# Population Genetics of the Endangered Wetland Endemic, *Xyris tennesseensis* Kral (Xyridaceae)

Kela M. Downey

# Population Genetics of the Endangered Wetland Endemic, *Xyris tennesseensis* Kral (Xyridaceae)

A Thesis

Presented to

The College of Graduate Studies

Austin Peay State University

In Partial Fulfillment

Of the Requirements for the Degree

Master of Science in Biology

Kala Downey

August, 2016

August, 2016

To the College of Graduate Studies:

We are submitting a thesis written by Kala Downey entitled "Population Genetics of the Endangered Wetland Endemic, *Xyris tennesseensis* Kral (Xyridaceae)." We have examined the final copy of this thesis for form and content. We recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science in Biology.

Carol J. Baskauf

**Carol Baskauf**

**Research/Committee Advisor/Chair**

Jeff Lebkuecher

**Jeff Lebkuecher**

**Committee Member**

**Stefan Woltmann**

**Committee Member**

Accepted for the Graduate and Research Council

Raja Dakshinamurthy

Dean, College of Graduate Studies

(Original signatures on file with official student records.)

Statement of Permission to Use

In presenting this thesis in partial fulfillment of the requirements for the Master of Science in Biology at Austin Peay State University, I agree that the library shall make it available to borrowers under the rules of the library. Brief quotations from this field study are allowable without special permission, provided that accurate acknowledgement of the source is made. Permissions for extensive quotation or reproduction of this field study may be granted by my major professor, or in his/her absence, by the Head of the Interlibrary Services when, in the opinion of either, the proposed use of the material is for scholarly purposes. Any copying or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Kala Downey

08/17/2016
Date

**Table of Contents**

Acknowledgements

A number of people were instrumental in the completion of this thesis and deserve my gratitude. First, I must thank my advisor, Dr. Carol Baskauf, who has been with me every step of the way. She advised in the lab, guided in the field, and pulled almost as many sleepless nights for this project as I did. I could not have accomplished this without her. My committee members, Dr. Stefan Woltmann and Dr. Jeff Lebkuecher, were also very helpful in sharing their experience with unfamiliar programs and encouraging more thorough exploration of the implications of my findings.

I could not have completed my collections without the aid of many people and organizations. Andrea Bishop from the Tennessee Department of Environment and Conservation (TDEC) and the Tennessee Natural Heritage Program, and Allen Trently, also a member of the Natural Heritage Program, provided information on Tennessee site locations and in several instances acted as guides during collection trips. Dr. Steve Baskauf also kindly donated his time to escort my companions and me to two Tennessee localities. David Lincicome and Roger McCoy (TDEC) and Geoff Call of the U.S. Fish and Wildlife Service (USFWS) helped me to obtain collection permits for Tennessee and put me in contact with a USFWS representative in Alabama, Shannon Holbrook. Leah Storino allowed us access to two Alabama sites and acted as a guide at these sites as well. Dr. Bob Boyd of Auburn University put me in contact with many of the people who gave me information about sampling locations or guided me to sites. He also tirelessly answered many follow-up questions throughout this research. Al Schotz and Michael Barbour helped me get in touch with the Hogue family. This couple graciously permitted sampling on their property in Alabama. Dr. Mike Hardig led us to the site owned by Montebello

University. Dr. Mincy Moffett, a U.S. Fish and Wildlife representative in Georgia, enthusiastically shared his expertise on *Xyris tennesseensis,* in addition to guiding my field assistants and me to every Georgia site included in this study. I must also express my personal appreciation for his help in finding the local Urgent Care following an unfortunate encounter with fire ants at one the first sampling locations.

Lastly, I would like to show my gratitude for the participation of several of my classmates. In particular, Nick Willis, who attended nearly all of my sampling trips, as well as Nadilen Bannis, who was irreplaceable in the laboratory, patiently and cheerfully wading through every setback with me. Though he was unable to join me as often as he wanted, Richard Hankenson took part in one of our Tennessee collection trips and joined me in the lab a few times. Shawn Settle requested vacation time off of work to join me for the two-day sampling trip to Georgia. I cannot thank these individuals enough for their assistance.

## ABSTRACT

KALA MARIE DOWNEY. Population Genetics of the Endangered Wetland Endemic, *Xyris tennesseensis* Kral (Xyridaceae) (under the direction of DR. CAROL BASKAUF).

Genetic variation is required by species for long-term adaptive evolution, and loss of variability can cause inbreeding depression and impair growth and disease resistance in the short-term. Information about the distribution and levels of genetic diversity in rare species is important for prioritizing the protection of populations so as to represent the bulk of the species' diversity. *Xyris tennesseensis* is a federally endangered obligate wetland plant found in critically imperiled seep and riparian habitats, with a highly disjunct distribution across 10 counties in Tennessee, Alabama, and Georgia. The genetic diversity of this species was examined using 14 nuclear microsatellite markers that were polymorphic for the species. Five Tennessee, four Alabama, and five Georgia populations were sampled. Results showed limited variability, with a population mean of only 16% polymorphic loci. Observed heterozygosity in all populations was unexpectedly low, ranging from 0.000 to 0.095, with an average of 0.017. This resulted in a relatively high $F_{IS}$ of 0.71, suggesting that *X. tennesseensis* experiences high levels of inbreeding. Two Alabama populations contain the majority of within-population diversity seen in *X. tennesseensis*, whereas most others (including all in Tennessee and all but one in Georgia) appear to be genetically depauperate. Analysis of molecular variance estimated that differentiation among the regions (states) accounted for the majority (60%) of the genetic variation in this species, with an additional 31% due to genetic differences among the populations of each state and only 9% due to variability within populations. These data indicate *X. tennesseesis* possesses low levels of genetic diversity and there is very little, if any, gene flow

among the populations. To preserve what remains of the genetic variation of this species, it is essential that multiple populations be protected, with priority given to the largest and most diverse populations in each region.

LIST OF FIGURES

FIGURE PAGE

# LIST OF TABLES

TABLE PAGE

# CHAPTER I

## Introduction

Understanding the levels and distribution of genetic diversity of endangered species can be integral to the success of conservation biology goals – goals such as the preservation and management of rare species, including their reintroduction into appropriate habitat (Frankham 1995). Genetic diversity provides the building blocks for species to adapt to ever changing environments (Soule 1980, Simberloff 1988, Hamrick and Godt 1989, Barrett and Kohn 1991, Huenneke 1991). With high levels of genetic variability comes a higher probability of potentially beneficial alleles, which could increase the population's or species' chances of survival in the long-term (Hamrick and Godt 1989, Barrett and Kohn 1991, Huenneke 1991). Even from a short-term perspective, the loss of genetic variation can have a negative impact on a species' growth rate, development, and response to disease (Beardmore 1983, Huennke 1991, Ellstrand and Elam 1993, Reed and Frankham 2003). Studying the various aspects of genetic diversity can provide insight that might allow one to predict how a species may change over time (Weir 1990) or determine how a population of an endangered species might be managed to assist in its conservation (Frankham 1995). For example, populations that have drastically decreased in size and are thus experiencing high rates of inbreeding might be expected to decline due to the harmful effects of inbreeding depression. Therefore, genetic material from a population possessing different or more alleles might be used to augment the inbred population in an attempt to increase the genetic diversity available in the population.

## *Xyris tennesseensis*

### *Description*

*Xyris tennesseensis* Kral (Xyridaceae), Tennessee yellow-eyed grass, is a rare perennial monocot endemic to the southeastern United States (Kral 1978, U.S. Fish and Wildlife Service 1994, Boyd and Moffett 2010). Plants typically occur in a clumping formation (Fig. 1). Mature leaves exhibit the twisted shape prevalent in the Xyridaceae family (Fig. 1), and can range from 14 to 45 cm in height (Kral 1978, U.S. Fish and Wildlife Service 1991). The leaves originate from a rounded base, which often displays a pink or purple tint (Kral 1978).

Figure 1. Photo of *Xyris tennesseensis* leaves (photo courtesy of Carol Baskauf).

*Xyris tennesseensis* most often flowers from August to September (Kral 1978), though flowering has been seen to occur as early as July (Boyd et al. 2011). The inflorescence (Fig. 2) is a small brown cone-like structure ranging from 1 to 1.5 cm in length found at the tips of narrow brown stalks ranging from 30 to 70 cm tall (U.S. Fish and Wildlife Service 1991). One or two pale to bright yellow flowers emerge from the bracts at mid to late-morning and wilt by mid-afternoon (U.S. Fish and Wildlife Service 1991, Wall et al. 2002). Fruits are dehiscent capsules hidden within the bracts with many seeds that are released at the end of the flowering season when the cone bracts open (Kral 1978, Wall et al. 2002, Boyd et al. 2011). Vegetative growth is accomplished by axillary bud formation at the leaf base, contributing the clumping arrangement of ramets (Kral 1990).

Figure 2. Photo of *Xyris tennesseensis* inflorescence (photo courtesy of Shawn Settle).

*Xyris tennesseensis*, much like the rest of the 25 *Xyris* species found in North America north of the Mexico border, thrives in seeps and along the edges of streams under full sun to light shade. This rare species differs from other Xyrids in that it requires the more basic pH levels found in calcareous soils associated with shale, limestone, and dolomite bedrock, rather than the commonly preferred acidic soils of bogs (Kral 1979, Kral 2000). The initial survey of the species by Kral (1978) documented seven extant sites occurring over the Western Highland Rim of Tennessee, the Cumberland Plateau of Alabama, and the Valley and Ridge Province of Georgia. All but two of these sites (one in Franklin County Alabama, and one in Natchez Trace State Park in Tennessee) were entirely on privately owned and, thus, unprotected land. *Xyris tennesseensis* was placed under the protection of the Endangered Species Act due to the natural and man-made threats resulting from the unusually specific habitat requirements (U.S. Fish and Wildlife Service 1991). The current range of the species remains in the three regions identified by Kral (1978), with populations now known from one county in Tennessee, five in Alabama, and five in Georgia (Fig. 3). A total of 22 extant populations are known from these three states (Boyd and Moffett 2010).

The wetland habitat of *X. tennesseensis* is also severely threatened. A U.S. Fish and Wildlife Service aerial photography survey of mainland United States wetlands from the 1950s through the 1970s determined that approximately 50% of those wetlands fell within the southeastern states (Hefner and Brown 1984, Opheim 1997). The survey revealed a large loss of national wetlands, 84% occurring in the southeast (Hefner and Brown 1984). In Tennessee, *X. tennesseensis* is located in Highland Rim Seepage Fens, which are considered critically

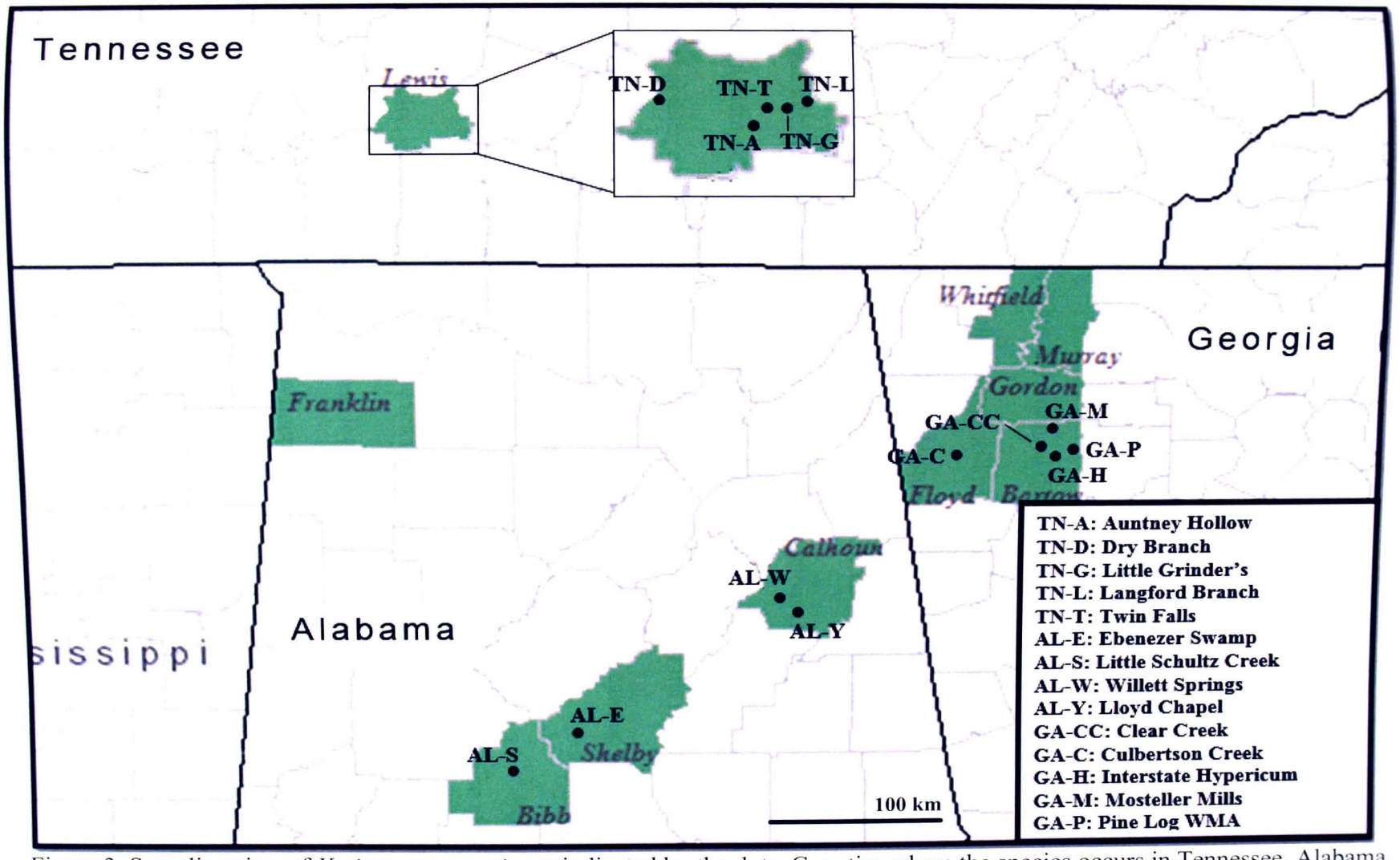



Figure 3. Sampling sites of *Xyris tennesseensis* are indicated by the dots. Counties where the species occurs in Tennessee, Alabama, and Georgia are shown in green.

imperiled and known only from roughly ten locations in three counties (NatureServe 2013). Both Alabama and Georgia possess a corresponding habitat referred to as Southern Ridge and Valley Seepage Fen. The number of locations with suitable *X. tennesseensis* habitat in Alabama and Georgia is currently unknown.

While conservation efforts are now being made to preserve wetland habitats, driven by the implementation of the "No Net Loss" program in the early 1990s (Heimlich et al. 1997, Morgan and Roberts 2003), it has been argued that the current definition of "wetlands" is too narrow. Because of this, new policies may still result in mass losses of biodiversity in these fragile ecosystems that are not yet fully understood (Kentula 1996, Opheim 1997, Heimlich et al. 1997, Whigham 1999, Morgan and Roberts 2003). To achieve the goal of conserving wetland habitat, private, state, and government organizations must have a thorough understanding of the type of wetland in question. In order to successfully recreate or restore these habitats, a record of the inhabiting species and knowledge of the ecology and the genetics of these species — especially those with endangered or threatened status — is vital if the existing biodiversity is to be maintained.

There is some evidence that the genetic diversity within one species can also affect the productivity, growth, and inter-specific interactions within communities, as well as affecting ecosystem-level processes. For instance, levels of diversity in one plant species can cause the populations of another plant species to fluctuate in numbers, which can in turn affect nutrient availability in an ecosystem (Hughes et al. 2008). Booth and Grime (2003) showed that multi-species communities with higher genetic diversity within each species maintained higher overall species diversity over time than those communities hosting the same species composition with lower genetic diversity. Studies such as these suggest that the level of

diversity seen in one species can be of unexpected significance and should be considered when attempting to preserve the overall biodiversity of threatened ecosystems. Thus, information gathered from a population genetics study of *X. tennesseensis* could aid in the goal of not only conserving the species, but the overall habitat as well.

*Breeding Ecology*

Boyd et al. (2011) observed the breeding ecology of an Alabama *X. tennesseensis* population. Although there is anecdotal evidence that many *Xyris* species may produce seeds apomictically (Kral 1983), Boyd et al. (2011) determined that *X. tennesseensis* could produce viable seeds from both self-fertilization and cross-pollination. This mixed breeding system may result in higher levels of genetic variation in populations that are able to participate in outcrossing (Barrett and Kohn 1991), while smaller populations that could be more dependent on frequent inbreeding, including selfing, may exhibit lower levels of heterozygosity within the population. In the Boyd et al. (2011) study, five groups were compared: self-pollinated (stamen removed with tweezers and brushed against stigma); not pollinated (flowers were bagged to block insect pollination and no hand-pollination occurred); outcrossed once (no insect activity permitted, but flowers were outcrossed by hand); outcrossed twice (outcrossed hand-pollination performed twice); and open pollinated (flowers were not bagged and no hand-pollination occurred). Seed germination rates were compared for all treatments, and one finding was that seeds resulting from any type of selfing (manipulated selfing from the "self-pollinated" treatment or natural selfing from the "not pollinated" treatment) germinated better than any manipulated outcrossing (outcrossed once

or twice). Comparisons of selfing treatments with "open pollinated" seeds were inconsistent (germination rate of seeds produced by open pollination were significantly lower than rates from the manipulated selfing treatment but not from the natural selfing treatment). Boyd et al. (2011) proposed no explanation for the increased germination of seeds from selfing.

Frequent gene flow among all the *X. tennesseensis* populations via pollen is likely limited by the small body size (approximately 0.5 cm) of *Lasioglossum zephyrum* Smith, the primary *X. tennesseensis* pollinator, observed by both Boyd et al. (2011) and Wall et al. (2002). Other observed pollinators that performed pollen-collecting behaviors were of a similarly small size (Boyd et al. 2011). Whether or not gene flow by means of pollinators can occur between some populations within Tennessee, Georgia, and Alabama is unknown, but there is virtually no chance that gene flow will occur from one state to the next with such small pollinators. In a study involving 16 bee species, Gathmann and Tscharntke (2002) found a positive correlation between body size and travel distance as well as trip duration. Foraging locations were limited by the location of the nesting site, and in several species, 90% of the bees were unable to find their way back to the nest when displaced more than 510 m.

Gene flow could also occur through seed dispersal between populations connected by waterways, with genetic material moving downstream from one population to the next (Baskin and Baskin 2003). All of the Tennessee populations are found on separate streams that eventually drain into Big Swan Creek, with the exception of Dry Branch, which drains into Cane Creek. However, as there are no *X. tennesseensis* populations found on Big Swan Creek or Cane Creek, there is no indication that the Tennessee populations share genetic material in this way. None of the populations from Alabama and Georgia included in this study are located downstream of another population.

Because *X. tennesseensis* seeds are very tiny, dispersal by wind could be possible, perhaps blowing seeds into the nearby stream. It is unlikely, though, that the seeds would travel far enough on wind alone for this method to play an important role in gene flow, especially since most populations are surrounded by forests and/or thick vegetation. The seeds have no special adaptations for animal dispersal. There are no reports of herbivores actively targeting *X. tennesseensis* for consumption. Some populations are found in grazing pastures, but the cattle found there are not free-range and are therefore unlikely to transport seeds among populations. While a wild animal could easily brush against a seed cone, the seeds of *X. tennesseensis* do not have clinging mechanisms to attach to fur for long periods of time.

*Endangered Species Status*

With most populations exhibiting chronically low numbers of individuals (U.S. Fish and Wildlife Service 1994), *X. tennesseensis* is at a high risk for inbreeding and for allele loss resulting from random genetic drift (Barrett and Kohn 1991). Three of the most commonly used techniques to combat dwindling numbers in populations of rare species are the protection of high priority sites, the bolstering of struggling populations using their own greenhouse-raised offspring (e.g. Moffett and Boyd 2013) or at least those of similar ecotypes, and the creation of new populations using offspring from one or more extant populations (Philippart 1995, Robert 2009). It is important that conservationists consider genetic information in such attempts. The more genetic diversity occurring in a population or species, the greater the chance there will be of having beneficial alleles for natural selection

(Weir 1990, Barrett and Kohn 1991), so it could be advantageous to utilize individuals from more diverse populations in creating new populations or augmenting existing ones. On the other hand, a manager may want to avoid combining individuals from populations that are genetically very different from each other, in case the genetic differences detected correlate with local adaptation to each population's respective habitat.

Following recommendations in the official U.S. Fish and Wildlife Service recovery plan (1994) for *X. tennesseensis*, studies have been conducted to investigate seed propagation and germination (Baskin and Baskin 2003), root micropropagation and seed cryopreservation (Johnson et al. 2012), general floral ecology (Boyd et al. 2011), and shrub-growth reduction at several known sites (Moffett and Boyd 2013). The plan also recommends genetic research for this species, but as of yet no work has been published to address this aspect.

Understanding the population genetics of an endangered species is an integral part of protecting the species (Fukunaga et al. 2005, Segarra-Moragues et al. 2005, Pettengill and Need 2011, Oleas et al. 2012). Because a population genetics study can reveal the magnitude and distribution of a species' genetic variation, these data were originally requested in *X. tennesseensis* ' recovery plan as part of the criteria for ranking sites in order of protection priority, the intent being that populations with higher levels of genetic diversity or those presenting unique alleles receive closer monitoring and/or more stringent protection policies (U.S. Fish and Wildlife Service 1994). It would be useful to quantify the amount of genetic variation within and among populations of *X. tennesseensis* to determine whether or not *X. tennesseensis* exhibits the reduced genetic variability typical for populations of a rare species (Hamrick and Godt 1989, Barret and Kohn 1991, Gitzendanner and Soltis 2000, Cole 2003, Nybom 2004) and to evaluate the prevalence of genetic drift and outcrossing vs. inbreeding.

Although a preliminary study using isozymes was conducted recently (N. Willis, unpubl. data), only eight isozymes were resolved and almost no genetic variability was detected (no within-population and very little among-population variability). Clearly, for this species a more variable, higher resolution genetic marker would be useful. Thus, the current study utilizes microsatellites as genetic markers to carry out a population genetics study, as recommended in the recovery plan for *X. tennesseensis* (U.S. Fish and Wildlife Service 1994).

**Forces Affecting Genetic Variation**

The levels and distribution of genetic diversity in a species can be impacted by many evolutionary forces, such as mutation, gene flow, natural selection, random genetic drift, and nonrandom mating (Wright 1978, Barrett and Kohn 1991, Hartl and Clark 2007). Mutations introduce brand new alleles into a genome and can accumulate to have prominent effects on the species over time. Gene flow spreads alleles from one population to another, decreasing genetic differentiation among populations but increasing variation within populations. For plants, gene flow takes place by pollen or seed dispersal. Natural selection favors certain phenotypes of individuals that are therefore able to survive long enough to produce offspring. Random genetic drift is the alteration of allele frequencies due to pure chance, such as can occur because of a genetic bottleneck (e.g. a dramatic decrease in population size due to a natural disaster). Changes in allele frequencies due to random genetic drift can include the loss of alleles in a population, and thus the loss of genetic variability.

Some of these forces can exert stronger pressures on rare species than widespread ones, especially when the species consists of small populations (Barrett and Kohn 1991, Ellstrand and Elam 1993). Random genetic drift plays an increasingly important role as populations get smaller, and thus small populations are vulnerable to loss of alleles. Lande (1988) defines effective population size as "the number of individuals in an ideal population that would give the same rate of random genetic drift as in the actual population", and the "effective" population sizes for rare species will be even smaller than the physical population sizes. In such instances, the effects of genetic drift can outweigh the effects of natural selection and act as the predominant factor determining population genetic structure (Barrett and Kohn 1991, Kimura and Crow 1964, Hedrick and Hurt 2012). Gene flow can be more limited among populations of rare species, particularly if the species has experienced a recent reduction in the number and/or sizes of its populations (Kimura and Crow 1964). Along with this decreased gene flow, the occurrence of inbreeding may also be more prevalent in small populations as there are fewer individuals with which to cross-pollinate, and those in the surrounding vicinity may be closely related (Mustajarvi et al. 2001).

The breeding system of a species plays an important role in the levels and distribution of genetic variability within a species (Beardmore 1983, Brown 1990, Karron 1991). Predominantly outcrossing species produce offspring from the gametes of two separate individuals and typically display higher levels of heterozygosity. Inbreeding decreases heterozygosity in a population and the offspring may be more vulnerable to the harmful reproductive and fitness effects of inbreeding depression (Hartl 1981, Barrett and Kohn 1991, Gathmann and Tscharntke 2002). Selfing, the union of gametes from one individual, is the most extreme form of inbreeding. The ability to self can be beneficial in the event of low

pollinator visitation, however. Alternatively, some species can carry out vegetative reproduction, allowing the propagation of the population without pollination and the energetic expense of gametes. Plants produced this way are genetically identical to the parent, resulting in populations that may be unable to survive the introduction of a pathogen because they are genetically homogenous (Beardmore 1983, Ellstand and Elam 1993) or to adapt in the event of environmental change over time (Beardmore 1983, Barrett and Kohn 1991).

## Determining Genetic Variation

There are a number of ways to evaluate levels of genetic variation such as estimating the number of alleles per locus, the percentage of polymorphic loci, observed and expected heterozygosity levels, and Wright's *F*-statistics. Nei's unbiased genetic identity (or genetic distance) indicates genetic similarity (or dissimilarity) between pairs of populations, while various sorting or clustering methods can indicate genetic similarity among multiple populations.

### *Percentage of polymorphic loci and alleles per locus*

The percent of polymorphic loci (P) can be calculated as the proportion of all sampled loci that possess more than one allele per locus. Alternatively, sometimes the percent of polymorphic loci is reported as the proportion of sampled loci where the frequency of the

most common allele does not exceed a certain arbitrary level (either 95% or 99%), in order to de-emphasize rare alleles (Hartl and Clark 2007).

For a study involving multiple populations, calculations can be carried out at the population level and at the species level. For example, the proportion of polymorphic loci can be determined for each population and an average across all populations calculated. The data for all populations can also be combined and percent polymorphic loci calculated across the entire species (Hartl 1981). Similarly, the mean number of alleles per locus can be determined for each population and then these values can be averaged across populations, for a "population level" analysis. One can also add up how many different alleles are found for all loci throughout all populations of a species and divide that total by the number of loci to get a "species-level" analysis.

*Observed and expected heterozygosity*

Observed heterozygosity ($H_o$) is the proportion of heterozygotes found at each locus. Expected heterozygosity ($H_e$) is the proportion of heterozygotes that would be expected if the population were in "Hardy-Weinberg equilibrium" conditions. Mating must be random for a population to be in Hardy-Weinberg equilibrium, and the population must not be experiencing genetic mutations, natural selection, random genetic drift or gene flow from other populations. Thus, under equilibrium conditions a population is able to maintain constant allele and genotype frequencies across multiple generations.

The Hardy-Weinberg equation uses observed alleles frequencies to calculate the expected genotype frequencies for a population under equilibrium conditions. Although it is unlikely

that a natural population will be in perfect "equilibrium" conditions, the Hardy-Weinberg equation estimates can still be used as a baseline, for comparison purposes. A goodness-of-fit test is used to determine if the observed genotype frequencies of a population fit with those expected by Hardy-Weinberg. If the data do not fit Hardy-Weinberg expectations, then one or more Hardy-Weinberg equilibrium assumptions are not met — i.e. one or more evolutionary forces must be affecting allele frequencies at that locus. It should be noted that technical problems such null alleles, missing data, and incorrectly interpreted stutter peaks are all artifacts that could also give the impression that Hardy-Weinberg expectations are not being met. This can also result from sampling errors, such as mistakenly combining two populations with an unrecognized barrier to gene flow into one population, referred to as the Wahlund effect.

*Wright's* F-*statistics*

Wright's *F*-statistics can provide information about the distribution of genetic variability within and among populations. Wright's *F*-statistics (e.g. $F_{IS}$, $F_{ST}$) are calculated for polymorphic loci to compare the observed heterozygosity value to Hardy-Weinberg expectations at different levels of analysis (Wright 1978, Hartl and Clark 2007). $F_{IS}$, often referred to as the inbreeding coefficient, compares the observed heterozygosity levels for individuals within a population to Hardy-Weinberg expectations and ranges from -1 to 1. Values approaching 1 are due to a deficiency of heterozygotes and may indicate a prevalence of inbreeding, or may instead be due to the Wahlund effect. Negative values would indicate an excess of heterozygotes, perhaps due to selective pressure in favor of heterozygotes. A

value of 0 would indicate that the observed heterozygosity matches the expectations of Hardy-Weinberg exactly. Species that are obligate outcrossers often have $F_{IS}$ values relatively close to 0.

$F_{ST}$ estimates the proportion of the species' genetic diversity that is due to populations differing from each other ("among-population" genetic variability). $F_{ST}$ can range from 0 to 1 when using genetic markers with relatively few alleles per locus. If $F_{ST}$ values are low, the populations are genetically similar to each other, and the majority of the species' genetic diversity is due to within-population variation (i.e. individuals from the same population differing genetically from one another). Such genetic similarity among populations could be the result of significant gene flow occurring regularly among the populations, or a separation of the populations that only occurred fairly recently, and/or low levels of genetic drift following separation (Barrett and Kohn 1991, Gitzendanner and Soltis 2000). High $F_{ST}$ values indicate greater levels of among-population genetic variation, where populations differ from each other in terms of allele frequencies (Wright 1978, Reed and Frankham 2003). High population differentiation may indicate long-term isolation of the populations with little gene flow between them paired with the occurrence of genetic drift and/or the existence of varying selective pressures (Epperson 1989, Barrett and Kohn 1991).

Comparisons of $F_{ST}$ values among loci and among species must be made with caution as highly variable loci can make interpretation of $F_{ST}$ difficult. The higher the expected heterozygosity, the lower the maximum possible $F_{ST}$ will be. As high levels of heterozygosity do not necessarily mean that populations share the same alleles, this could result in a low $F_{ST}$ for populations that are in fact very different genetically (Meirmans and Hedrick 2011). This problem can be corrected for by calculating $F'_{ST}$ (Meirmans 2006).

which is $F_{ST}$ standardized by the maximum possible value for $F_{ST}$ based on the expected heterozygosity at each marker. Thus $F'_{ST}$ is more comparable across loci and species of differing variability levels. However, Whitlock (2011) argues that $F_{ST}$ is the better measure of the effects of population structure, and Meirmans and Hedrick (2011), pointing out that comparison with past studies is important, state that both statistics need to be reported. $G''_{ST}$, the unbiased estimator of $F'_{ST}$ (Meirmans and Hedrick 2011), can also be included for the sake of comparison.

*Nei's genetic distance and identity*

Nei's unbiased genetic distance and genetic identity (Nei 1978) can be used to determine genetic differentiation or similarity of population pairs, respectively. These calculations include all loci, monomorphic and polymorphic. The more similar two populations are, the closer the genetic identity score is to 1, whereas greater differentiation results in genetic identities closer to 0. If no alleles are shared between two populations, genetic identity is 0 and the genetic distance score is infinity (Nei 1978).

## Methods to Collect Genetic Data

Various genetic markers can be used to assess the genetic diversity of a species. Protein electrophoresis typically utilizes enzymes referred to as "isozymes" or "allozymes" (Hamrick and Godt 1989). Technically, isozymes are enzymes coded for by different gene loci whereas allozymes are different forms of the enzyme due to different alleles at one locus, but both

terms are often used synonymously in the literature. Variability in these enzymes indicates genetic variation in the populations. Since isozymes are codominant markers, heterozygote individuals can be detected (provided the allozymes are unequal in mass and/or charge), which is beneficial for population genetics studies. DNA-based methods, such as amplified fragment length polymorphisms (AFLPs) and microsatellites (short, tandem repeats of two to six nucleotides), are becoming more cost-effective and thus more popular in population genetics studies. These markers often display higher rates of polymorphism than isozymes, as they can be developed from sections of DNA that are not expressed and thus are more likely to be variable. This is particularly important when studying species with low levels of genetic variability, as is often the case for many rare species (Barret and Kohn 1991, Ellstrand and Elam 1993, Gitzendanner and Soltis 2000). Such genetic markers are considered to be selectively neutral, an important assumption for Hardy-Weinberg conditions (Liu and Cordes 2004) and can provide information on gene flow, genetic drift, and breeding system. Of the three types of markers mentioned above, only microsatellite loci are both codominant and highly variable, which make them particularly useful for population genetics studies in general and for rare species in particular. Though these are considered to be neutral markers and thus may not provide an accurate representation of the adaptive variation in *X. tennesseensis*, some studies have noted correlations between estimates made using neutral markers and adaptive trait diversity (Reed and Frankham 2001, Leinonen et al. 2008).

# CHAPTER II

## Materials and Methods

### Population Sampling

Geographically distant populations may be relatively different from each other genetically, and genetic diversity tends to be higher in larger populations. Thus, in an attempt to include as much of the species' genetic variability as possible, populations were sampled throughout the entire species range (Table 1, Fig. 3) and included most of the largest and "highest quality" (Moffett 2008) sites in all three states (although we were unable to sample the Franklin County, AL population which is also one of the larger sites). Some smaller sites were also sampled within each state. Roughly 58% of extant sites were utilized in the study. Five of the 6 populations in Tennessee were sampled, 4 of the 10 in Alabama, and 5 of the 6 extant sites in Georgia.

Leaf tissue collections occurred during the summer of 2013 and 2014. When collecting leaves, a portion of the tissue was placed in silica gel packets immediately for rapid drying. The rest of the tissue was stored in Ziploc® bags on ice in the field, and then refrigerated at 4° C in a laboratory setting until DNA was extracted.

*Xyris tennesseensis* tends to grow in small clumps, and since the species can grow vegetatively through axillary buds (Kral 1990), it is difficult to estimate population sizes. Boyd and Moffett (2010) recorded the number of flowering spikes along transects as an indication of *X. tennesseensis* population sizes (Table 1). It should be noted that the number of flowering spikes is only a rough estimate of population size, as not all plants present in a

population produce flowering spikes and individual plants can produce more than one flowering spike.

Table 1. Number of flowering spikes per *X. tennesseensis* populations during a 2008-2009 survey (Boyd and Moffett 2010).

| State | County | Population Name | Population Abbreviation | Flowering Spikes |
|---|---|---|---|---|
| Tennessee | Lewis | Little Grinders Creek | TN-G | 2997 |
| | | Auntney Hollow | TN-A | 361 |
| | | Twin Falls | TN-T | 1500 |
| | | Langford Branch | TN-L | 159 |
| | | Dry Branch | TN-D | 1459 |
| Alabama | Calhoun | Lloyd Chapel Swale | AL-Y | 22 * |
| | | Willett Springs | AL-W | 700 |
| | Bibb | Little Schultz Creek | AL-S | 8064 |
| | Shelby | Ebenezer Swamp | AL-E | 366 |
| Georgia | Bartow | Mosteller Mill Springs | GA-M | 9793 |
| | | Interstate Hypericum | GA-H | 771 |
| | | Clear Creek Lake** | GA-CC | *** |
| | | Clear Creek Spring** | GA-CC | 1360 |
| | | Pine Log Springs (WMA) | GA-P | 127 |
| | Floyd | Colbertson Spring | GA-C | 252 |

* Many seedlings also present.
**Clear Creek Lake and Clear Creek Spring are now considered Clear Creek Spring Upstream and Downstream, respectively, and are treated as one population in this study.
***Plants were not flowering at time of census, estimated >10,000 ramets.

During the current study's 2013-2014 leaf collections, a clump of ramets was treated as an "individual" for sampling purposes. Population sizes ranged from just a few sparse clumps to thousands of ramets forming a dense carpet (i.e. for the Nix Branch part of TN-G, and for GA-CC). As vegetative growth does not occur using rhizomes, sampled clumps were separated by at least 30 cm in an attempt to ensure that distinct genets were collected, often

leaving several unsampled clumps in between collections. Samples were collected throughout a population, with a minimum of 30 clumps per population sampled from populations where there were at least that many clumps. Whereas some populations were densely packed on hillside seeps or riparian zones, others followed along streambeds for several kilometers, in which case both upstream and downstream sections of the stream were sampled.

**DNA Extraction**

DNA extractions were performed following the DNeasy plant mini kit protocols (Qiagen, Valencia, CA) with fresh tissue for two TN populations (TN-G and TN-L), and the E.Z.N.A. SP Plant DNA kit protocols (Omega Bio-tek, Inc., Norcross, GA) with dry tissue for the remaining 12 populations. Using a Retsch (Newton, PA) MM301 bead mill, fresh leaf tissue was ground for five minutes in a microcentrifuge tube with one tungsten-carbide bead, whereas dry tissue was ground for 2.5 minutes with two beads. DNA elutions were stored at -80° C after concentrations and purity were determined using a NanoDrop-1000 Spectrophotometer (Thermo Scientific, Wilmington, DE). DNA was extracted from 399 plants (averaging 28.5 plants per population) and all these individuals were used in the microsatellite analysis.

**Laboratory Methods**

To develop a microsatellite library, a *X. tennesseensis* leaf from the type-locality (TN-L) was submitted to Steve Bogdanowicz of the Cornell University (Ithaca, NY) Evolutionary Genetic Core Facility (EGCF), where DNA for the creation of the genomic library enriched for microsatellites was extracted and the concentration was determined using a Qubit 2.0 fluorometer. The library was sequenced using an Illumina MiSeq (Illumina, Inc., San Diego, CA) at the Sequencing and Genotyping Facility at the Cornell Life Sciences Core Laboratory Center. The EGCF scanned for dimeric through hexameric repeats using Msatcommander 1.0.3 (Faircloth, 2008). The DNA (50-100ng) was digested using restriction enzymes *AluI*, *RsaI*, *Hpyl66II*. The digests were then combined for adenylation with Klenow (exo-) and dATP, after which ATP to 1mM was added and T4DNA ligase was used to ligate an Illumina Y-adaptor. Resulting fragments were enriched by hybridization to 3'-biotinylated probes and then collected using magnetic beads with a streptavidin-coat. Amplification was performed using Platinum *Taq* polymerase and two Illumina primers (a universal and an index). Successful amplification was verified by agarose gel electrophoresis and PCR product concentration quantified using the Qubit 2.0 fluorometer. This procedure provided more than 3000 possible microsatellite primer pairs, from which dimeric to tetrameric microsatellites ranging between 150 to 400 bp were selected.

Using DNA for two plants from each of twelve populations (all but TN-A and AL-Y) throughout the species' range, 50 primer pairs from the microsatellite library were surveyed to determine which markers could be amplified and resolved, and which of those that could be resolved were variable. Only those found to be polymorphic were used in this study. The protocol for fluorescent labeling of polymerase chain reaction (PCR) products generally

followed Schuelke (2000) as modified by Baskauf et al. (2014) who used an 18 bp M-13 tag (5'-CAC GAC GTT GTA AAA CGA-3'). The M-13 tag was attached to the fluorophore as well as the 5' end of the forward primer. In addition, a 7-base pair short tag (5'-GTTTCTT-3') was added to the 5' end of the reverse primer. The components of the 15μL PCR solution were as follows: 1.5 μL tricine buffer (300 mM tricine, 500 mM KCl, 20 mM $MgCl_2$), 0.06 μL dNTPs (25 mM each), 0.04 μM M-13 tagged forward primer (0.6 pmoles), 0.2 μM short-tagged reverse primer (3 pmoles), 0.2 μM M-13 tagged fluorophore, 0.03 units of *Taq* DNA polymerase (GoTaq Flexi DNA Polymerase, Promega, Madison, WI), and 1 μL DNA template (2-35 ng/μL). PCR conditions were as described in Baskauf et al. (2014).

PCR products were run on 1.2% agarose gels at 90V for 45 to 60 min to check for amplification. PCR products (i.e. amplified microsatellites) were placed in a 1:10 or 1:15 dilution with molecular biology grade USP sterile purified water (Mediatech, Inc., Manassas, VA), depending on the heights of peaks from previous runs. One microliter of this diluted microsatellite solution was added to a 10 μL mixture of GeneScan LIZ-500 size standard (Applied Biosystems, Foster City, CA) and Hi-Di formamide (Applied Biosystems, Foster City, CA), which were mixed in a 0.25 : 9.25 ratio following the manufacturer's protocol. The mixture of PCR product and size standard was then heated to 95°C for five minutes and shipped overnight on ice to Tom Cunningham at the University of Tennessee Health Science Center (Memphis, TN) for autosequencing on an ABI 3130XL DNA Analyzer (Applied Biosystems, Foster City, CA). Fragments were viewed using GeneMarker v1.97 (SoftGenetics, LLC, State College, PA, 2010) to identify alleles, which were verified by two people.

**Population Analyses**

Unless otherwise indicated, GenAlEx v6.502 (Peakall and Smouse 2012) was used for calculating population genetics statistics, including the percentage of polymorphic loci, levels of observed and expected heterozygosity, and Nei's (1978) genetic identities and genetic distances between populations. GENEPOP v4.5.1 (Rousset 2008) was used to carry out a chi-squared goodness-of-fit test to determine the significance of any deviations of genotype frequencies from Hardy-Weinberg expectations under equilibrium conditions, and linkage disequilibrium were calculated, using 10000 Markov Chain Monte Carlo permutations to test for significance, to identify any non-random associations between alleles from different loci. To adjust for multiple comparisons, the significance tests for linkage disequilibrium and Hardy-Weinberg were both corrected using a sequential Bonferroni method. Genetic structure was investigated using analysis of molecular variance (AMOVA) in GenAlEx, with the codominant allelic distance matrix input option, and the allelic distance matrix option was used in estimating Wright's *F*-statistics (Weir and Cockerham 1984) to determine the distribution of genetic variability within and among populations and regions (states, here). The standardized $F'_{ST}$ (Meirmans 2006) and pairwise $F_{ST}$ comparisons between populations were also estimated from the AMOVA. The estimator $G''_{ST}$ was calculated in the *G*-statistics option for codominant loci. Significance in AMOVA and $G''_{ST}$ were determined using 9999 Markov Chain Monte Carlo permutations. Using the non-standardized distance matrix input option, principal coordinate analysis generated a graphical representation of the genetic distance matrix derived from the data set.

In order to establish the number of genetic populations found in *X. tennesseensis* the Bayesian clustering software package STRUCTURE v2.3.4 (Pritchard et al. 2000) was

employed. By comparing their multilocus genotypes, individuals are assigned to *K* population genetic clusters without prior information about population membership being utilized for the analysis. Using the default parameter settings, 20 iterations per *K* value were carried out, examining *K*=1-16 clusters with a burn-in period of 200,000 and 1,000,000 Markov chain Monte Carlo repetitions. The Delta*K* method (Evanno et al. 2005) was used to identify the most appropriate number of genetic clusters for the data, using STRUCTURE HARVESTER (Earl and vonHoldt 2012) to graph likelihood values and Delta*K*. Probabilities were averaged across 20 runs at the best modal K using CLUMPP (Jakobsson and Rosenberg 2007) and a graphical interpretation of the results was generated with STRUCTURE PLOT (Ramasamy et al. 2014).

# CHAPTER III

## Results

Twenty-eight microsatellite loci were resolved, 14 (50%) of which were found to be polymorphic for the species. The average number of alleles per polymorphic locus was 3.4 at the species level. Microsatellite descriptions, including primer sequences, repeat motifs and range of allele sizes are detailed in Table 2, and population level allele frequencies for all loci are shown in Appendix 2.

Four loci pairs from among three populations (one pair of loci in GA-CC, one in AL-E, and two in AL-W) displayed significant linkage disequilibrium, after a sequential Bonferroni correction for multiple testing. As each of the four significant results involved different pairs of loci, these results do not indicate any consistent nonrandom associations among loci for this species. Chi-squared goodness-of-fit tests revealed that the genotype frequencies of the majority of polymorphic loci did not meet the expectations of Hardy-Weinberg equilibrium ($P<0.005$): 89% (8/9) in AL-E, 67% (6/9) in AL-W, 50% (1/2) in AL-Y, 75% (3/4) in GA-CC, and the one locus that is polymorphic for GA-H and AL-S (Appendix 1). It is unlikely these deviations are due to technical error as there was no missing data and all markers that produced excessive stutter peak were removed from the study.

Although only loci that were polymorphic at the species level were assayed for this study, not all of these loci were polymorphic in all populations. On average, only 16.3% of these loci were polymorphic in *X. tennesseensis* populations (Table 3). At the regional (state) level, Alabama populations displayed the highest levels of genetic variability while Tennessee populations had the lowest values for all genetic variability measures (alleles per locus, %

polymorphic loci, and heterozygosity; Table 3). Alabama populations averaged 37.5 % polymorphic loci, which was almost four to seven times higher than for Georgia and

Table 2. Characterization of microsatellites including primer sequence, repeat motif, number of alleles, allele size, and fluorophore tag utilized in this study.

| Locus | Primer Sequence (5'-3') | Repeat motif | No. of alleles | Size (bp) | Fluorophore |
|---|---|---|---|---|---|
| 33 | F-AGACTTGATAGCATCGTGGAGAC | $(ACAT)_5$ | 3 | 263-271 | 6FAM |
| | R-TCGATGCGCATGGTAATAAGAAG | | | | |
| 57 | F-GCAATTGGCGCAAATATATCTCG | $(ACC)_5$ | 2 | 346-351 | 6FAM |
| | R-AGGTATTGGCAGATCATCCTCTC | | | | |
| 153 | F-AACTCTGATAGAAGCTCCACTGG | $(AG)_{10}$ | 2 | 350-354 | NED |
| | R-ACCAGAATTTCATCCTCTTTCAGC | | | | |
| 175 | F-GATGCTTTGCTGAATGTTTCTGG | $(AG)_{16}$ | 5 | 200-208 | NED |
| | R-CATTGTTCACACTCGCATCAATG | | | | |
| 1066 | F-CGGAAGGAAGAGAAGAACATGTG | $(AG)_{13}$ | 4 | 385-391 | VIC |
| | R-ATGATCACAACCTCCCTAACTCC | | | | |
| 1245 | F-ATACTTGTCTCCTTTCTCTGGCG | $(AG)_6$ | 2 | 242-244 | PET |
| | R-ACGGATTTCTAGGGTTATTTGCTG | | | | |
| 1418 | F-CGGTTTGTAAATGCCATGAACAC | $(AG)_9$ | 4 | 322-330 | NED |
| | R-TATTCCAGTCGCGTTGAACTATC | | | | |
| 1558 | F-AACTCACTTAATCTTCCACGCAG | $(AG)_{14}$ | 3 | 215-219 | NED |
| | R-TCTTGGTCATATTGCTCTTCGTTG | | | | |
| 1570 | F-AAAGCGAAAGGATGGTCTTACG | $(AG)_{20}$ | 7 | 221-243 | 6FAM |
| | R-TTGCGAACATTCACTATCTCAGC | | | | |
| 1670 | F-TGTTTGCTTCCACTGATTGTCTC | $(ACT)_6$ | 2 | 344-347 | PET |
| | R-ATGCCCAAATCAAGACTAGGTTG | | | | |
| 1712 | F-ATTTCTATCTCTGCCTTGCGTTC | $(AG)_{13}$ | 4 | 230-236 | VIC |
| | R-CAACTGGGCTGTTAGGTTGAAC | | | | |
| 2032 | F-CTATGGAGTAGGATCAGCATGTG | $(AG)_9$ | 2 | 184-186 | PET |
| | R-TCTTTCTCGCGTACCTTCATATTG | | | | |
| 2846 | F-CAAGTATTCCACCATCTGCCATG | $(AG)_{20}$ | 6 | 205-234 | VIC |
| | R-CAGCACTCACTAAGCCACTTATG | | | | |
| 7809 | F-CTCGATCAAAGGCTCTGAATTCC | $(ACT)_7$ | 2 | 182-185 | PET |
| | R-TCAAATACAGCGCTTACTCAACC | | | | |

Table 3. Genetic variability averaged across 14 polymorphic loci for all sampled *X. tennesseensis* populations (standard error provided in parentheses). Mean sample size (N), average number of alleles per locus (A), private alleles ($A_P$), mean polymorphic loci (P), mean polymorphic loci with 95% occurrence ($P_{95}$), mean observed heterozygosity ($H_O$), unbiased (Nei, 1978) mean expected heterozygosity ($H_e$).

| State-Population | N | | A | | $A_P$ | | P | | $P_{95}$ | | $H_o$ | | $H_e$ | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TN-L | 30.0 | (0.00) | 1.1 | (0.07) | 0.0 | | 7.1 | | 0 | | 0.002 | (0.002) | 0.002 | (0.002) |
| TN-T | 30.0 | (0.00) | 1.1 | (0.07) | 0.0 | | 7.1 | | 7.1 | | 0.005 | (0.005) | 0.005 | (0.005) |
| TN-G | 30.0 | (0.00) | 1.1 | (0.07) | 0.0 | | 7.1 | | 7.1 | | 0.007 | (0.007) | 0.007 | (0.007) |
| TN-D | 30.0 | (0.00) | 1.1 | (0.01) | 0.0 | | 7.1 | | 7.1 | | 0.010 | (0.010) | 0.009 | (0.009) |
| TN-A | 18.0 | (0.00) | 1.0 | (0.00) | 0.0 | | 0.0 | | 0 | | 0.000 | (0.000) | 0.000 | (0.000) |
| **Mean TN** | **27.6** | | **1.06** | | **0.00** | | **5.68** | | **4.26** | | **0.005** | | **0.005** | |
| AL-E | 30.0 | (0.00) | 1.7 | (0.16) | 7.0 | | 64.3 | | 64.3 | | 0.071 | (0.021) | 0.251 | (0.057) |
| AL-W | 30.0 | (0.00) | 1.6 | (0.13) | 2.0 | | 64.3 | | 64.3 | | 0.095 | (0.029) | 0.279 | (0.058) |
| AL-S | 30.0 | (0.00) | 1.1 | (0.07) | 0.0 | | 7.1 | | 7.1 | | 0.000 | (0.000) | 0.035 | (0.035) |
| AL-Y | 30.0 | (0.00) | 1.1 | (0.10) | 0.0 | | 14.3 | | 14.3 | | 0.005 | (0.005) | 0.036 | (0.032) |
| **Mean AL** | **30.0** | | **1.39** | | **2.25** | | **37.5** | | **37.5** | | **0.043** | | **0.150** | |
| GA-H | 30.0 | (0.00) | 1.1 | (0.07) | 1.0 | | 7.1 | | 7.1 | | 0.000 | (0.000) | 0.028 | (0.028) |
| GA-M | 30.0 | (0.00) | 1.1 | (0.07) | 1.0 | | 7.1 | | 7.1 | | 0.005 | (0.005) | 0.005 | (0.005) |
| GA-C | 30.0 | (0.00) | 1.0 | (0.00) | 0.0 | | 0.0 | | 0.0 | | 0.000 | (0.000) | 0.000 | (0.000) |
| GA-CC | 30.0 | (0.00) | 1.3 | (0.13) | 0.0 | | 28.6 | | 28.6 | | 0.031 | (0.018) | 0.101 | (0.050) |
| GA-P | 21.0 | (0.00) | 1.1 | (0.07) | 1.0 | | 7.1 | | 0.0 | | 0.003 | (0.003) | 0.003 | (0.003) |
| **Mean GA** | **28.2** | | **1.10** | | **0.60** | | **10.0** | | **8.56** | | **0.008** | | **0.027** | |
| **Mean for species** (all populations) | **28.6** | (0.27) | **1.17** | (0.03) | **0.857** | (0.50) | **16.3** | (5.72) | **15.3** | (5.89) | **0.017** | (0.004) | **0.054** | (0.01) |

Tennessee, respectively. Observed ($H_o$=0.043) and expected ($H_e$=0.150) heterozygosity for Alabama populations averaged more than five times higher than for Georgia populations, whereas $H_o$ was almost 9 time higher and $H_e$ was 30 times higher for Alabama than for Tennessee populations ($H_o$=0.005 and $H_e$=0.005). The inbreeding coefficient, $F_{IS}$, reflected this overall lack of heterozygotes within populations of this species ($F_{IS}$ =0.707, estimated from AMOVA, $P$<0.001).

Of the 14 *X. tennesseensis* populations sampled, Alabama populations AL-E and AL-W were the most variable. Tennessee populations were all fixed for the same allele at 11 of the 14 loci, with TN-A being invariant for all loci. The other four Tennessee populations were each variable at just one of the remaining loci, having a second allele of relatively low frequency (2%-7%) at the variable locus. Four of the five populations in Georgia were fixed at either all of the loci (GA-C) or at all but one locus (GA-H, GA-M, GA-P). GA-CC was polymorphic at 4 loci, but no Georgia population had more than two alleles at any of the variable loci. Although Alabama populations tended to be more variable, AL-S was fixed for all but one locus and AL-Y for all but two. AL-E and AL-W were each polymorphic at 10 loci. With 3 alleles at locus 2846, AL-E was the only *X. tennesseensis* population to have more than two alleles at a variable locus. Among the 30 plants sampled from each of the two most variable populations, there were 24 unique multilocus genotoypes within AL-E and 28 within AL-W.

Divergence among populations accounted for approximately 91% of genetic variability in *X. tennesseensis* ($F_{ST}$ = 0.906, estimated from AMOVA, $P$<0.001). Including regions (states) in the AMOVA hierarchical analysis, differentiation among the states accounted for the

majority (60%) of the genetic variation in this species, with an additional 31% due to genetic differences among the populations within each state, and only 9% due to variability within populations. Both Meirmans' (2006) $F'_{ST}$ and Meirmans and Hedrick's (2011) $G''_{ST}$ were marginally higher at 95%.

While Tennessee populations did not possess any private alleles, three populations from Georgia and two from Alabama had one or more (Table 3). In Georgia, GA-H, GA-M, and GA-P each had one private allele at a locus (a different locus for each population). In Alabama there were nine private alleles split between AL-E (seven private alleles across six loci) and AL-W (two private alleles, one at each of two loci). As a region, Tennessee did have 7 unique alleles not seen in either Alabama or Georgia.

Estimates of Nei's (1978) unbiased genetic identity (Table 4) for pairs of Tennessee populations round to 1.000 (genetic distance of 0.000) because, with the exception of the few heterozygotes at three loci, Tennessee populations are genetically nearly identical. What little differentiation was available among these populations was more noticeable in the pairwise $F_{ST}$ comparisons (values ranging from 0.003 to 0.034) (Table 5). Alabama populations showed more divergence within their region, which is reflected in identity values ranging from 0.557 to 0.833 (genetic distance range of 0.183 to 0.586) and pairwise $F_{ST}$ values ranging from 0.189 to 0.856. Georgia populations displayed a similar spread of genetic identities within their region, ranging from 0.532 to 0.817 (genetic distance range of 0.203 to 0.632), although pairwise $F_{ST}$ showed the higher levels of differentiation between Georgia populations, ranging from 0.432 to 0.987. Overall, Georgia populations were more differentiated from Tennessee than from Alabama populations, with GA-P and GA-C being the least genetically similar to Tennessee according to both Nei's genetic identities and the

Table 4. Below the diagonal are Nei's (1978) unbiased genetic identity values; above are Nei's (1978) unbiased genetic distance values for each *Xyris tennesseensis* population pair.

| State-Population | TN-L | TN-T | TN-G | TN-D | TN-A | AL-E | AL-W | AL-S | AL-Y | GA-H | GA-M | GA-C | GA-CC | GA-P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **TN-L** | — | 0.000 | 0.000 | 0.000 | 0.000 | 0.771 | 0.967 | 0.903 | 1.160 | 0.877 | 1.019 | 1.252 | 0.832 | 1.931 |
| **TN-T** | 1.000 | — | 0.000 | 0.000 | 0.000 | 0.770 | 0.966 | 0.902 | 1.159 | 0.876 | 1.018 | 1.250 | 0.831 | 1.930 |
| **TN-G** | 1.000 | 1.000 | — | 0.000 | 0.000 | 0.766 | 0.963 | 0.910 | 1.170 | 0.884 | 1.027 | 1.262 | 0.829 | 1.954 |
| **TN-D** | 1.000 | 1.000 | 1.000 | — | 0.000 | 0.757 | 0.949 | 0.888 | 1.172 | 0.862 | 1.003 | 1.232 | 0.818 | 1.897 |
| **TN-A** | 1.000 | 1.000 | 1.000 | 1.000 | — | 0.772 | 0.968 | 0.904 | 1.162 | 0.878 | 1.021 | 1.253 | 0.834 | 1.932 |
| **AL-E** | 0.463 | 0.463 | 0.465 | 0.469 | 0.462 | — | 0.221 | 0.183 | 0.547 | 0.423 | 0.975 | 1.033 | 0.556 | 1.037 |
| **AL-W** | 0.380 | 0.381 | 0.382 | 0.387 | 0.380 | 0.802 | — | 0.296 | 0.413 | 0.508 | 0.891 | 0.914 | 0.592 | 0.859 |
| **AL-S** | 0.405 | 0.406 | 0.403 | 0.412 | 0.405 | 0.833 | 0.743 | — | 0.586 | 0.352 | 0.752 | 0.759 | 0.518 | 0.761 |
| **AL-Y** | 0.313 | 0.314 | 0.310 | 0.310 | 0.313 | 0.579 | 0.661 | 0.557 | — | 0.741 | 0.721 | 1.312 | 0.938 | 0.874 |
| **GA-H** | 0.416 | 0.416 | 0.413 | 0.422 | 0.415 | 0.655 | 0.602 | 0.703 | 0.477 | — | 0.506 | 0.512 | 0.203 | 0.514 |
| **GA-M** | 0.361 | 0.361 | 0.358 | 0.367 | 0.360 | 0.377 | 0.410 | 0.471 | 0.486 | 0.603 | — | 0.436 | 0.264 | 0.563 |
| **GA-C** | 0.286 | 0.286 | 0.283 | 0.292 | 0.286 | 0.356 | 0.401 | 0.468 | 0.269 | 0.599 | 0.647 | — | 0.307 | 0.242 |
| **GA-CC** | 0.435 | 0.436 | 0.437 | 0.442 | 0.434 | 0.574 | 0.553 | 0.596 | 0.392 | 0.817 | 0.768 | 0.736 | — | 0.632 |
| **GA-P** | 0.145 | 0.145 | 0.142 | 0.150 | 0.145 | 0.355 | 0.424 | 0.467 | 0.417 | 0.598 | 0.570 | 0.785 | 0.532 | — |

Table 5. Pairwise comparison of $F_{ST}$ between populations below the diagonal, significance values above.

| State-Population | TN-L | TN-T | TN-G | TN-D | TN-A | AL-E | AL-W | AL-S | AL-Y | GA-H | GA-M | GA-C | GA-CC | GA-P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| TN-L | — | 1.000 | 0.200 | 0.200 | 1.000 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 |
| TN-T | 0.003 | — | 0.060 | 0.140 | 0.550 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 |
| TN-G | 0.021 | 0.022 | — | 0.100 | 0.290 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 |
| TN-D | 0.029 | 0.029 | 0.031 | — | 0.150 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 |
| TN-A | 0.008 | 0.017 | 0.026 | 0.034 | — | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 |
| AL-E | 0.652 | 0.650 | 0.647 | 0.643 | 0.655 | — | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 |
| AL-W | 0.656 | 0.654 | 0.652 | 0.648 | 0.659 | 0.220 | — | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 |
| AL-S | 0.940 | 0.936 | 0.933 | 0.929 | 0.943 | 0.346 | 0.417 | — | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 |
| AL-Y | 0.946 | 0.943 | 0.940 | 0.937 | 0.949 | 0.561 | 0.481 | 0.857 | — | 0.010 | 0.010 | 0.010 | 0.010 | 0.010 |
| GA-H | 0.950 | 0.946 | 0.943 | 0.939 | 0.954 | 0.522 | 0.529 | 0.820 | 0.887 | — | 0.010 | 0.010 | 0.010 | 0.010 |
| GA-M | 0.989 | 0.986 | 0.982 | 0.979 | 0.993 | 0.682 | 0.643 | 0.929 | 0.925 | 0.923 | — | 0.010 | 0.010 | 0.010 |
| GA-C | 0.997 | 0.994 | 0.991 | 0.988 | 1.000 | 0.694 | 0.651 | 0.937 | 0.952 | 0.934 | 0.987 | — | 0.010 | 0.010 |
| GA-CC | 0.841 | 0.837 | 0.834 | 0.830 | 0.844 | 0.504 | 0.492 | 0.737 | 0.807 | 0.576 | 0.680 | 0.717 | — | 0.010 |
| GA-P | 0.993 | 0.991 | 0.988 | 0.986 | 0.996 | 0.691 | 0.639 | 0.932 | 0.935 | 0.927 | 0.982 | 0.985 | 0.812 | — |

pairwise $F_{ST}$ comparisons. Alabama populations were also less similar to Tennessee than to Georgia populations. Comparing Tennessee with the other two states, AL-E, AL-W, and GA-CC are the populations most similar to Tennessee populations (although specific ranking differs between genetic identities vs. pairwise $F_{ST}$). Comparing Alabama and Georgia populations, AL-S and GA-H have the highest genetic identity values, whereas AL-W and GA-CC have the lowest $F_{ST}$ values.

Principal coordinate analysis (Fig. 4) displayed many of the same trends shown by Nei's (1978) unbiased genetic identity. The first axis of principal coordinate analysis accounted for 48.6% of variation in the data, placing the Tennessee populations on one side, with Georgia and the majority of Alabama populations together on the other side. The second axis explained 22.2% of the variation separating Alabama (below the axis) from Georgia and Tennessee (both above the axis). Symbols representing the individuals of Tennessee populations are clustered tightly together, while the more differentiated populations of Georgia and Alabama are spread more widely. The lack of within-population diversity can be seen for all Tennessee and most Georgia populations, as well as an Alabama population, with individuals so genetically similar that population symbols overlap so as to appear only once or twice. In contrast, AL-E and AL-W in Alabama and GA-CC in Georgia have much less overlap among individuals, indicating higher within-population differentiation.

STRUCTURE results (Fig. 5 and 6) reflect the regional grouping evident in the principal coordinate analysis. The modal value of Delta*K*, derived from the log likelihood plot (example in Fig. 5a), was K=3 for nine of 10 STRUCTURE analyses performed, often with a singular high peak (example in Fig 5b), indicative of the strength of the support that this is

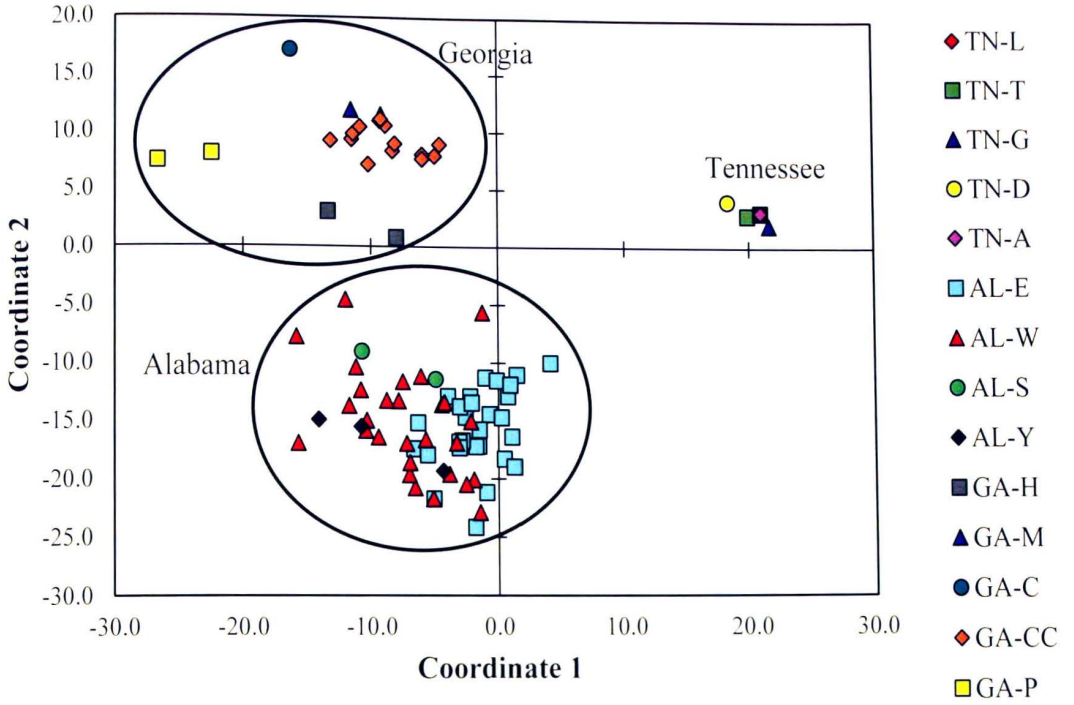
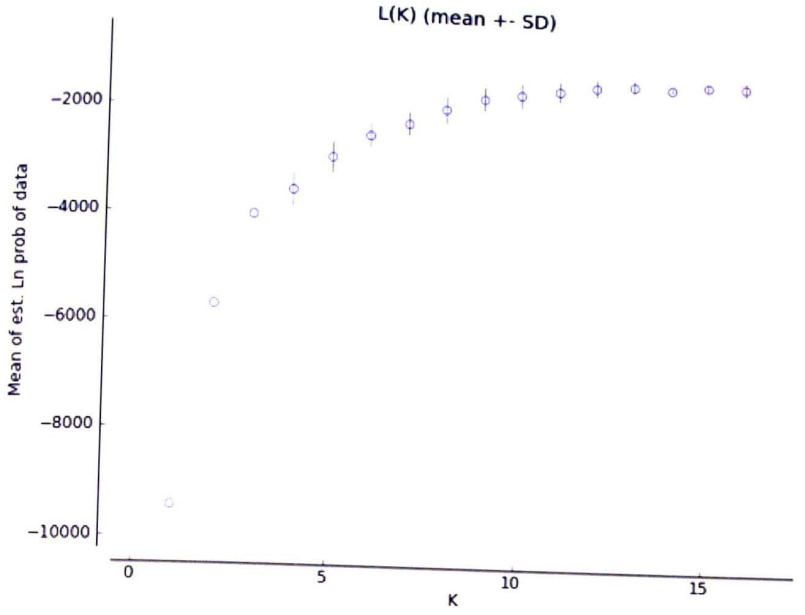



Figure 4. Principal coordinate analysis using the genetic distance matrix, without data standardization, for *Xyris tennesseensis* individuals. Population abbreviations are given in the legend. Coordinate 1 explains 48.5% and coordinate 2 explains 22.2% of the variation (70.7% total).

the true K (Evanno et al. 2005). These 3 clusters represented the 3 states, suggesting the populations found in each state combine to form 3 genetic populations. The next most likely hypothesis by the Delta*K* method was K=2, although this was not nearly as well supported as K=3. With only two clusters, Alabama and Georgia populations are grouped together while Tennessee populations are assigned to a separate cluster, a reflection of the fact that populations from these two states are more genetically similar to each other than they are to populations in Tennessee. The log likelihood plot (Fig. 5a) shows that log likelihood values plateau at high K values (e.g. around K 8-10), with none of them significantly different from each other.

(a)




(b)

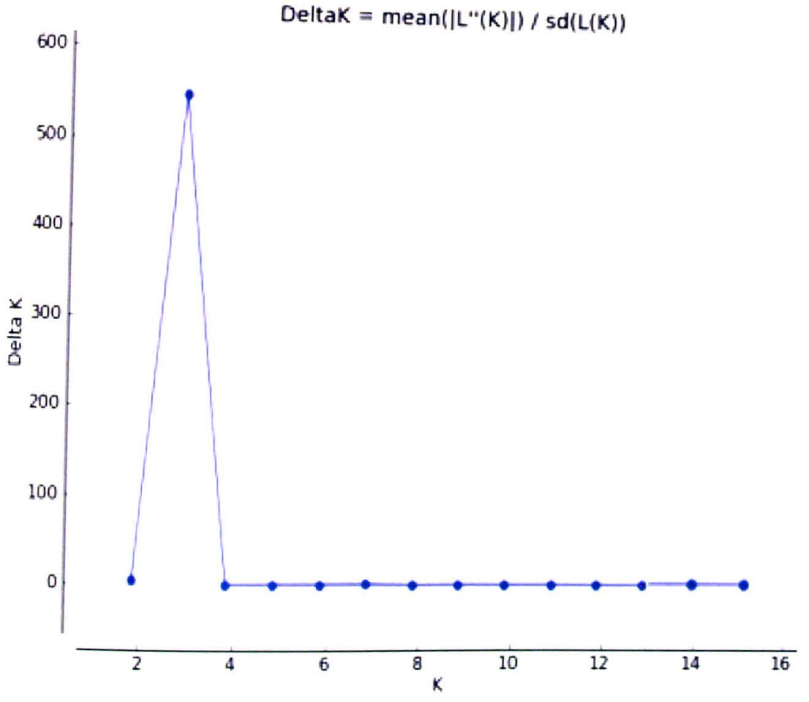



Figure 5. Examples of typical output from STRUCTURE HARVESTER (a) Mean estimated log likelihood (± SD) for 20 STRUCTURE runs for hypotheses K=1-16. (b) Delta*K*, the second order increase in likelihood for K=2-15.

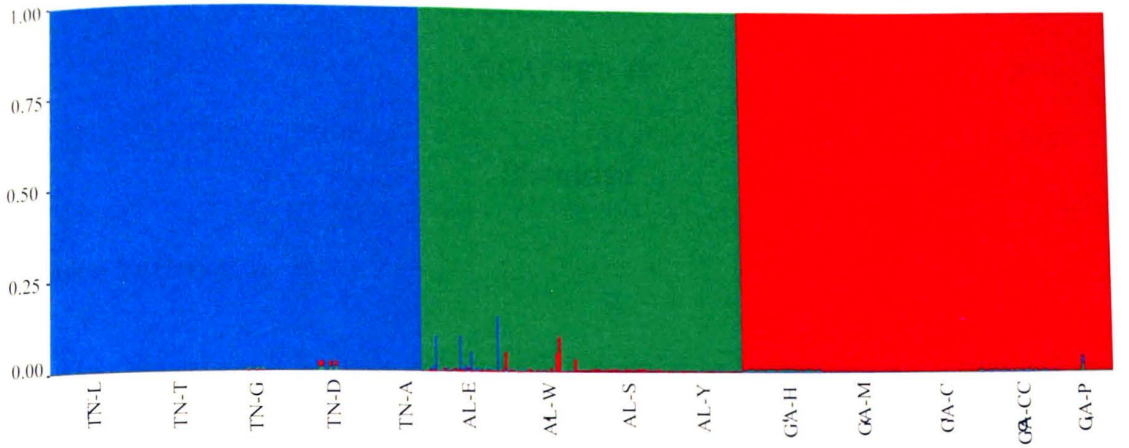



Figure 6. Genetic structure of *X. tennesseensis* based on the average of nine trials of 20 runs if $K$=3 using the software STRUCTURE.

# CHAPTER IV

## Discussion

### Genetic Variation

*Xyris tennesseensis* has very little genetic diversity at 14 polymorphic microsatellite loci. Population-level polymorphism averaged only 16%, with about 1.2 alleles per locus for the species. Tennessee populations had almost no variability, with only a few heterozygotes at the three loci that were polymorphic for the region. Georgia populations, with the exception of GA-CC, had similarly low levels of polymorphism, although Georgia populations did differ from each other more than did the nearly uniform Tennessee populations. Although AL-W and AL-E, the most genetically diverse of all the *X. tennesseensis* populations, were both invariant at only five of the 14 loci, AL-S and AL-Y were fixed for nearly all loci. Very few alleles were observed per population for those loci that were polymorphic (out of 14 *X. tennesseensis* populations, only AL-E, had more than 2 alleles at a polymorphic locus), and observed levels of heterozygosity were extremely low in all populations. Such low levels of genetic variability in *X. tennesseensis* populations suggest that many of these populations may have experienced severe genetic bottlenecks at some point in time, resulting in a loss of alleles.

It is very likely that many *X. tennesseensis* populations may experience periodic population genetic bottlenecks due to the flooding events that can occur for streamside habitats. With the combination of loose, mucky soil and relatively shallow roots, *X. tennesseensis* individuals can be easily uprooted with sufficient force. Fast moving flood

waters could erode soil and dislodge entire plants, and if a large portion of the population was destroyed in this manner, such flooding events could cause severe genetic bottlenecks, with a resultant loss of alleles. Low diversity populations AL-S, GA-C, TN-A, TN-D, and part of TN-G are likely frequently subjected to flooding conditions as they are found very close to the edge of streams. TN-T is on a steep hillside beside the stream, thus it is likely that only the lower portions of the population would be affected by flooding. Drought could be a stronger environmental pressure for GA-P, TN-L, TN-T, and the section of TN-G not located next to the stream. The water source for these populations is almost strictly ground water, thus extended droughts may lead to loss of individuals. In 2000, a severe genetic bottleneck occurred in AL-Y when the water-source, a nearby spring, failed and the majority of the population was lost (Moffett 2008).

The most diverse populations may not be as susceptible to flooding and drought. AL-E, one of the two most diverse of all sampled populations, is on the periphery of a swamp near the mouth of a spring. Though flooding may occur, it would result in standing water, instead of the faster currents that may be seen in streams, and the water-source at this site may be reliable enough than droughts are less common than in sites such as AL-Y. The other of the two most diverse populations, AL-W, surrounds a spring-fed pond on flat ground and thus is also not likely to experience violent flooding events that could carry away large portions of the population. The most diverse Georgia population (and the largest population sampled), GA-CC, is again found close to a springhead on flat ground.

Even for a rare endemic, diversity levels as low as those seen in *X. tennesseensis* are unusual in studies utilizing microsatellites markers, although Peakall et al. (2003) found no

genetic variability at all at microsatellite loci for the relictual Australian conifer, *Wollemia nobilis*. However, many rare species have more variability. For example, among recent studies of rare herbaceous endemics, *Primula boveana* (Jimenez et al. 2014), *Helianthus verticillatus* (Ellis et al. 2006) and two species of *Limnanthes* (Sloop et al. 2011; Sloop et al. 2012) all had polymorphism levels that were over two to six times higher than those of *X. tennesseensis*. The average number of alleles per locus for both *Limnanthes* species and *H. verticillatus* was 3 to 10 times higher than for *X. tennesseensis* populations, although the average for *P. boveana* was only slightly higher, at 1.42 alleles per locus. The average polymorphism of another endangered southeastern U.S. endemic, *Boechera perstellata*, was 28% (Baskauf et al. 2014), still nearly two times higher than *X. tennesseensis*, although the average alleles per locus was similar at 1.3.

Expected heterozygosity is also unusually low for *X. tennesseensis*. In a review of 106 microsatellite studies, Nybom (2004) reported expected heterozygosity levels for various categories of plant species, including ones with breeding systems and ecological and life history strategies similar to *X. tennesseensis*: endemics (5 species), selfing species (4), mixed breeding species (13), short-lived perennials (20), and species that utilized wind/water seed dispersal (26). The average expected heterozygosity for species in each of these categories was at least 3 to 12 times higher than that of *X. tennesseensis*, as were the mean expected heterozygosity levels reported by Sloop et al. (2011) and Sloop et al. (2012). Expected heterozygosities were more than 8 times higher for *Helianthus verticillatus* (Ellis et al. 2006) and *H. porteri* (locally abundant on a rare habitat) (Gevaert et al. 2013) than for *X. tennesseensis*. On the other hand, expected heterozygosity for *Boechera perstellata*

($H_e$=0.059, Baskauf et al. 2014) and *Primula boveana* ($H_e$=0.085, Jimenez et al. 2014) were more comparable to *X. tennesseensis* ($H_e$=0.054).

Most *X. tennesseensis* populations had extremely low levels of observed heterozygosity at most polymorphic loci (average $H_o$=0.017). Average observed heterozygosity levels in each of the five aforementioned categories in Nybom's (2004) review were at least 20 times higher than that of *X. tennesseensis*, with the exception of the selfing species category, for which values were still 3 times higher. Observed heterozygosity for *X. tennesseensis* deviated significantly from Hardy-Weinberg equilibrium in 20 out of 32 tests, resulting in a high value for the inbreeding coefficient ($F_{IS}$=0.707). Boyd et al. (2011) indicate that this species is able to self successfully, and although it has a mixed breeding system, the significant deficit of heterozygotes displayed by these populations suggests that the species engages in much more inbreeding than outcrossing. The inbreeding coefficient for *Boechera perstellata* was even higher ($F_{IS}$=0.933, Baskauf et al. 2014), as was that of *Primula boveana* ($F_{IS}$=0.862, Jimenez et al. 2014) and *Limnanthes floccosa* ($F_{IS}$=0.82, Sloop et al. 2011). All three authors attributed the high $F_{IS}$ values to a potentially high prevalence of selfing, as may be the case for *X. tennesseensis*. Jimenez et al. (2014) postulates that there may have been a historical reliance on selfing when pollinators were scarce. It is possible that pollinator scarcity could also affect the rate of selfing in *X. tennesseensis*.

Despite their relatively low variability in *X. tennesseensis*, microsatellite markers were more variable and afforded better resolution for the population structure of this species than did isozymes. Only 25% (2/8) of assayed isozymes were polymorphic at the species level in Willis' survey (unpubl. data), whereas 50% of 28 microsatellites were polymorphic at the

species level. At the population level, there were no polymorphic isozyme loci. Thus while observed and expected heterozygosity levels were low at microsatellite loci, at isozyme loci $H_o$ and $H_e$ were both zero due to all isozyme loci being fixed for a single allele within each population. There was no consistent pattern of fixations when comparing isozyme and microsatellite results (AL-Y and GA-C were fixed for the same less common allele at one isozyme locus and AL-W was fixed for a unique allele at the second polymorphic isozyme). Although it is common for there to be less variability at isozyme loci than at microsatellite loci, the Willis study revealed unusually low genetic variability in *X. tennesseensis*, even for isozyme loci, and even compared to other endemic species (see review of plant isozyme studies in Hamrick and Godt, 1989).

**Population Structure**

Geographically isolated populations with limited means of seed dispersal tend to be more highly differentiated (Hamrick and Godt 1996). *Xyris tennesseensis* does have very high population level differentiation, as is reflected in its high $F_{ST}$ value ($F_{ST}$=0.906,). This value is unusually high when compared with many other species, including other species with low variability. $F_{ST}$ was two times higher than estimates included for selfing species in Nybom's (2004) review of microsatellite studies, and was three to seven times higher than values reported for the other associated categories such as endemics, short-lived perennials and mixed breeding. Keeping in mind that caution is needed when comparing across species or loci with very different levels of diversity, comparisons with other low diversity species still indicates that *X. tennesseensis* has unusually high levels of population differentiation.

The annual vernal pool endemic, *Limnanthes floccosa* sp. *californica*, possessing the same mixed mating system, with a prevalence of selfing, and utilizing water for seed dispersal much like *X. tennesseensis*, has a lower $F_{ST}$ ($F_{ST}$=0.65, Sloop et al. 2011). *Primula boveana*, despite having experienced what the authors postulate were likely severe genetic bottlenecks at the only four existing populations in addition to relying largely on selfing, still has lower among-population variation ($F_{ST}$=0.737, Jimenez et al. 2014). However, *Boechera perstellata* has the same level of differentiation among its populations ($F_{ST}$= 0.906, Baskauf et al. 2014) as *X. tennesseensis*. Like *X. tennesseensis*, *B. perstellata* is another species which has very low genetic diversity and a disjunct distribution, and which the authors suspect to be primarily selfing. It should be noted that none of the above studies report $F'_{ST}$ values.

Principal coordinate analysis illustrated that a significant amount of the variation in the species was due to regional (state) differences and that Tennessee was the most divergent region (Fig. 4). Despite the lack of a consistent pattern between genetic similarity and geographic proximity for specific population pairs, the average genetic identity between Alabama and Georgia populations (0.472) was higher than the average between Tennessee and either Alabama (0.392) or Georgia (0.329). It is possible that at one time the distribution of the seepage fen habitat preferred by *X. tennesseensis* was more continuous between Alabama and Georgia, allowing gene flow to occur and leading to higher levels of genetic similarity. There may have been a more continuous distribution throughout Tennessee and Alabama as well, although genetic identities suggest that the separation of these two regions took place long before gene flow stopped among Georgia and Alabama populations. With the Appalachian Mountains acting as a geographic barrier between Tennessee and Georgia populations, it is perhaps unsurprising that comparisons between these regions give the

lowest average genetic similarity. Currently, the largest gap between any two neighboring sampled populations within the Alabama-Georgia band is ~120 km (from AL-E to AL-Y). In contrast, the shortest distance from any Tennessee population to an Alabama or a Georgia population within the band is 289 km (TN-A to AL-W) or 325 km (TN-A to GA-C), respectively, which is similar to the geographical span between the most distant populations in the Alabama-Georgia band (302 km from AL-S to GA-P). Over time, with continued habitat destruction and fragmentation due to anthropogenic activities, genetic drift and increased isolation may lead to further differentiation among the Georgia and Alabama populations.

The Delta*K* method of Evanno et al. (2005) typically detects the highest level of structure in a data set, which is the regional (state) level in the case of K=3 for this study. It should be noted that Pritchard et al. (2010) cautioned that inbreeding, which the data suggests occurs in all of the sampled *X. tennesseensis* populations, could lead to an overestimation of K by STRUCTURE. Thus, for this species which appears to rely heavily on inbreeding, it may not be surprising that increasingly higher K values displayed in the log likelihood plot (Fig. 5a) appear as more likely hypotheses than the best modal K of 3. It should also be noted that although the term "genetic population" is used to refer to genetically similar clusters formed by analyses such as STRUCTURE, there are various genetic and ecological definitions of "population" (Waples and Gaggiotti 2006). Of course, genetic differences exist among the populations even within states, and with their disjunct distribution the likelihood of gene flow among them is extremely low. While a "K=3" STRUCTURE solution indicates the relatively higher similarity of populations within a state than among the states, it does not negate the fact that it is extremely unlikely that gene flow is currently occurring among most

populations within a region, or that it has occurred for a very long time in the past, considering the geographic distances between the populations and the fact that many populations within Alabama or Georgia have private alleles and are fixed or nearly fixed for different alleles at some loci.

## Conservation Implications

Habitat loss is a major threat to *X. tennesseensis* populations, contributed to by improper or insufficient management of *X. tennesseensis* sites on both private and public lands. In 2008, the privately-owned GA-M was determined to be the largest Georgia population (Moffett 2008). However, during the 2014 leaf collection at this site, it appeared that much of the population was destroyed due to the owner using a back-hoe to scrape out the stream and the surrounding riparian zone in an attempt to make the stream less hospitable to reptiles such as *Agkistrodon piscivorus*. Likewise, even though the GA-P population is located in a wildlife management area it was nearly extirpated by wild hogs between the 2008 census and the 2014 collection. This population, which once hosted many more ramets (Boyd and Moffett 2010), was reduced to a mere 21 ramets, none of which had flowering stalks.

Only six sampled *X. tennesseensis* sites are located on public lands (TN-A, TN-D, AL-W, AL-Y, GA-H and GA-P), where they are afforded some degree of protection, at least from human activities. The private owners of AL-E, TN-L, TN-G either protect the site or have agreements with government agencies to permit monitoring. However, the largest and most diverse Georgia population (GA-CC) is on privately owned land, and the upstream and downstream portions are owned by different individuals. The landowner of the upstream

portion, where the vast majority of the population is concentrated, does actively protect the population and in 2014 agreed to refrain from mowing so close to the stream during the flowering season for *X. tennesseensis* (M. Moffett, pers. comm.). The much smaller downstream portion of the population is threatened because it surrounds a watering hole in a cattle field, and that land owner is unlikely to make changes to protect the plant. The remainder of the *X. tennesseensis* populations occur on privately owned lands where landowners have not made any official agreements to protect the plants.

In the interest of preserving the genetic diversity of the species, more populations need to be protected. Those which are currently on located on public lands should be more closely monitored and maintenance efforts, such as invasive species removal (i.e. wild hogs, competing invasive plants) and thinning of encroaching native plants, should be carried out for all populations found on federal or otherwise protected land. In two Tennessee populations, TN-A and TN-T, vegetation of the threatened *Parnassia grandifolia* is very thick and could potentially overcrowd *X. tennesseensis*, preventing vegetative spread and root growth (D. Estes, pers. comm.). *Impatiens capensis* growth found on either side of the TN-G hillside population may eventually begin to shade out *X. tennesseensis* individuals found on the periphery (A. Bishop, pers. comm.). Competition through shading is also occurring in the GA-H population in Georgia. At one point over 50% of GA-H was in full shade due to overhanging branches of *Hypericum interior* and *X. tennesseensis* individuals in the population had not bloomed for two years. Clipping back the over-hanging branches by hand is the most effective way to remove the *Hypericum* shrubs without damaging the habitat or the existing *X. tennesseensis* plants, although it is a labor-intensive process. In the first year following treatment there was a significant increase in the number of flowering spikes,

but a decline in the next two years, most likely because the shrubs began to grow back and other herbaceous species were overcrowding the *X. tennesseensis* clusters (Moffett 2008). Any future management plans should consider this when scheduling overstory trimming.

Results from this study indicate that continued protection of the AL-E and AL-W populations in Alabama should be a high priority. Not only are these the largest populations in Alabama, but they also have the highest levels of genetic diversity of any sampled *X. tennesseensis* populations. Fortunately, both AL-E and AL-W occur on protected lands, with AL-E owned by the University of Montevallo and AL-W found on Fort McClellan Military Reservation. They should continue to be regularly monitored and surrounding shrubbery, such as the dense thickets found lining the spring at AL-E, should be pruned if competing plants threaten to crowd out patches of *X. tennesseensis*. The GA-CC population in Georgia ranked third for in terms of genetic diversity. Open communication should be maintained with the owner of the upstream portion of GA-CC to protect this largest and most genetically diverse Georgia population. The fourth most variable site, GA-H, located on protected public lands, receives regular maintenance to keep the over-hanging *Hypericum* limbs trimmed back.

Although the Tennessee populations have lower levels of genetic variability, these populations are the most genetically divergent group of populations, having seven alleles at 6 (43%) of the loci that are unique or nearly unique to the Tennessee populations. Whether or not these genetic differences correlate with unidentified regional ecotypes is unknown. Among Tennessee's populations, TN-T, TN-G, and TN-D (one of the only Tennessee population protected on publicly owned land) are the largest, and they are also the top three

in terms of heterozygosity levels. Although TN-T is not the most variable population by far, it is a fairly large population that is reported to have the highest quality *X. tennesseensis* habitat of all the populations in all three regions (Boyd and Moffett 2010).

## Future Research

Because it is impossible to determine in the field whether a clump is one genetic individual or several, it is difficult to accurately estimate population sizes. Thus, it would be useful to determine how common cloning is and how likely it is that clumps consist of a single clone. It should be noted that the microsatellites included in this study were not variable enough to address this question. In future studies, it would be beneficial to employ different genetic markers that are much more variable (either more variable microsatellites if such exist, or some different kind of marker), so that the prevalence of clonal growth in this species could be assessed by subsampling clusters of ramets. It would be best to include the most genetically diverse populations in such a study, as populations that are predominantly monomorphic at microsatellite markers used for this study may not provide any discernable variation between clusters even if more variable microsatellites or other types of markers are employed. The population in Franklin County, AL, should also be sampled in a future study, as there is currently no knowledge of its genetic diversity.

With the loss of wetlands across the southeastern United States due to anthropogenic activities, forces such as random genetic drift and population fragmentation are likely affecting multiple species within these habitats, not just *X. tennesseensis*. A study including

several species in this wetland habitat may reveal that multiple species have similarly low levels of genetic diversity, lowering the overall genetic diversity and possibly the productivity of the community (Booth and Grime 2003). Determining which species, if any, are genetically depauperate and assessing what may be the cause of the low diversity may also allow conservationist to more accurately ascertain what pressures, environmental or otherwise, the community is experiencing.

Many factors beyond rarity such as geographic isolation, breeding system, seed dispersal, and life form can have effects on genetic diversity (Hamrick and Godt 1989, Gitzendanner and Soltis 2000, Nybom 2004). It would be informative to compare the variability of *X. tennesseensis* to a widespread congener such as *X. torta* to determine the presence and possible significance of any correlation between the rarity of *X. tennesseensis* and its low genetic variability. The widespread species thrives in habitats very similar to that of *X. tennesseensis*, with the exception that the soil is more acidic than circum-neutral.

At the end of the flowering season, it would be useful to collect seeds from cones of the most genetically diverse populations to perform a greenhouse study on the ecotypes of those populations, perhaps determining if individuals from the Southern Ridge and Valley Seepage Fens are able to thrive in soil from the Western Highland Rim Seepage Fens and vice versa. In addition, greenhouse cross-breeding experiments could evaluate whether interpopulation crosses among long-isolated populations would be harmful (e.g. result in outbreeding depression) or beneficial (e.g. improve fitness-related traits such as seed production or seedling survival). Determining the effects of such crosses would be crucial to deciding if offspring from a mixture of populations should be used to create an artificial population or

augment an existing one. The Missouri Botanical Garden and Atlanta Botanical Garden are currently storing seeds, though not from all populations. Future seed collections should include the cones of the populations with the highest diversity (in particular AL-E and AL-W), and those hosting unique alleles.

# CHAPTER V

## Conclusions

The rare plant, *Xyris tennesseensis* has very low genetic diversity at microsatellite loci — even compared with other rare species — with low percent polymorphic loci, number of alleles per locus, and observed and expected heterozygosity. Such low levels of diversity could put this rare species at a higher risk for extinction by lowering fitness in the short-term and impairing the species' ability to adapt in the long-term. Thus it is especially important to preserve as much genetic variability as possible in this species. Despite having the lowest genetic variation of the three regions, Tennessee populations are genetically distinct from Alabama and Georgia, with several unique alleles in the region, and the state hosts one of the highest quality sites for the species. Therefore, these populations warrant monitoring and invasive and/or competing species-removal maintenance. Georgia populations show limited variability, with only 4 polymorphic loci in the region, but several populations have private alleles, and the region boasts the largest and best maintained site at this point (GA-CC). Alabama is the region with some of the most genetically variable populations, although the majority of that variation could be attributed to only two of the four sampled populations. Fortunately, those two populations occur on protected land, and hopefully they will continue to be protected and managed properly. Efforts should also be focused on protecting Georgia and Tennessee populations in order to preserve as much genetic diversity as possible at both the species and population levels.

Habitat loss is the most immediate threat to this species' survival, so much so that the habitat itself is considered imperiled (NatureServe 2013). Efforts must be made toward

habitat preservation and removal of shading competition in extant sites. Surveys should be conducted to locate suitable sites for future artificial populations while taking any ecotypic patterns into consideration, perhaps beginning with the known Highland Rim Seepage Fens sites in Tennessee. Attention should be paid to which populations, if any, are combined in these new sites to prevent outbreeding depression.

## LITERATURE CITED


Barrett, S.C.H., J.R. Kohn. 1991. Genetic and evolutionary consequences of small population size in plants: implications for conservation. *Genetics and Conservation of Rare Plants*, D.A. Falk and K.E. Holsinger (eds.), pp. 3–30. Oxford University Press, New York, NY. USA.

Baskauf, C.J., N.C. Jinks, J.R. Mandel, D.E. McCauley. 2014. Population Genetics of Braun's Rockcress (*Boechera perstellata*, Brassicaceae), an Endangered Plant with a Disjunct Distribution. *Journal of Heredity* 104: 124-137.

Baskin, C.C., J.M. Baskin. 2003. Seed germination and propagation of *Xyris tennesseensis*, a federal endangered wetland species. *Wetlands* 23: 116-124.

Beardmore, J.A. 1983. Extinction, survival, and genetic variation. In: *Genetics and Conservation*, C.M. Schoenmwald-Cox, S.M. Chambers, B. MacBryde, and L. Thomas (eds.) pp. 125-151. Benjamin-Cummings, Menlo Park, CA., USA.

Booth, R.E., J.P. Grime. 2003. Effects of genetic impoverishment on plant community diversity. *Journal of Ecology* 91:721-730.

Boyd, R.S., A. Teem, M.A. Wall. 2011. Floral biology of an Alabama population of the federally endangered plant, *Xyris tennesseensis* Kral (Xyridaceae). BioONE 76(3):255-265.

Boyd, R.S., J.M. Moffett. 2010. Population survey of the federally endangered Tennessee Yellow-Eyed Grass (*Xyris tennesseensis*). Unpublished report to the U.S. Fish and Wildlife Service, Jackson, Mississippi. 13 pp.

Brown, A.H.D. 1990. Genetic characterization of plant mating systems. In: *Plant Population Genetics, Breeding, and Genetic Resources*, A.D.G Brown, M.T. Clegg, A.L. Kahler,

and B.S. Weir (eds.), pp. 145-162. Sinauer Associates Publishing, Sunderland, MA, USA.

Cole, C.T. 2003. Genetic variation in rare and common plants. *Annual Review of Ecology, Evolution and Systematics* 34:213-237.

Earl, Dent A. and vonHoldt, Bridgett M. (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources vol. 4 (2) pp. 359-361 doi: 10.1007/s12686-011-9548-7.

Ellis, J.R., C.H. Pashley, J.M. Burke, D.E. McCauley. 2006. High genetic diversity in a rare and endangered sunflower as compared to a common congener. *Molecular Ecology* 15:2345-2355.

Ellstrand, N.C., D.R. Elam. 1993. Population genetic consequences of small population size: implications for plant conservation. *Annual Review of Ecology Systematics* 24:217-242.

Epperson, B.K. 1989. Spatial patterns of genetic variation within plant populations. In: *Plant Population Genetics, Breeding, and Genetic Resources*, A.D.G Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir (eds.), pp. 229-253. Sinauer Associates Publishing, Sunderland, MA, USA.

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

Frankham, R. 1995. Conservation genetics. *Annual Review of Genetics* 29:305-320.

Fukunaga, K., J. Hill, Y. Vigouroux, Y. Matsuoka, J.G. Sanchez, K. Liu, E.S. Buckler, J. Doebley. 2005. Genetic diversity and population structure of teosinte. *Genetics* 169:2241-2254.

Gathmann, A., T. Tscharntke. 2002. Foraging ranges of solitary bees. *Journal of Animal Ecology* 71:757-764.

Gevaert, S.D, J.R. Mandel, J.M. Burke, L.A. Donovan. 2013. High genetic diversity and low population structure in Porter's Sunflower (*Helianthus porter*). *Journal of Heredity* 104:407-415.

Gitzendanner, M.A., P.S. Soltis. 2000. Patterns of genetic variation in rare and widespread plant congeners. *American Journal of Botany* 87:783-792.

Hamrick, J.L., M.J.W. Godt. 1989. Allozyme diversity in plant species. In: *Plant Population Genetics, Breeding, and Genetic Resources*, A.D.G Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir (eds.), pp. 43-63. Sinauer Associates Publishing, Sunderland, MA, USA.

Hartl, D.L. 1981. *A primer of population genetics*. Sinauer Associates Publishing, Sunderland, MA, USA.

Hartl, D.L., A.G. Clark. 2007. *Principles of Population Genetics*. Sinauer Associates Publishing, Sunderland, MA, USA.

Hedrick, P.W. 2011. *Genetics of populations*. Jones and Bartlett Publishers, Sudbury, MA, USA.

Hedrick, P.W., C.R. Hurt. 2012. Conservation genetics and evolution in an endangered species: research in Sonoran topminnows. *Evolutionary Applications* 5:806-819.

Hefner, J.M., J.D. Brown. 1984. Wetland trends in the southeastern United States. *Wetlands* 4(1):1-11.

Heimlich, R.E., K.D. Wiebe, R.L. Claassen. 1997. Sustaining our wetland gains. *National Wetlands Newsletter* 19(4):5-9.

Huennke, L.F. 1991. Ecological implications of genetic variation in plant populations. In: *Genetics and Conservation of Rare Plants*, D.A. Falk, and K.E. Holsinger (eds.), pp. 31-44. Oxford University Press, USA.

Hughes, A.R., B.D. Inouye, M.T.J. Johnson, N. Underwood, M. Vellend. 2008. Ecological consequences of genetic diversity. *Ecological Letters* 11:609-623.

Jakobsson, M., N.A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801-1806.

Jimenez, A., H. Mansour, B. Keller, E. Conti. 2014. Low genetic diversity and high levels of inbreeding in the Sinai primrose (*Primula boveana*), a species on the brink of extinction. *Plant Systematic Evolution* 300:119-1208.

Johnson, T., J.M. Cruse-Sanders, G.S. Pullman. 2012. Micropropagation and seed cryopreservation of the critically endangered species Tennessee yellow-eyed grass, *Xyris tennesseensis* Kral. *In Vitro Cellular and Development Biology—Plant* 48:369-376.

Karron, J.D. 1991. Patterns of genetic variation and breeding systems in rare plant species. In: *Genetics and Conservation of Rare Plants*, D.A. Falk and K.E. Holsinger (eds.), pp. 87-98. Oxford University Press, New York, NY. USA.

Kentula, M.E. 1996. Wetland restoration and creation. In: J.D. Fretwell, J.S. Williams, P.J. Redman (eds.) National water summary on wetlands resources. United States Geological Survey. Water-Supply Paper 2425:87-92.

Kimura, M., J.F. Crow. 1964. The number of alleles that can be maintained in a finite population. *Genetics* 49: 725-738.

Kral, R. 1978. A new species of *Xyris* (Sect. Xyris) from Tennessee and Northwestern Georgia. *Rhodora* 80: 444-447.

Kral, R. 1979. Xyridaceae. Pages 479-502 *in* R. K. Godfrey and J. M. Wooten, editors. Aquatic and Wetland Plants of the Southeastern United States, Vol. 1 (Monocotyledons). University of Georgia Press, Athens, GA, USA.

Kral, R. 1983. The Xyridaceae in the southeastern United States. *Journal of the Arnold Arboretum* 64:421–429.

Kral, R. 1990. A status report on *Xyris tennesseensis*. Unpublished report submitted to U.S. Fish and Wildlife Service, Jackson, Mississippi. 22 pp.

Kral, R. 2000. *Xyris*. Volume 22. p. 154–167. In: Flora of North America Editorial Committee (eds.). Flora of North America north of Mexico. Oxford University Press, New York, NY, USA.

Lande, R. 1988. Genetics and demography in biological conservation. *Science*. 241:1455-1460.

Leinonen, T., R.B. O'Hara, J.M. Cano, J. Merilä. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. Journal of Evolutionary Biology 21:1-17.

Liu, Z.J., J.F. Cordes. 2004. DNA marker technologies and their applications in aquaculture genetics. *Aquaculture* 238:1-37.

Meirmans, P.G. 2006. Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution* 60:2399-2402.

Meirmans, P.G., P.W. Hedrick. 2011. Assessing population structure: FST and related measures. Molecular Ecology Resources 11, 5-18.

Moffett, J.M. Jr. 2008. *Xyris tennesseensis*: status survey, habitat restoration/management concerns, and relation to a new Xyrid, *Xyris spathifolia*. Ph.D. dissertation, Auburn University, Auburn, AL.

Moffett, J.M., R.S. Boyd. 2013. Management of a population of the federally endangered *Xyris tennesseensis* (Tennessee Yellow-Eyed Grass). *Castanea* 78:198-212.

Morgan, K.L., T.H. Roberts. 2003. Characterization of wetland mitigation projects in Tennessee, USA. *Wetlands* 23(1):65-69.

Mustajarvi, K., P. Siikamaki, S. Rytkonen, A. Lammi. 2001. Consequences of plant population size and density for plant-pollinator interactions and plant performance. *Journal of Ecology* 89:80-87.

NatureServe (2013) NatureServe explorer: an online encyclopedia of life (web application), version 7.1, NatureAerve, Arlington. http://explorer.nartureserve.org

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

Nybom, H. 2004. Comparison of different nuclear DNA markers for estimating intraspecific genetic diversity in plants. *Molecular Ecology* 13:1143-1155.

Oleas, N., A.W. Meerow, J. Fransico-Ortega. 2012. Population dynamics of the endangered plant, *Phaedranassa tunguraguae*, from the tropical Andean hotspot. *Journal of Heredity* 103:557-569.

Opheim, T. 1997. Wetland losses continue but have slowed. *National Wetlands Newsletter* 19(6):7-9.

Peakall, R., D. Ebert, L.J. Scott, P.F. Meagher. C.A. Offord. 2003. Comparative genetic study confirms exceptionally low genetic variation in the ancient and endangered relictual conifer, *Wollemia nobilis* (Araucariaceae). *Molecular Ecology* 12:2331-2343.

Peakall, R., P.E. Smouse. 2006. GENALEX 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics* 28:2537-2539.

Pettengill, J.B., M.C. Neel. 2011. A sequential approach using genetic and morphological analyses to test species status: the case of the federally endangered *Agalinis acuta* (Orobanchaceae). *American Journal of Botany* 98:859-871.

Philippart, J.C. 1995. Is captive breeding an effective solution for the preservation of endemic species? *Biological Conservation* 72:281-295.

Pritchard, J.K, M. Stephens, P. Donnelly. 2000. Inference of population structure using multilocus genotype data. *Genetics* 155:945-959.

Ramasamy, R.K., S. Ramasamy, B.B. Bindroo, V.G. Naik. 2014. STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *Springerplus* 3:431.

Reed, D.H., R. Frankham. 2001. How closely correlated are molecular and quantitative measures of genetic variation? A metaanalysis. Evolution 55:1095–1103

Reed, D.H., R. Frankham. 2003. Correlation between fitness and genetic diversity. *Conservation Biology* 17:230-237.

Robert, A. 2009. Captive breeding genetics and reintroduction success. *Biological Conservation* 142:2915-2922.

Rousset, F. 2008. Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. Mol. Ecol. Resources 8: 103-106.

Schuelke M. 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature Biotechnol*ogy. 18:233–234.

Simberloff, D. 1988. The contribution of population and community biology to conservation science. *Annual Review of Ecology and Systematics* 19:473-511.

Sloop, C.M., C. Pickens, S.P. Gordon. 2011. Conservation genetics of Butte County meadowfoam (*Limnanthes floccose* ssp. *Californica* Arroyo), an endangered vernal pool endemic. *Conservation Genetics* 12:311-323.

Sloop, C.M., R. Eberl, D.R. Ayres. 2012. Genetic diversity and structure in the annual vernal pool endemic *Limnanthis vinculans* Ornduff (Limnanthaceae): implications of breeding system and restoration practices. *Conservation Genetics* 13:1365-1379.

Segarra-Moragues, J.G., J.M. Iriondo, P. Catalan. 2005. Genetic fingerprinting of germplasm accessions as an aid for species conservation: a case study with *Borderea chouardii* (Dioscoreaceae), one of the most critically endangered Iberian plants. *Annals of Botany* 96:1283-1292.

Soule, M.E. 1980. Thresholds for survival: maintaining fitness and evolutionary potential. In: *Conservation Biology: An Evolutionary-Ecological Perspective*, ed. M.E. Soule, B.A. Wilcox (eds.), pp. 151-169. Sinauer Associates Publishing, Sunderland, MA, USA.

U.S. Fish and Wildlife Service. 1994. Recovery Plan for Tennessee yellow-eyed grass *(Xyris tennesseensis* Kral). Jackson, MS. 24 pp.

U.S. Fish and Wildlife Service. 1991. Endangered and threatened wildlife plants; endangered status for the plant *Xyris tennesseensis* (Tennessee yellow-eyed grass). Fed Reg. 56:34151-34154.

Wall, M.A., A.P. Teem, R.S. Boyd. 2002. Floral manipulation by *Lasioglossum zephyrum* (Hymenoptera: Halictidae) ensures first access to floral rewards by initiating premature anthesis of *Xyris tennesseensis* (Xyridaceae) flowers. *Florida Entomologist* 85:290-291.

Wall, W.A., N.A. Douglas, W.A. Hoffman, T.R. Wentworth, J.B. Gray, Q.J. Xiang, B.K. Knaus, M.G. Hohmann. 2014. Evidence of population bottleneck in *Astragalus michauxii* (Fabaceae), a narrow endemic of the southeastern United States. *Conservation Genetics* 15:153-164.

Weir, B.S. 1990. Sampling properties of gene diversity. In: *Plant Population Genetics, Breeding, and Genetic Resources*, A.D.G. Brown, M.T. Clegg, A.L. Kahler, and B.S. Weir (eds.), pp. 23-42. Sinauer Associates Publishing, Sunderland, MA, USA.

Whigham, D.F. 1999. Ecological issues related to wetland preservation, restoration, creation, and assessment. *The Science of the Total Environment* 240: 31-40.

Whitlock, M.C. 2011. $G'_{ST}$ and $D$ do not replace $F_{ST}$. *Molecular Ecology* 20:1083-1091.

Wright, S. 1978. Variability within and among natural populations, vol 4. *Evolution and the genetics of populations*. The University of Chicago Press, Chicago, IL, USA.

Appendix 1. Polymorphic loci deviating significantly from Hardy-Weinberg expectations, by population.

| State-Population | Number of Polymorphic Loci | Deviating Loci | *P*-value |
|---|---|---|---|
| AL-E | 9 | 175, 1066, 1418, 1558, 1570, 1712, 2846, 7809 | $P<0.0017$ |
| AL-W | 9 | 153, 175, 1066, 1418, 1670, 1712 | $P<0.0020$ |
| AL-S | 1 | 153 | $P<0.0000$ |
| AL-Y | 2 | 153 | $P<0.0000$ |
| GA-H | 1 | 153 | $P<0.0000$ |
| GA-CC | 4 | 153, 175, 7809 | $P<0.0045$ |

*Appendix 2. Allele frequencies at 14 loci for Xyris tennesseensis.* Sample size per population (N) is 30, except for TN-A (N=18) and GA-P (N=21). (bp=base pairs)

| Locus | Allele Size (bp) | TN-L | TN-T | TN-G | TN-D | TN-A | AL-E | AL-W | AL-S | AL-Y | GA-H | GA-M | GA-C | GA-CC | GA-P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **33** | 263 | 0.00 | 0.00 | 0.00 | 0.07 | 0.00 | 0.88 | 1.00 | 1.00 | 0.03 | 1.00 | 1.00 | 1.00 | 1.00 | 0.98 |
| | 267 | 1.00 | 1.00 | 1.00 | 0.93 | 1.00 | 0.00 | 0.00 | 0.00 | 0.97 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 |
| | 271 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.12 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **57** | 346 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 352 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **153** | 350 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.27 | 0.43 | 0.67 | 0.27 | 1.00 | 1.00 | 0.50 | 1.00 |
| | 354 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.73 | 0.57 | 0.33 | 0.73 | 0.00 | 0.00 | 0.50 | 0.00 |
| **175** | 200 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.75 | 0.65 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 202 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.25 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 204 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.92 | 0.00 |
| | 206 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.35 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 1.00 |
| | 208 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.08 | 0.00 |
| **1066** | 385 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.67 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 387 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.70 | 0.33 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 389 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 391 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.30 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **1245** | 242 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 | 0.00 |
| | 244 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 1.00 |
| **1418** | 322 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 324 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 0.72 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 328 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.18 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 330 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.82 | 0.28 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

| Locus | Allele Size (bp) | TN-L | TN-T | TN-G | TN-D | TN-A | AL-E | AL-W | AL-S | AL-Y | GA-H | GA-M | GA-C | GA-CC | GA-P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1558** | 215 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.50 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 217 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.50 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | 219 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| **1570** | 221 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.20 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 223 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.80 | 0.72 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 225 | 0.02 | 0.03 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 233 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 235 | 0.98 | 0.97 | 1.00 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 241 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.28 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 0.00 |
| | 243 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| **1670** | 344 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.28 | 0.00 | 1.00 | 0.00 | 0.97 | 0.00 | 0.15 | 1.00 |
| | 347 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.72 | 1.00 | 0.00 | 1.00 | 0.03 | 1.00 | 0.85 | 0.00 |
| **1712** | 230 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.43 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 232 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.77 | 0.57 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 1.00 | 0.00 |
| | 234 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.23 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 236 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 1.00 |
| **2032** | 184 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 1.00 |
| | 186 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 1.00 | 1.00 | 0.00 |
| **2846** | 205 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.10 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 222 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 |
| | 224 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.55 | 0.70 | 1.00 | 0.00 | 1.00 | 0.00 | 1.00 | 1.00 | 1.00 |
| | 226 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 230 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.30 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| | 234 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.35 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |

| Locus | Allele Size (bp) | TN-L | TN-T | TN-G | TN-D | TN-A | AL-E | AL-W | AL-S | AL-Y | GA-H | GA-M | GA-C | GA-CC | GA-P |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **7809** | 182 | 0.00 | 0.00 | 0.05 | 0.00 | 0.00 | 0.63 | 0.57 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.58 | 0.00 |
| | 185 | 1.00 | 1.00 | 0.95 | 1.00 | 1.00 | 0.37 | 0.43 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.42 | 1.00 |