonawicz mined the data for potential microsatellite loci using MSATCOMMANDER 1.0.3 (Faircloth, 2008). Data received from this service included a spreadsheet of more than 3000 potential primer pairs for the amplification of nuclear microsatellite loci identified in *A. bibullatus*. A total of 46 primer pairs were screened for potential use in this study. Seven of the loci were selected based on clarity of signal and polymorphism (i.e., more than one allele).

2.3. Genotyping

Genotyping methods followed those outlined in Morris et al. (2016). DNAs were extracted from all samples collected from *A. bibullatus* populations included in this study using the Qiagen DNeasy Plant Mini Kit (Qiagen, Valencia, California). Each of the seven nuclear microsatellite loci selected for this work were amplified in singleplex following the methods described in Morris et al. (2016). Reactions followed the three-primer approach of Schuelke (2000), with a 17-base tail (5'-GTAAAACGACGGCCAGT-3') added to the 5' end of each forward primer and a 7-base 'pigtail' (5'-GTTTCTT-3') added to the 5' end of each reverse primer. The third primer in each reaction was identical to the forward tail and was fluorescently labeled with one of three standard tags (FAM, VIC, or NED). Final reaction concentrations were as follows: 1× Platinum Taq buffer (Life Technologies), 2 mM $MgCl_2$, 0.5 μM forward primer with

the 5'- M13 tail, 0.15 μM fluorescently-labeled M13 primer, 0.2 μM pig-tailed reverse primer, 0.2 mM dNTPs, 0.5 U Platinum Taq (Life Technologies, Foster City, CA USA), and 1 μL DNA. A touchdown cycling protocol was used as described in Morris et al. (2016). PCR fragment sizes were sized with the ABI 3730 (Life Technologies, Foster City, CA USA) using the LIZ-500 size standard using the GeneMarker MTP software v. 2.6.0 (Softgenetics LLC, State College, PA USA).

2.4. Data analyses

Measures of genetic diversity were calculated using GenAlEx 6.5 (Peakall and Smouse, 2006, 2012). Summary statistics were calculated for each locus over all populations, as well as for each population over all loci. These statistics included number of alleles (N_a), effective number of alleles (N_e), observed heterozygosity (H_o), unbiased expected heterozygosity (H_e), and F_{IS}. Deviation from Hardy-Weinberg equilibrium (HWE) was tested for each population by locus. Genetic structure was further assessed by comparison of pairwise population F_{ST} values, and Analysis of Molecular Variance (AMOVA) was used to assess hierarchical variation within and among populations.

A Bayesian method of assignment was implemented in STRUCTURE v 2.3.4 (Pritchard et al., 2000). The assignment in STRUCTURE uses a Monte Carlo Markov Chain (MCMC) algorithm to assign individuals to a cluster based on their posterior probability of a multi-locus genotype in that individual, regardless of the site from which the samples were obtained. The likelihood of the posterior probability distributions was computed for each cluster (K = 1-10). Each model was simulated ten times, after the

specified burn-in (500 000 generations), assuming admixture and correlated allele frequencies. Structure Harvester (Earl and Vonholdt, 2012) was used to find ln likelihood (L(*K*) and Δ*K*), which is estimated based on the rate of change of L(*K*) between successive *K* values (Evanno et al., 2005). Additionally, Principal coordinate analysis (PCoA) was performed in GenAlEx 6.5 using an individual genetic distance matrix to further visual structure among individuals and populations.

CHAPTER 3. RESULTS

A total of 193 individuals from five populations were genotyped using seven microsatellite loci. Summary statistics are provided in Table 1 and the Appendix. All loci included here were polymorphic, with the mean number of alleles per population ranging from 4.714 (Davenport East) to 5.571 (Manus). One locus (*AB51*) exhibited deviation from HWE in two populations, while three other loci (*AB83*, *AB95*, and *AB170*) exhibited deviation from HWE in one population each (Appendix).

Pairwise population F_{ST} values varied from 0.007 between Davenport East and West, to 0.056 between Airport and Davis (Table 2). AMOVA results indicate that the majority of variation occurs within individuals (86%), while only 7% of variation occurs among sampled sites (Table 3). F_{ST} among populations was 0.066 ($P < 0.05$). PCoA results indicate that 31.40% of genetic variation among individuals was explained by the first three axes combined (Figure 1).

The maximum ln likelihood of the posterior probability distribution implemented in STRUCTURE reached a maximum value when $K = 3$ (Figure 2). At the threshold for probability of assignment (0.80), 31 of 52 individuals from the Manus site were assigned to cluster 1, one individual from Manus was assigned to cluster 2, and all other individuals from Manus were unassigned. For the Airport site, 24 of 51 individuals were assigned to cluster 2, while the remaining individuals were unassigned. At Davis, 9 of 23 individuals were assigned to cluster 3, one was assigned to cluster 2, and all remaining

individuals were unassigned. Davenport East and West were largely unassigned (22 of 30, and 29 of 37, respectively).

CHAPTER 4. DISCUSSION

Astragalus bibullatus is a federally endangered legume that appears to exhibit limited genetic structure within and among naturally occurring populations (Baskauf and Snapp, 1998; Morris et al., 2002; Baskauf and Burke, 2009). For the current study, we developed novel nuclear microsatellite loci in this species to further investigate genetic variation within and among populations (Morris et al., 2016). Previous work used allozymes (Baskauf and Snapp, 1998; Morris et al., 2002) and amplified fragment length polymorphisms (AFLP) (Baskauf and Burke, 2009). It is important to note that that analyses of these different marker systems are only comparable when used on the same plant tissues. While we sampled some of the same populations as previous studies, there is not complete overlap between populations sampled. Therefore, it does not make sense here to directly compare diversity measures, but instead to discuss trends in the results.

As suggested by previous work (Baskauf and Snapp, 1998; Morris et al., 2002; Baskauf and Burke, 2009), the data presented here indicate that naturally occurring populations of *A. bibullatus* are genetically similar, with pairwise F_{ST} values of 0.056 or less (Table 2). These values are considered low or moderate at best. Additionally, the PCoA results presented here show nearly complete overlap of individuals from different populations, further suggesting limited differentiation among sites (Figure 1). However, Bayesian assignment tests in STRUCTURE supported recognition of three possible clusters, although these clusters do not seem to correlate with geographic distribution of sites (Figure 2). This combined pattern of limited differentiation between sites coupled

with structure that is inconsistent with geographic distance between sites would suggest that genetic drift is likely playing a noticeable role in at least some populations (i.e., Airport, Davis, and Manus). These findings are somewhat similar to those of Breinholt et al. (2009) who used AFLPs to evaluate genetic structure in *A. ampullarioides*, an endangered species endemic to Washington County, Utah, in and around Zion National Park. In that study, gene flow was determined to be low among populations based on population genetics statistics. Some geographic structure was identified, but the authors concluded that low gene flow and geographic isolation, in addition to the fact that this species is self-compatible, have ultimately amplified the impact of genetic drift.

Since these kinds of genetic surveys show a snapshot in time, and seed dormancy is perceived to be long lasting, additional genetic assessment of this species is needed. Based on soil seed bank studies in *A. bibullatus*, Morris at al. (2004) inferred a loss of genetic variability over time. We know very little about the length of seed dormancy in this species, and it is possible that the bulk of seeds recruited from the seed bank are those from recent years that have not yet made it into the soil. Future efforts will focus on determining the extent of the soil seed bank at selected sites and determining the potential of the soil seed bank to serve as a source for novel genetic variants in reintroduction efforts.

The markers used for this study were nuclear microsatellite loci. Given the relatively close geographic proximity of the sites sampled here and the limited genetic structure recovered, movement of pollinators within and among sites does not appear to be limited. However, given the short-distance seed dispersal these plants are known to

exhibit, it might be beneficial to use chloroplast microsatellite loci as a second approach as they can be used to trace maternal inheritance within and among sites, providing insight into how limited seed dispersal influences population structure in this species.

One primary goal of this thesis work was to assess genetic diversity within and among populations of *A. bibullatus* as a tool for making more informed decisions with respect to reintroduction and conservation efforts. The main challenges in making such decisions are the limited number of seed source sites and the rapid expansion of human development in Rutherford County, the only county in which this species is known to occur. The oldest known herbarium specimen of *A. bibullatus* was collected by Gattinger in 1881, but it was originally identified as another species (Barneby and Bridges, 1987). This plant would not be named *A. bibullatus* until 1987 when it was rediscovered by Milo Pyne, for whom the plant is named (Pyne's ground plum). Given the history of the taxonomy of the species, and given the human history of rapid growth of this region, it is likely that the glades themselves, and therefore *A. bibullatus*, once encompassed a larger area than they do now. Additional work is underway in the Morris Lab to determine genetic structure within and among natural and reintroduced populations of *A. bibullatus*. With this additional information, it may be possible to better understand the limited genetic diversity observed among natural populations here.

REFERENCES

ALBRECHT, M. A., AND J. C. Z. PENAGOS. 2012. Seed germination ecology of three imperiled plants of rock outcrops in the southeastern United States. *Journal of the Torrey Botanical Society* 139: 86-95.

ALBRECHT, M. A., R. E. BECKNELL, AND Q. LONG. 2016. Habitat change in insular grasslands: Woody encroachment alters the population dynamics of a rare ecotonal plant. *Biological Conservation* 196: 93-102.

BARNEBY, R. C., AND E. L. BRIDGES. 1987. A new species of *Astragalus* (Fabaceae) from Tennessee Central Basin. *Brittonia* 39: 358-363.

BASKAUF, C. J., AND S. SNAPP. 1998. Population genetics of the cedar-glade endemic *Astragalus bibullatus* (Fabaceae) using isozymes. *Annals of the Missouri Botanical Garden* 85: 90-96.

BASKAUF, C. J., AND J. M. BURKE. 2009. Population genetics of *Astragalus bibullatus* (Fabaceae) using AFLPs. *Journal of Heredity* 100: 424-431.

BASKIN, J. M., AND C. C. BASKIN. 1999. Cedar glades of the southeastern United States. *In* R. C. Anderson, J. S. Fralish, AND J. M. Baskin [eds.], Savannas, barrens, and rock outcrop plant communities of North America, 206-219. Cambridge University Press.

BASKIN, J. M., AND C. C. BASKIN. 2003. The vascular flora of cedar glades of the southeastern United States and its phytogeographical relationships. *Journal of the Torrey Botanical Society* 130: 101-118.

BASKIN, J. M., C. C. BASKIN, AND P. J. LAWLESS. 2007. Calcareous rock outcrop vegetation of eastern North America (Exclusive of the Nashville Basin), with particular reference to use of the term "cedar glades". *Botanical Review* 73: 303-325.

BREINHOLT, J. W., R. VAN BUREN, O. R. KOPP, AND C. L. STEPHEN. 2009. Population genetic structure of an endangered Utah endemic, *Astragalus ampullarioides* (Fabaceae). *American Journal of Botany* 96: 661-667.

DUNCAN, R. S., C. B. ANDERSON, H. N. SELLERS, AND E. E. ROBBINS. 2007. The effect of fire reintroduction on endemic and rare plants of a southeastern glade ecosystem. *Restoration Ecology* 16: 39-49.

EARL, D. A., AND B. M. VONHOLDT. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4: 359-361.

EDWARDS, A. L., B. WILTSHIRE, AND D. L. NICKRENT. 2004. Genetic diversity in *Astragalus tennesseensis* and the federal endangered *Dalea foliosa* (Fabaceae). *Journal of the Torrey Botanical Society* 131: 279-291.

EVANNO, G., S. REGNAUT, AND J. GOUDET. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* 14: 2611-2620.

FAIRCLOTH, B. C. 2008. MSATCOMMANDER: detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8: 92-94.

FRODIN, D. G. 2004. History and concepts of big plant genera. *Taxon* 53: 753-776.

LEVEY, D. J. 1988. Tropical wet forest treefall gaps and distributions of understory birds and plants. *Ecology* 69: 1076-1089.

MORRIS, A. B., R. S. BAUCOM, AND M. B. CRUZAN. 2002. Stratified analysis of the soil seed bank in the cedar glade endemic *Astragalus bibullatus*: Evidence for historical changes in genetic structure. *American Journal of Botany* 89: 29-36.

MORRIS, A. B., C. SCALF, A. BURLEYSON, L. JOHNSON, AND K. TROSTEL. 2016. Development and characterization of microsatellite primers in the federally endangered *Astragalus bibullatus* (Fabaceae). *Applications in Plant Sciences* 4.

MYERS, R. K., AND D. H. VAN LEAR. 1998. Hurricane-fire interactions in coastal forests of the South: a review and hypothesis. *Forest Ecology and Management* 103: 265-276.

NEEL, M. C. 2008. Patch connectivity and genetic diversity conservation in the federally endangered and narrowly endemic plant species *Astragalus albens* (Fabaceae). *Biological Conservation* 141: 938-955.

PEAKALL, R., AND P. E. SMOUSE. 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288-295.

PEAKALL, R., AND P. E. SMOUSE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28: 2537-2539.

PRITCHARD, J. K., M. STEPHENS, AND P. DONNELLY. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.

SCHUELKE, M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnology* 18: 233-234.

U.S. FISH AND WILDLIFE SERVICE. 1991. Endangered and threatened wildlife and plants; *Astragalus bibullatus* (Guthrie's Ground-plum) determined to be endangered. *Federal Register* 56: 48748-48751.

U.S. FISH AND WILDLIFE SERVICE. 2009. Technical/Agency draft recovery plan for *Astragalus bibullatus* (Pyne's Ground-plum), Atlanta, Georgia.

VICENTE, M. J., F. SEGURA, M. AGUADO, D. MIGLIARO, J. A. FRANCO, AND J. J. MARTINEZ-SANCHEZ. 2011. Genetic diversity of *Astragalus nitidiflorus*, a critically endangered endemic of SE Spain, and implications for its conservation. *Biochemical Systematics and Ecology* 39: 175-182.

APPENDICES

APPENDIX A: TABLES AND FIGURES

Table 1. Summary statistics for five naturally occurring populations of *Astragalus bibullatus* based on 193 individuals averaged across seven nuclear microsatellite loci. All loci were previously published by Morris et al. (2016).

Population	N_a	N_e	H_o	H_e	F_{IS}
DAVW	5.143	2.935	0.590	0.638	0.076
	(0.884)	(0.362)	(0.070)	(0.043)	(0.064)
DAVE	4.857	2.917	0.655	0.622	-0.089
	(0.829)	(0.416)	(0.052)	(0.058)	(0.060)
AIR	5.000	2.231	0.459	0.494	0.049
	(1.113)	(0.299)	(0.079)	(0.080)	(0.053)
DAVIS	4.714	2.679	0.574	0.590	0.003
	(0.837)	(0.408)	(0.058)	(0.057)	(0.048)
MANUS	5.571	3.071	0.599	0.583	-0.026
	(1.462)	(0.719)	(0.094)	(0.079)	(0.044)

Populations listed in order of geographic proximity: DAVW = Davenport West and DAVE = Davenport East, both of which are on Flat Rock State Natural Area; AIR = Airport; DAVIS = Davis; MANUS = Manus Road State Natural Area. N_a = mean number of alleles over loci for each population (and standard error); N_e = effective number of alleles calculated as $1/(\Sigma\, p_i^2)$ which is the sum of the squared population allele frequencies (and standard error); H_o = observed heterozygosity calculated here as mean (and standard error) over all loci for each population; H_e = unbiased expected heterozygosity calculated here as $(2N / (2N\text{-}1))^* H_e$ where H_e is calculated as mean (and standard error) over all loci for each population; F_{IS} = fixation index calculated over all loci for each population (and standard error).

Table 2. Population pairwise F_{ST} values for five naturally occurring populations of *Astragalus bibullatus* based on 193 individuals averaged across seven nuclear microsatellite loci.

Population	DAVW	DAVE	AIR	DAVIS	MANUS
DAVW	0.000				
DAVE	0.007	0.000			
AIR	0.047	0.039	0.000		
DAVIS	0.043	0.042	0.056	0.000	
MANUS	0.028	0.035	0.048	0.054	0.000

Populations listed in order of geographic proximity: DAVW = Davenport West and DAVE = Davenport East, both of which are on Flat Rock State Natural Area; AIR = Airport; DAVIS = Davis; MANUS = Manus Road State Natural Area.

Table 3. Results of analysis of molecular variance (AMOVA) based on seven nuclear microsatellite loci amplified in 193 individuals of *Astragalus bibullatus* from five naturally occurring populations in Rutherford County, Tennessee.

Source	df	Sum of squares	Variance components	% of variation	F_{ST}
Among pops	4	53.168	13.292	7	0.066 (0.001)*
Among indiv	188	420.511	2.237	7	
Within indiv	193	369.000	1.912	86	
Total	385	842.679		100	

df = degrees of freedom; * ($P < 0.05$)

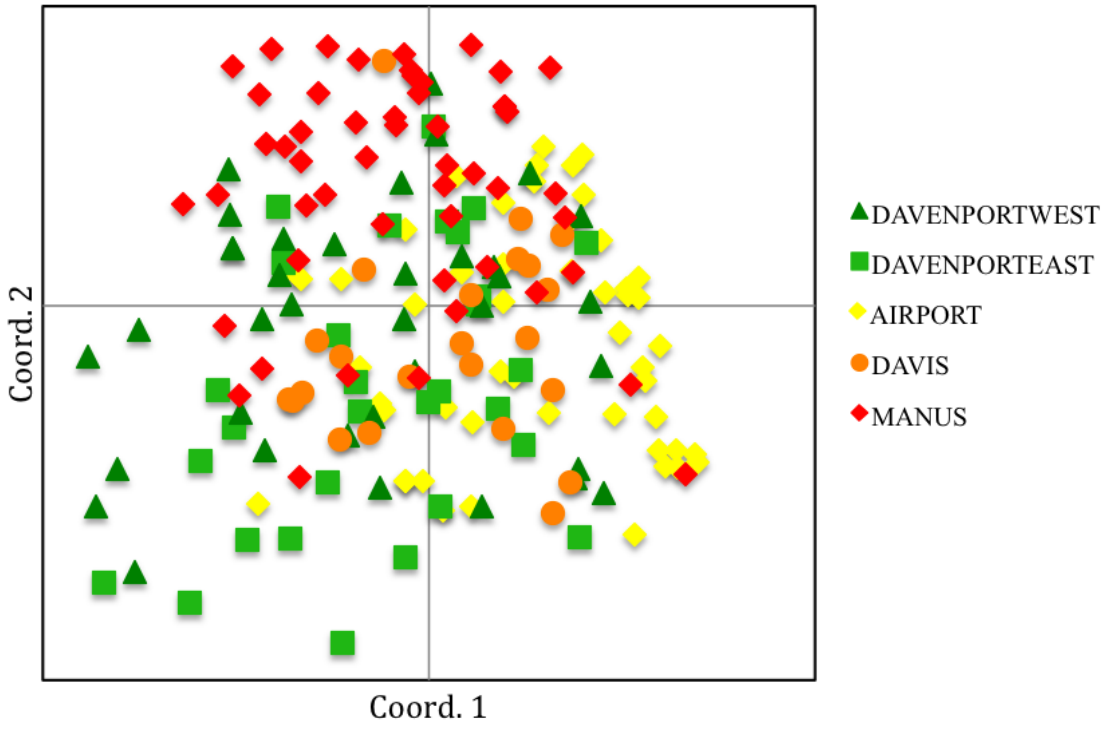

Figure 1. Principal coordinates analysis (PCoA) of five naturally occurring populations of *Astragalus bibullatus* based on 193 individuals averaged across seven nuclear microsatellite loci. Populations are listed in legend in order of geographic proximity. Percentage of variation explained by first three axes was 12.85, 9.49, and 9.06, respectively, or 31.40 cumulatively.

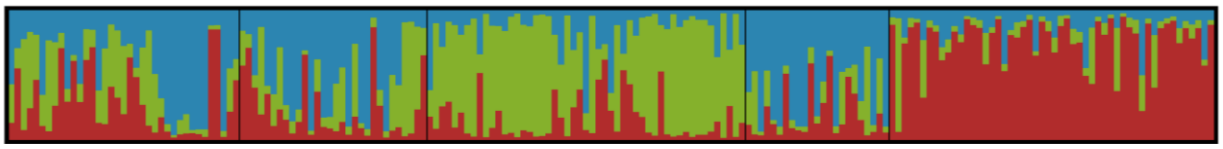

Figure 2. Clustering analysis of genetic structure within and among natural populations of *Astraglus bibullatus*. Populations sampled in the current study are listed geographically from north to south. Manus is geographically the most distant from all other populations, although all five populations are within approximately five miles of each other. The maximum ln likelihood of the posterior probability distribution implemented in STRUCTURE for these populations reached a maximum value when $K = 3$, with each cluster indicated here as a separate color.

APPENDIX B: ADDITIONAL STATISTICS

Summary statistics for seven nuclear microsatellite loci amplified in 193 individuals of *Astragalus bibullatus* from five naturally occurring populations in Rutherford County, Tennessee.

Locus	N_a	N_e	H_o	H_e	F_{IS}
AB51	10.600 (0.927)	4.219 (0.700)	0.751 (0.058)	0.755 (0.033)	-0.013
AB59	5.000 (0.000)	3.420 (0.464)	0.687 (0.081)	0.684 (0.063)	-0.019
AB83	4.600 (0.245)	3.611 (0.325)	0.701 (0.024)	0.724 (0.029)	0.018
AB93	2.400 (0.245)	1.663 (0.157)	0.380 (0.075)	0.380 (0.067)	-0.015
AB95	4.400 (0.245)	1.759 (0.157)	0.385 (0.036)	0.420 (0.054)	0.068
AB113	5.400 (0.245)	2.641 (0.172)	0.655 (0.052)	0.624 (0.027)	-0.066
AB170	3.000 (0.000)	2.051 (0.150)	0.468 (0.061)	0.510 (0.033)	0.068

N_a = mean number of alleles over populations for each locus (and standard error); N_e = effective number of alleles calculated as $1/(\Sigma\, p_i^2)$ which is the sum of the squared population allele frequencies (and standard error); H_o = observed heterozygosity calculated here as mean (and standard error) over all populations for each locus; H_e = unbiased expected heterozygosity calculated here as $(2N / (2N\text{-}1))^* H_e$ where H_e is calculated as mean (and standard error) over all populations for each locus; F_{IS} = fixation index calculated over all populations for each locus.