THE EVOLUTIONARY HISTORY OF PAYSONIA

(BRASSICACEAE)

By

GUADALUPE BORJA

Bachelor of Science in Botany

University of Oklahoma

Norman, Oklahoma

2008

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July, 2013

THE EVOLUTIONARY HISTORY OF PAYSONIA

(BRASSICACEAE)

Thesis Approved:

Dr. Andrew Doust

Thesis Adviser

Dr. Mark Fishbein

Dr. Linda Watson

ACKNOWLEDGEMENTS

I would like to express my deepest appreciation to all those who supported me through the duration of my master studies. First off, I want to begin by thanking my advisory committee and most especially my mentor, Dr. Andrew Doust, for his patience, support, and supervision during the course of my studies. I am also thankful for all members of the Doust lab who have at one time helped me by providing me with advice and lab training.

Additionally, my fieldwork would not have been successful had it not been for the help of Mary Gard, who accompanied me on an adventurous trip to Tennessee to find and collect *Paysonia*. I also want to acknowledge Dr. Dwayne Estes from Austin Peay State University in Tennessee. He took time out of his busy schedule to meet with me and aid me in locating various populations of *Paysonia*.

My graduate studies and particularly my research work was made possible due to the financial support of teaching assistantships offered by the Oklahoma State Botany Department, an Oklahoma State Graduate College Summer Research Fellowship, the Genetics Section of the Botanical Society of America Masters Student Award, a Botanical Society of America Graduate Student Research Award, and an American Society of Plant Taxonomists Graduate Student Research Award.

Last but not least, I would also like to extend a special thanks to my family. I would especially like to thank my father Sergio Borja for his encouragement and example, who

taught me the importance of education and introduced me to science. I want to also acknowledge my lovely mother, Rosalina Glover, for always believing in me. Last but not least, I sincerely like to thank my dearest friend and soon-to-be husband, Chris Ivener, for his unconditional love, support, and patience during the duration of my studies.

Name: GUADALUPE BORJA

Date of Degree: JULY 2013

Title of Study: THE EVOLUTIONARY HISTORY OF PAYSONA (BRASSICACEAE)

Major Field: BOTANY

ABSTRACT: *Paysonia* is a small genus of annuals consisting of eight species; three are found in Texas and Oklahoma and the remaining five are in the Central Basin of Tennessee and northern Alabama. Species are morphologically distinct, but interspecific relationships are uncertain. To investigate evolutionary relationships in a phylogenetic context, the internal transcribed spacer (ITS) region of nuclear ribosomal DNA and three chloroplast markers were sequenced for multiple accessions from different populations of each species. Although little phylogenetic resolution was found among the Tennessee/Alabama species, each of the Texas/Oklahoma species is monophyletic. These phylogenetic analyses suggest that both incomplete lineage sorting and gene flow may be complicating the recovery of evolutionary relationships among the southeastern species. To test for evidence of present day hybridization in Tennessee, multiple microsatellite markers were used to document gene flow among populations of each species. Results indicate a complex combination of relationships and present day gene flow in these species; the data in combination suggest that each species is a separately evolving metapopulation lineage.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

CHAPTER I

INTRODUCTION

Paysonia O'Kane & Al-Shehbaz (Brassicaceae) is a genus of herbaceous, annual plants ranging across four states of the southern United States and parts of northern Mexico (Fig 1.1). Eight diploid species comprise the group: *P. auriculata* (Engelm. & A. Gray) O'Kane and Al-Shehbaz, *P. densipila* (Rollins) O'Kane and Al-Shehbaz, *P. grandiflora* (Hook.) O'Kane and Al-Shehbaz, *P. lasiocarpa* (Hook. ex A. Gray) O'Kane and Al-Shehbaz, *P. lescurii* (A. Gray) O'Kane and Al-Shehbaz, *P. lyrata* (Rollins) O'Kane and Al-Shehbaz 2002, *P. perforata* (Rollins) O'Kane and Al-Shehbaz, and *P. stonensis* (Rollins) O'Kane and Al-Shehbaz. Historically, the eight species of *Paysonia* were classified under the much larger genus, *Lesquerella* S. Watson. Reed C. Rollins (1955;1973), however, recognized that these "auriculate-leaved *Lesquerellas*" were morphologically distinct from the rest of the genus and Rollins and Shaw (1973) classified them all in an informal group within *Lesquerella*. Even so, no official taxonomic recognition of this group was proposed until 2002 when the species were formally transferred to a new genus, *Paysonia* (O'Kane and Al-Shehbaz 2002), based on ITS sequence data. The genus was named to honor Edwin B. Payson, the first to monograph *Lesquerella* (Payson 1922). *Physaria* (Nuttal ex Torrey & A. Gray), a related genus, was expanded to incorporate the remaining ninety-one species from *Lesquerella* (Al-Shehbaz and O'Kane 2002). The main characteristics that set *Paysonia* apart from *Physaria* are sessile auriculate cauline leaves and seeds containing wing-like lateral outgrowths.

1.1. Taxonomic Classification

Paysonia belongs to the Brassicaceae, a large, primarily herbaceous family comprised of 3710 species and 338 genera (Warwick et al. 2010). It is the largest family in the order Brassicales, an order that is well-known for its mustard-oil glucosides and myrosin cells (Al-Shehbaz 2011). At the generic level, this family has been taxonomically difficult to classify. Various molecular studies (Al-Shehbaz et al. 2006; Bailey et al. 2006; Beilstein et al. 2008; Couvreur et al. 2010; Warwick et al. 2010) have shown that the previously-held classification of the family based on morphological characters alone was largely artificial due to the convergent evolution of many of the characters used to define the genera, such as fruit morphology and seed embryo type (Al-Shehbaz et al. 2006). Currently, about 93% of the family is arranged within 48 tribes (Al-Shehbaz 2011). There is still some confusion on the classification of some of the basal genera, but there is consensus on the recognition of three major lineages within the family.

Currently, *Paysonia* is recognized as being one of seven genera within the tribe Physarieae B.L. Rob, along with *Dithyrea* Harv., *Dimorphocarpa* Rollins, *Nerisyrenia* Greene, *Lyrocarpa* Hook. & Harv., *Synthlipsis* A. Gray, and *Physaria* (Nutt. ex Torr. & A. Gray) A. Gray (Fuentes-Soriano and Al-Shehbaz 2013). Physarieae is classified in one of the three major lineages (Lineage I) that also include the scientifically important species, *Arabidopsis thaliana* (L.) Heynh. in the tribe Camelineae. Physarieae is the only tribe in the family in which all of its species have multi-colpate pollen (Fuentes-Soriano and Al-Shehbaz 2013). The most current molecular data, based on chloroplast markers, supports the sister relationship of *Paysonia* to *Physaria* (Fuentes-Soriano and Al-Shehbaz 2013). Together they form a clade which is sister to a lineage containing the remaining 5 genera in the Physarieae (Fuentes-Soriano and Al-Shehbaz 2013). These latest results, however, conflict with the ITS data of Bailey et al. (2006), which resolved *Paysonia* as sister to the entire Physarieae.

Although the position of *Paysonia* within the tribe remains unclear, its monophyly of *Paysonia* is uncontested. The recognition of *Paysonia* as a monophyletic lineage is well-supported by both morphology and DNA sequence data (O'Kane and Al-Shehbaz 2002; Al-Shehbaz et al. 2006; Bailey et al. 2006; Beilstein et al. 2008; Fuentes-Soriano and Al-Shehbaz 2013). The most extensive studies that have included all eight species in *Paysonia* examined sequence data of the chloroplast gene *ndh*F (Fuentes-Soriano and Al-Shehbaz 2013) and nuclear regions LUMINIDEPENDENS and ITS (Fuentes-Soriano unpublished data 2010). While all 8 species of *Paysonia* were found to be monophyletic, relationships within the genus remain incompletely resolved, in particular for the clade containing *P. densipila*, *P. lescurii*, *P. lyrata*, *P. perforata*, and *P. stonensis* (Fig 1.3).

1.2. Geographic Distribution

Paysonia species are distributed across four states of the southern United States and part of northern Mexico (Fig 1.1). The genus has an east-west disjunction with three species in the "southwestern" group (*P. lasiocarpa*, *P. grandiflora*, and *P. auriculata*) and five species in the "southeastern" group (*P. lyrata*, *P. densipila*, *P. lescurii*, *P. stonensis*, and *P. perforata*).

In the southwestern group, there are two species with a primary distribution in Texas. *Paysonia lasiocarpa* occurs predominately along the eastern coastal region in Texas, growing mainly on sandy or gravelly soils. The range of this species also extends to the mountains and foothills of northeastern Mexico (Rollins 1955). Within *P. lasiocarpa* there are currently three recognized subspecies: *P. lasiocarpa* ssp. *lasiocarpa* (Hook. ex A. Gray) O'Kane and Al-Shehbaz, *P. lasiocarpa* ssp. *berlandieri* (A. Gray) O'Kane and Al-Shehbaz, and *P. lasiocarpa* ssp. *heterochroma* (S. Watson) O'Kane and Al-Shehbaz 2002. *P. lasiocarpa* ssp. *lasiocarpa* is found in Texas, and populations have been recorded in Coahuila, Nuevo Leon, and Tamaulipas, Mexico. *Paysonia lasiocarpa* ssp. *berlandieri* grows in Texas and records list it as occurring in

Nuevo Leon, Tamaulipas, and Veracruz, Mexico. *Paysonia lasiocarpa ssp. heterochroma* is only known from high elevations and moist habitats in Nuevo Leon, Mexico (Rollins 1955). The second predominantly Texas species, *P. grandiflora,* grows almost exclusively on loose and well-drained sandy soils and can be found in south-central Texas, east of the Edwards Plateau (Rollins 1955).

The third southwestern species is *P. auriculata*. It is mainly found in Oklahoma, although there have been populations that have been recorded in northern Texas. It is also known in southern Kansas growing in the Chikaskia watershed; these Kansas populations are thought to be critically imperiled (NatureServe 2013). It is possible that *P. auriculata* is not as continuously distributed as it once was, maybe likely due to anthropogenic disturbances. The current conservation status in Oklahoma has not yet been assessed. This species has been observed growing on pastures and roadsides, mainly on sandy loam type soils in regions dominated by Permian red shales and sandstones. Populations have also been found on some limestone outcrops in Oklahoma (Rollins and Shaw 1973).

The southeastern group is composed of *P. densipila*, *P. lescurii*, *P. lyrata*, *P. perforata,* and *P. stonensis*. All but *P. lyrata* are found in the Central Basin of Tennessee, around Nashville. *Paysonia lyrata* is known from only a few cedar glade localities in northern Alabama, with one extant population each in Colbert, Lawrence and Franklin Counties (Service 1996). The Colbert County population is found in the Spring Creek watershed and the Lawrence and Franklin County populations are located in the Town Creek watershed, which empties upstream of the mouth of Spring Creek into the Tennessee River. The range of *P. lyrata* is highly restricted, and thus is considered an endangered species (U.S. Fish and Wildlife Service 1990). There is a recovery plan in place that aims to protect the few extant populations and to increase the number of populations to nine (Service 1996). Some disturbance is required to maintain the species, probably to prevent ruderal perennial species from outcompeting them (Rollins and Shaw 1973).

It is believed that one of the main reasons the species is less common than in pre-settlement times is due to the lack of natural disturbances in glades.

Paysonia lescurii is predominantly found growing along the lower Cumberland River in various counties located in the northern part of the Central Basin. Populations have been found in southern Kentucky, along the Cumberland River in Trigg County (NatureServe 2013). This species has been recorded growing on a variety of sites including hillsides, cedar glades, flood plains, fields, and pastures (O'Kane, S.L., Jr. 2010).

The northern limit of *P. densipila* is the West Fork of the Stones River, and the species is found southwestward into Giles County in Tennessee. Some historical records list it in Morgan and Lawrence Counties in Alabama. It grows predominantly on cedar glades, open alluvial sites, stream bottoms, and fallow fields.

Paysonia stonensis has a restricted distribution along the East Fork of the Stones River in Rutherford County. It is seen growing in pastures, fields, roadsides, and stream banks. This species is federally-listed as threatened (U.S. Fish and Wildlife Service 1999), and is thus of conservation concern.

Paysonia perforata also has a restricted distribution. This species is found along three creeks (Spring, Bartons, and Cedar) outside the town of Lebanon, Tennessee in Wilson County. It is known from only 21 sites in a five-mile radius (U.S. Fish and Wildlife Service 2006). This species is federally-listed as endangered because of its restricted distribution (U.S. Fish and Wildlife Service 1996). Although it grows abundantly where found in open fields, pastures, flood plains, and road sides, the conversion of land to uses other than cultivation, and rapid commercial, residential, and industrial development in the county is drastically reducing the availability of suitable habitat (U.S. Fish and Wildlife Service 2006; Fitch et al. 2007). There is a recovery plan in place until 2025 that aims to have a minimum of 25 occurrences of *P. perforata*,

with at least five along each of the three creeks (U.S. Fish and Wildlife Service 2006). These sites must have an average of 500 plants over a ten-year period to be considered an occurrence. Despite the restricted habitat of the species, the genetic variation for the extant populations is similar to that of more widespread species (Baskauf 2002). The work described in this thesis will contribute to the monitoring of the different sites.

1.3. Morphology

Paysonia species are differentiated from the remainder of the Physarieae on the basis of auriculate-cauline leaves and flattened seeds with distinctive lateral outgrowths. The morphological differences between the species are presented in Table 1.1. As with most Brassicaceae, the major differences are in fruit form, trichome type, and trichome distribution (Fig 1.4). Floral morphology among species does not differ drastically; all species have yellow obovate petals with the exception of *P. stonensis* and *P. perforata* with white petals.

The character of auriculate-cauline leaves is constant for all species expect *P. lasiocarpa*, where plants in some populations do not exhibit clear auricles on their cauline leaves. Because of this, Rollins (1955) was unsure of where to place *P. lasiocarpa* within *Lesquerella*. He thought it was closely related, and possibly a link to the species of the ditypic genus, *Synthlipsis* A. Gray because they hold some features in common, such as a medially flattened, pubescent silique (Rollins 1955; Rollins and Shaw 1973). However, he recognized that within *Lesquerella, P. lasiocarpa* most closely resembles *P. grandiflora* (Rollins and Shaw 1973). This influenced his placement of *P. lasiocarpa* with the group of auriculate-leaved species (Rollins and Shaw 1973). Within *P. lasiocarpa*, there are currently three recognized taxa, but Rollins (1955) and O'Kane and Al-Shehbaz (2002) did not assign them to specific rank. The morphology of *P. lasiocarpa* ssp. *heterochroma* differs from the other subspecies in that it exhibits a perennial caespitose habit with a comparatively thick caudex, a short thickened perennial stem at the base of the plant. The

collections of this variety come from much moister and high elevation (above 1000 meters) habitats, which may allow for a perennial habit to be sustained (Rollins 1955). The other subspecies are found in lower elevations and mostly along the Gulf coastal plain (Rollins 1955; Rollins and Shaw 1973). The main difference between the other subspecies is that the fruits of subspecies *lasiocarpa* have only slightly flattened fruits, closer to being spherical in shape, while subspecies *berlandieri* exhibits strongly compressed fruits (Rollins 1955; Rollins and Shaw 1973; O'Kane and Al-Shehbaz 2002).

Ovule number varies in *Paysonia.* The highest number of ovules is observed in *P. lasiocarpa* (14-32 per ovary), followed next by *P. grandiflora* (16-28 per ovary). A slight reduction of ovules is seen in *P. auriculata* with 12-20 ovules per ovary. In *P. densipila*, *P. lescurii*, and *P. lyrata*, the range is 4-8, and 4-12 and 8-12 in *P. perforata* and *P. stonensis*, respectively.

The most common chromosome number in *Paysonia* is n=8. However, the chromosome count of *P. lasiocarpa* is n=7 and *P. grandiflora* is n=9. The southeastern species, *P. densipila*, *P. lescurii*, *P. lyrata*, *P. perforata*, and *P. stonensis*, all share a chromosome number of n=8.

Apart from differences in ovule and chromosome number, there are few clear morphological differences between *P. grandiflora* and *P. auriculata.* One difference is that the infructescence of *P. auriculata* is dense and short, while in *P. grandiflora* it elongates and becomes somewhat loose. In addition, the vegetative trichomes on *P. grandiflora* are stellate, while those on *P. lasiocarpa* are long simple, mixed with short, branched trichomes.

Rollins (1973) hypothesized that *P. lyrata* was the evolutionary link between *P. densipila* in the Central Basin, and *P. auriculata* in Oklahoma. This is because *P. lyrata* shares morphological similarities with both *P. auriculata* and *P. densipila*. All three contain inflated fruits with some small differences. Like *P. auriculata*, the exterior of the siliques of *P. lyrata* is glabrous. The greatest difference between *P. auriculata* and *P. lyrata*, however, is the reduction of ovules in the

ovary. The ovary of *P. auriculata* contains about 12-20 ovules, whereas *P. lyrata* the number of ovules ranges from 4-8. The siliques of *P. auriculata* are a bit longer than broad and elliptical in outline, as well as about 2 mm longer than those of *P. lyrata*. The siliques in *P. lyrata* are broader than long, with a slight depression at the base of the style. However, the general shape of the siliques of *P. lyrata* and *P. densiplila* are very similar, although *P. densipila* has a pilose indumentum on the silique exterior.

Paysonia densipila also shares many similarities with *P. stonensis.* Both *P. densipila* and *P. stonensis* have round, subglobose, and subsessile fruits. Both species have siliques with trichomes, although in *P. densipila* they can also be simple or branched but are only simple in *P. stonensis*. The interior of the valves in both species is glabrous. *P. stonensis*, however, shares many other similarities with *P. perforata* that are not shared with the other southeastern species, the main one being white flowers. While the rest of the taxa in the genus contain complete fruit septa, the septa of *P. stonensis* can range from complete to perforate. In *P. perforata*, the septum ranges from perforate to nearly absent. Rollins (1955) believed that it was possible that *P. stonensis* gave rise to *P. perforata*, and morphologically, this appears plausible.

Paysonia lescurii has distinctive flattened siliques and large bulbous-based trichomes on the valve exteriors. *Paysonia lescurii* shares a rare characteristic in *Paysonia* of branched trichomes on the valve interiors with *P. perforata*. The compressed fruits of *P. lescurii*, however, are distinct from the rest of the southeastern species. Historically, because the fruits were so unlike any of the other *Paysonia* species in the Central Basin, there were challenges to assigning the taxonomic position of this species. This species was once even assigned to the genus *Alyssum* by Asa Gray (1867), but was later moved by Watson (1888), who placed it with *P. auriculata* and *P. grandiflora* (at that time, the other Central Basin species were unknown). Payson (1922), in his monograph of the *Lesquerella*, thought *P. lescurii* should form a monotypic section. Rollins,

however, showed that this species was genetically closer to *P. densipila* when he discovered that they freely interbreed in the field (1955).

A comprehensive seed coat study has yet to be conducted in *Paysonia*. Preliminary data (Fuentes-Soriano unpublished data) suggests that this may prove to be another source of morphological variation between taxa, and provide insight into species relationships.

1.4. Ecology

With the exception of *P. lasiocarpa* subsp. *heterochroma,* all *Paysonia* species are winter annuals (Rollins 1955) and are self-incompatible and thus obligate outcrossers. *Paysonia lasiocarpa* subsp. *heterochroma* is a perennial. The introduced honeybee is one of the most important pollinators for *Paysonia*, but solitary bees and a variety of dipteran flies also visit the flowers (Rollins and Solbrig 1973). The southeastern species are known to form persistent seed banks, where seeds are thought to be viable for about 6 years (Baskin and Baskin 1990; Baskin 1992; Baskin 2000; Fitch et al. 2007). It is likely that before colonial settlement they grew on floodplains where flooding prevented the formation of a closed canopy of trees or on rocky exposures of limestone in cedar glades (Fitch et al. 2007). The strict winter annual growth and the presence of persistent seed banks likely help the species survive in the flood plains (Baskin and Baskin 1990). With the advent of settlement and agriculture in their southeastern range, they spread into anthropogenically disturbed habitats where certain agricultural practices, such as tilling and plowing, formed a suitable habitat for *Paysonia*. The species are considered to have weedy tendencies and may be found growing profusely in fields by the thousands under suitable conditions (Rollins 1955; Baskin and Baskin 1990; Baskin 1992; Baskin 2000; Fitch et al. 2007).

High temperatures are required for the seeds to break dormancy (Baskin 1992), and the seeds must be photostimulated in the late summer for a period of weeks for successful germination (Fitch et al. 2007). Plowing must be done before photostimulation, but after plants have

dispersed their seeds so as not to deplete the soil bank. Plowing while seeds are still dormant prevents other ruderal or woody plant species from invading, which would compete with *Paysonia* (Baskin and Baskin 1990). Seed germination usually takes place in September, and the plants over-winter and flower in the early spring.

1.5. Species Diversity and Biogeography

The southeastern United States harbors the majority of the species diversity of the genus as currently circumscribed, concentrated mainly in the Central Basin of Tennessee. In general, the Central Basin is regarded as a center of high endemism for plant species (Estill and Cruzan 2001). The Central Basin is known for its limestone cedar glades, which are considered most abundant and best developed in the Cumberland and Duck River drainage basins (Baskin et al. 2007). These cedar glades are characterized by high irradiance, high soil temperatures in the summer, and extremes in soil moisture content that range from inundation in winter to below the permanent wilting point in summer (Quarterman 1950; Baskin et al. 2007). Specialization to this extreme microclimate may explain the high number of endemic species present in the region (Estill and Cruzan 2001).

Geology, paleogeography, and climate are considered to be major factors that have impacted the glades of the Central Basin (Fig. 1.2). Prior to the Paleozoic, Middle Tennessee was a low-lying plain at a similar level to the current Highland Rim that surrounds the Central Basin (Delcourt et al. 1986). Throughout the Paleozoic Era, multiple tectonic collisions occurred that caused a major bulge in middle Tennessee, named the Nashville Dome. Erosion of the Nashville Dome since the end of the Paleozoic Era until now has caused the present day bowl-like shape of the basin. Cedar glades are found on limestone rock exposures from the Stones River Group Formation from the Mid-Ordovician, in the innermost part of the Central Basin (Wilson 1962). With the constant erosion the region experiences, it is expected that cedar glades will disappear

from the inner part of the Basin as the Stones River Group Formation is eroded away, but will most likely appear around the surrounding portions of the basin when erosion exposes that particular rock group along the slopes of the Basin (Roland 1926).

Speciation in middle Tennessee may have been influenced greatly by the fragmentation of the landscape by major drainage systems. The dissection of the landscape by the Cumberland, Duck, and Tennessee River Drainage Systems is believed to have been a dynamic process over the last two to ten million years (Starnes and Etnier 1986). Each of the middle Tennessee *Paysonia* species is associated with a particular drainage system; *P. densipila* with the Duck River, *P. lescurii* with the Cumberland River, *P. perforata* with Spring, Barton and Cedar Creeks, and *P. stonensis* with the East Fork of the Stones River. Studies in other organisms are able to support the hypothesis that the paleogeography of the drainage systems played a large role in the speciation of major groups in that region. For example, fish groups such as the *E. simoterum* complex (snubnose darters; Percidae: *Etheostoma*) contain species that differentiated along the three drainage systems. Using appropriate fossils for calibration points, Harrington and Near (2012) estimated divergence times between three species occurring respectively along the Tennessee, Cumberland, and Duck drainage systems. The analysis showed the Duck River species as sister to a clade comprised of the Tennessee and Cumberland River species. The age of the most recent common ancestor of the whole *E. simtoreum* complex was estimated to be 3.7 million years, while the most recent common ancestor of the Tennessee and Cumberland River clade was estimated as 2.9 million years (Harrington and Near 2012). This suggests that other species groups with similar geographic distribution along river systems, including *Paysonia*, may exhibit similar divergence times and that speciation along the river systems is a fairly recent occurrence.

In addition, climatic changes during the Last Glacial Maximum (LGM), approximately between 24,000-12,000 years ago, may have also influenced the distribution and speciation of Central

Basin taxa. Although an ice sheet did not cover the region, it experienced a cooler and wetter climate, which is considered unsuitable for cedar glade flora (Delcourt et al. 1986). Suitable cedar glade habitat could only be found south of 34°N latitude, and thus northern Alabama is believed to have served as a refugium for cedar glade endemics from middle Tennessee (Delcourt et al. 1986). Another possible effect of climate change is that populations of a more widespread ancestral species became fragmented within the region into restricted habitats with suitable microclimates that allowed them to survive the LGM, followed by speciation.

1.6. Species Concepts

The lack of phylogenetic resolution in the southeastern species of *Paysonia* (*P. densipila, P. lescurii, P. lyrata, P. perforata,* and *P. stonensis*) belies their morphological differences (Table 1.1). Their demonstrated capacity for interbreeding (Rollins 1957; Rollins and Solbrig 1973; Rollins 1988) casts doubt on whether they should be considered distinct species. Defining a species is perhaps one of the most passionate topics of debate among biologists. Species are one of the most fundamental units of biology, and there is currently no unified agreement on what constitutes a species. Presently, there exists a multitude of species concepts; Mayden (1997) listed at least 22. Perhaps the most utilized species concepts are the biological, phylogenetic, and morphological species concepts (Freeman and Herron 2001). While the biological species concept is one of the most commonly used, it is still a limiting concept, especially for more recent species (Mayr 1942). Under the biological species concept (Mayr 1942), all of the southwestern *Paysonia* species have acquired reproduction barriers. In the case of southeastern *Paysonia,* these species still have the ability to interbreed. Dobzhansky (1951) argued that genetic incompatibility was necessary in order to define a species. However, closely related allopatric species, such as in southeastern *Paysonia*, do not always form reproductive barriers, due to their spatial separation and lack of interaction.

De Queiroz (2007), however, argued that what determines a species is that they must be separately evolving metapopulation lineages. De Queiroz's idea of a 'unified species concept' explains that species do not need to fall under one specific species concept, as species are continually evolving, and thus the acquisition of the properties that are used to define species happen at different stages and in a random order. What we currently know about *Paysonia* is that only the southwestern species are reproductively isolated while the southeastern species still have the ability to interbreed. One species concept that applies to all eight species is the morphological species concept, currently used to describe the genus on the basis of morphological characters; under this concept, the southeastern *Paysonia* are considered separate entities, as they each have defining morphological characters that set them apart. The only intermediates are found in hybrid populations. Genetically, it is unknown how differentiated each species is, although previous molecular phylogenetic analyses (Fuentes-Soriano and Al-Shehbaz 2013) suggest that reciprocal monophyly has been achieved for the southwestern *Paysonia.* However, this still remains to be investigated.

1.7. Evolutionary processes affecting speciation

Rollins (1952) believed that *Paysonia* had a southwestern origin, with *P. auriculata* in Oklahoma and *P. lyrata* in Alabama as evolutionary links to the Tennessee species. Long distance dispersal between the southwest and southeast could explain the disjunction in *Paysonia*. Alternatively, it is possible that there was a widespread ancestral species whose range was fragmented, since it is believed that at one time the Nashville Dome was once connected to the Ozark Dome (Starnes and Etnier 1986). Some of the more recent geological and climatic changes mentioned in the previous section may have caused *Paysonia* to experience fragmentation that lead to speciation. This makes it likely that speciation of the southeastern *Paysonia* was a fairly recent occurrence. The recently published Physariae phylogeny of Sara-Fuentes and Al-Shehbaz (2013) shows no

resolution of the southeastern clade (Fig. 1.3) and provides further support for the hypothesis that speciation of the southeastern *Paysonia* was a recent event.

Paysonia has a similar life cycle and distribution pattern as another Brassicaceae genus, *Leavenwothia* Torr. *Leavenworthia*, like *Paysonia,* has a large number of rare endemic taxa in the Central Basin (Estill and Cruzan 2001) and a similar east-west disjunction (Rollins 1955; Beck et al. 2006). The most current molecular phylogeny of *Leavenworthia* displays a weakly supported clade of three Central Basin species (*L. exigua*, *L. stylosa*, and *L. torulosa*) where species relationships are unclear and each species is not reciprocally monophyletic (Beck et al. 2006). Such patterns can have a variety of causes, including incomplete lineage sorting, gene flow, and recombination (Maddison 1997; Degnan and Rosenberg 2009), as discussed further below.

Incomplete lineage sorting (ILS), the stochastic extinction of ancestral polymorphisms, is a common phenomenon in recently diverged lineages that is known to pose a challenge to inferring species-level phylogenies (Maddison and Knowles 2006). This is especially so when there are large population sizes and speciation events happen in rapid succession, as ancestral polymorphisms can be retained in each of the new lineages (Maddison 1997). ILS is also known as "deep coalescence", a term first coined by Maddison (1997) when the common ancestor of alleles sampled from different species are found to coalesce deeper in the phylogeny than the divergence of the species from which they were sampled (Felsenstein 2004). The random sorting of alleles through time is a coalescent process, and Kingman (1982) was the first to mathematically describe the coalescent. Hudson (1990) also made a significant contribution to this theory by outlining an algorithm that can simulate allelic data under different population models. Coalescent theory incorporates major assumptions such as non-overlapping generations, constant effective population size (N_e) within populations, and random mating, all of which are derived from the Wright-Fisher model (Wright 1931; Degnan and Rosenberg 2009).

Persistent seed banks, large effective population sizes (N_e), and a small number of generations between species divergence events are all factors that are known to increase the chances of ILS being present (Hudson 1990; Maddison 1997). All of these factors are observed in the southeastern *Paysonia*. There are many studies of recently diverged groups are that are affected by ILS, including *Liolaemus* lizards (Camargo et al. 2012), the Mediterranean *Linaria* plants (Blanco-Pastor et al. 2012), *Drosophila* (Pollard et al. 2006), *Melanoplus* grasshoppers (Carstens and Knowles 2007), and New Zealand alpine *Maoricicada* cidadas (Buckley et al. 2006). This phenomenon may cause gene trees to be discordant with the species tree, and it becomes a challenge to determine which genes correctly reflect the species history.

To understand how incongruence due to ILS between gene and species topologies can occur, it is important to recall that within populations, not all individuals are genetically identical as a given gene may have slightly different forms, or alleles. To illustrate this point, consider the species and gene tree outlined in Figure 1.5. The species tree (Fig. 1.5.A) shows a different history than the gene tree (Fig. 1.5.B), so, when both are compared, they are discordant. The short branches on the species tree indicate that not many generations have passed since speciation and the branch thickness indicates that there is a large effective population size; under both conditions ILS is more likely to be observed. Effective population size is a measure of the idealized population that would undergo the same magnitude of genetic drift as the population of interest (Conner and Hartl 2004). Computer simulations of mitochondrial lineages have shown that shortly after speciation, the probability is high for the alleles of sister taxa to be polyphyletic or paraphyletic (Neigel and Avise 1986). This probability decreases with further generations, so that by approximately $4N_e$ generations after speciation (where N_e represents the effective population size), the probability of the sister taxa to be reciprocally monophyletic on the gene tree is significantly increased (Neigel and Avise 1986). For nuclear genes, the time for sorting to occur

would be extended because the effective population sizes for nuclear loci are four-fold times larger since the genes are diploid and are biparentally inherited (Avise 1994).

Since genealogical histories do not always match the species' history, it is important to properly sample and analyze gene sequences to infer the evolutionary relationships among species. Several methods have been used to try and resolve phylogenies affected by ILS. Maddison and Knowles (2006) found that despite ILS, gene sequences retain enough phylogenetic signal needed to reconstruct an accurate phylogeny when a reasonable number of individuals per taxa and multiple genes are sampled. The number sampled is dependent on the depth of the species tree; e.g. a shallow species tree benefits with an increase of individuals sampled, whereas a deeper species tree can be accurately reconstructed utilizing a larger number of loci.

There have been a number of methods used to analyze multigene datasets. Concatenation approaches have been extensively used when dealing with multigene data sets (Rokas et al. 2003; Gadagkar et al. 2005; de Queiroz and Gatesy 2007). This method is able to generate well-supported and resolved phylogenetic trees utilizing a supermatrix formed with a large number of concatenated gene regions. Methods, such as maximum likelihood, can be used to analyze the "supergene" and the resulting tree is then assumed to reflect the true species phylogeny. However, this approach can be misleading because ILS can cause conflict among individual gene trees (Edwards et al. 2007; Kubatko and Degnan 2007; Degnan and Rosenberg 2009). Kubatko and Degnan (2007) showed how the concatenation of multiple genes with incongruent topologies could lead to an incorrect species tree, albeit with strong bootstrap support. This study demonstrated the importance of accounting for gene history heterogeneity when dealing with multilocus sequence data to accurately infer a species phylogeny.

An alternative to concatenation when dealing with multilocus sequence data is to incorporate a coalescence approach, where incomplete lineage sorting in each locus is modeled when deriving

species phylogenies. Coalescent methods trace the coalescence of gene copies back in time from the present day sample taxa, and are followed by approaches to minimize discordance between different genealogies within the constraints of a species tree (Rosenberg and Nordborg 2002; Degnan and Rosenberg 2009; Liu et al. 2009). This backward in time approach only takes into consideration current gene copies present in populations at the time of sampling. A forward approach would follow a sampling of alleles within populations through time, where many lineages may prove to be an evolutionary dead end due to extinction, although it's not possible to model this. There are now several phylogenetic programs that implement the coalescent for multilocus data, such as BEST (Liu 2008), *BEAST (Heled and Drummond 2010), and STEM (Kubatko et al. 2009).

Solely using coalescent methodology to estimate species phylogenies does not mean an accurate species tree will be inferred. The coalescent model only takes ILS into account as a source of discordance, but ILS is not the only cause of species and gene tree conflict. Other biological processes such as recombination and hybridization are also known to cause conflict in the data used to infer phylogenies (Maddison 1997; Degnan and Rosenberg 2009). This is significant for this study because the southeastern *Paysonia* species freely hybridize (Rollins 1988). Hybrid populations are known to occur between *P. densipila x P. lescurii*, *P. densipila x P. stonensis*, and *P. stonensis x P. lescurii* (Rollins 1955; Rollins and Shaw 1973). It is unclear what role hybridization and introgression have on speciation processes, but the presence of hybrid populations suggests that the Central Basin *Paysonia* species could also be undergoing gene flow effects, in addition to ILS. Hybridization causes similar effects as ILS on phylogeny reconstruction, making it difficult to differentiate between the two. Take, for example, the genealogy outlined in Figure 1.6. When two species hybridize, genes will be transferred between the two when there is backcrossing and introgression. The evolutionary histories of different alleles of species that have hybridized can take different paths through the two parental

populations. Figure 1.6.A shows the true species relationships outlined in bold, with C sister to the A,B clade. At some point in time, individuals of Species B and Species C hybridize, and thus exchange of genetic information occurs between populations. The gene tree of one gene shows that A diverges prior to both B and C, suggesting that B and C are sister taxa (Fig. 1.6.B). In this case, hybridization causes the gene tree to be in conflict with the species tree.

Methods have been derived that attempt to distinguish between hybridization and ILS (Joly et al. 2009; Kubatko 2009; Meng and Kubatko 2009; Gurushidze et al. 2010; Chung and Ané 2011; Yu et al. 2011), yet there are difficulties interpreting results when both of these evolutionary processes are taking place simultaneously. Many of these methods are limited in that the species tree topology must be known, only a single accession per species can be sampled, and no more than two hybridization events can be taken into account. More recent methods have been described to detect hybridization in the presence of ILS and to incorporate multiple accessions per species (Gerard et al. 2011; Yu et al. 2011; Yu et al. 2012), but require more than ten loci to be able to detect hybridization when it occurs. There are methods that can be used when dealing with fewer loci (Maureira-Butler et al. 2008; Joly et al. 2009), but when implemented, the results can still be difficult to interpret (Blanco-Pastor et al. 2012). Another shortcoming of using tree topologies to test for hybridization is that they are not able to estimate the magnitude of hybridization, but only the presence (Maureira-Butler et al. 2008; Joly et al. 2009; Blanco-Pastor et al. 2012).

Other researchers have used population genetic methods in combination with phylogenetics to derive a better understanding of the evolutionary history of specific groups of species when both processes are most likely taking place (Rieseberg et al. 1991; Hughes et al. 2005; Carling and Brumfield 2008; Pinzón and LaJeunesse 2011; de Villiers et al. 2013). There are challenges to understanding species boundaries of closely related species, especially in the face of hybridization. Using population genetics to examine genetic exchange between taxa could help

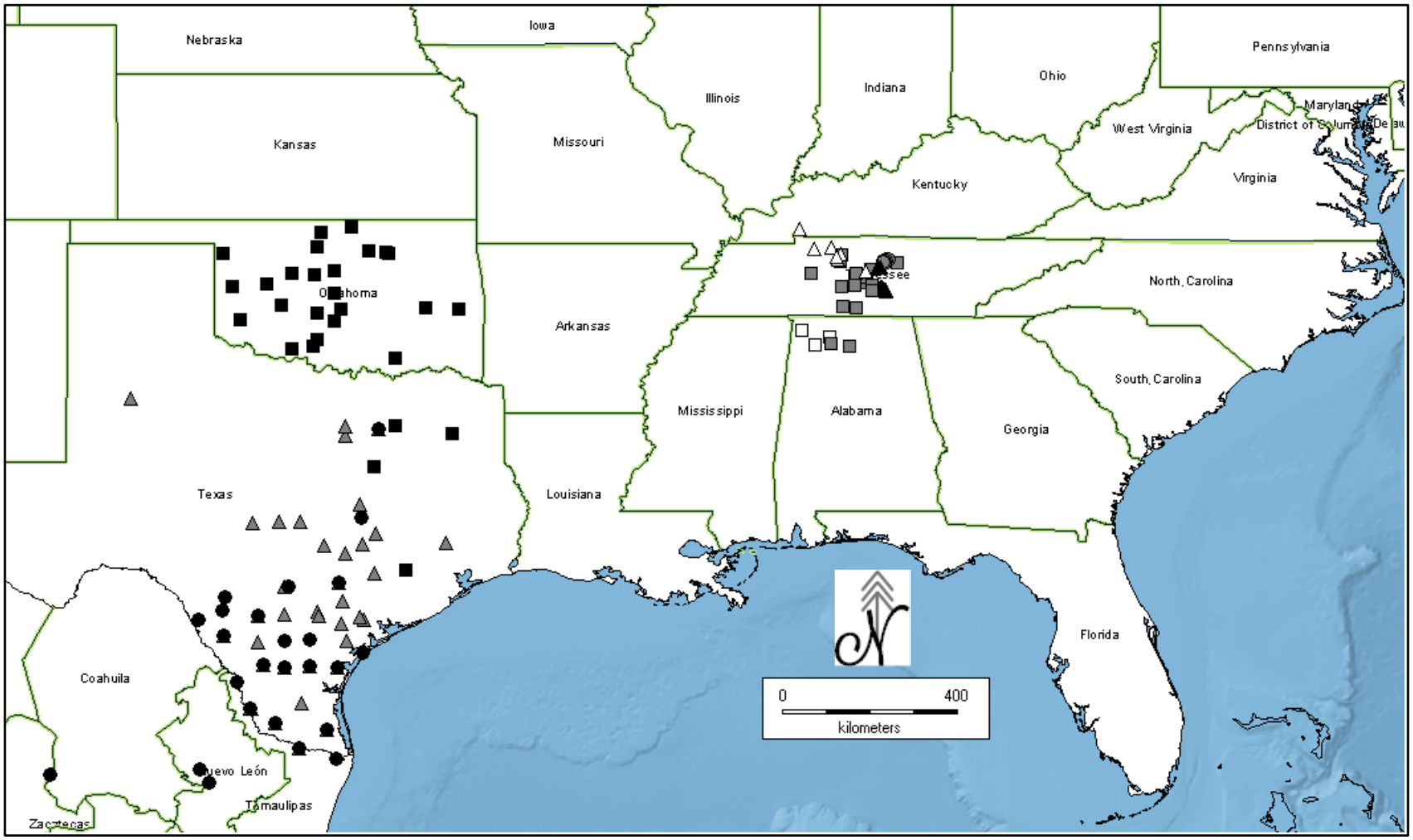
overcome that difficulty. An example of a group where both methods were combined is in *Streptocarpus* Lindl. (Gesneriaceae), a genus of herbaceous plants which contains a clade of species believed to have recently speciated in eastern South Africa during the Pleistocene. Population genetics and phylogenetics approaches have both been used to better understand the evolutionary history of the "Cape Primrose", South African *Streptocarpus* (Hughes et al. 2005; de Villiers et al. 2013). These studies have used both methods to test hypotheses of divergence, dispersal, species delimitations, and genetic structure of the species, and have obtained insight into processes that lead to speciation. Many of the same methods could be applied to understand the evolutionary history of *Paysonia*.

The other two processes that can cause gene trees to be in conflict with species trees are recombination events and gene duplications. A recombination event within a marker may produce a chimeric marker that contains two different molecular phylogenies (Maddison 1997). Gene duplications can be a problem when through inadequate sampling or gene loss, paralogs are compared rather than orthologs. Gene duplications yield a second locus, differing from ILS where multiple alleles compete for the same position at one locus. Duplicated genes evolve and descend independently of one another, which is why comparing paralogs may lead to discordant phylogenies (Maddison 1997). However, sufficient sampling should enable paralogs to be identified.

Paysonia appears to have undergone a recent radiation in the southeastern United States, with little resolution in previous phylogenetic trees (Fuentes-Soriano and Al-Shehbaz 2013; Fig. 1.3). The biology of *Paysonia* consists of many factors that are known to promote ILS, such as large population sizes and persistent seed banks (Degnan and Rosenberg 2009). In addition, some populations of the southeastern *Paysonia* consist of interspecific hybrids (Rollins 1955; Rollins and Shaw 1973), therefore hybridization and introgression must also be considered as possibly affecting species relationships in the genus. Combining population genetic approaches and

phylogenetics may allow better insight into the evolutionary history of the genus. The purpose of this master's thesis is to integrate both levels of study to arrive at a better understanding of the evolutionary history of *Paysonia*.

Figure 1.1: Distribution map of *Paysonia*. Locality information was compiled from recorded herbarium data in the Missouri Botanical Garden Tropicos Specimen Database, the Oklahoma Vascular Plant Database, and data provided by the Tennessee Department of Environment and Conservation. The base map was made using DIVA-GIS v7.5.0, software available at: http://www.diva-gis.org/.



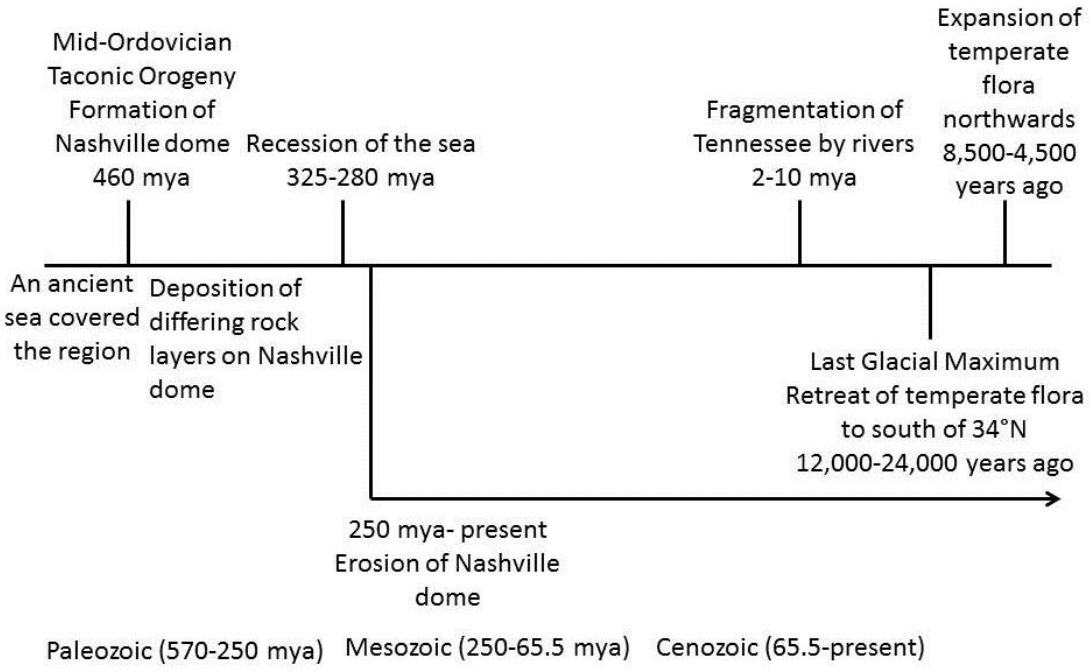


Figure 1.2: Climatic and geologic history of the Central Basin

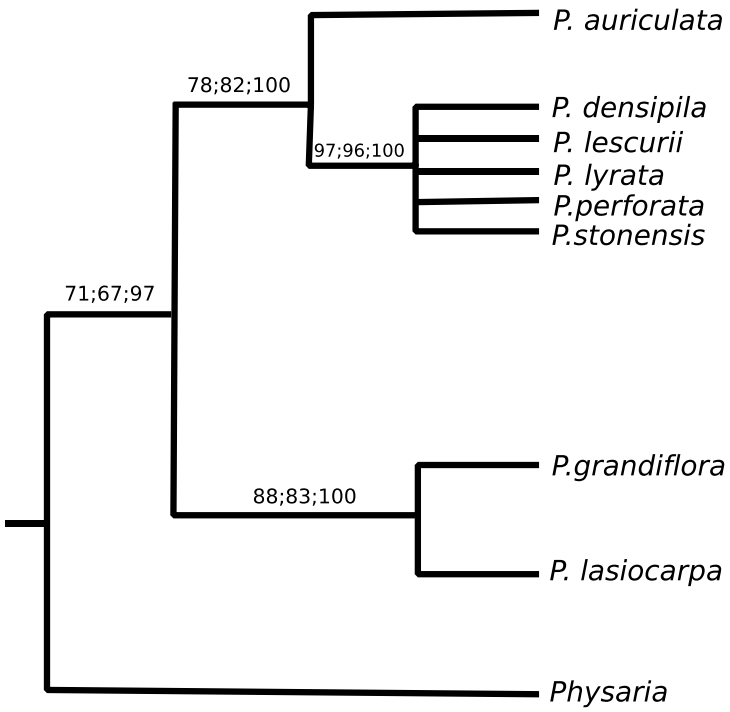


Figure 1.3: Relationships among *Paysonia* species as inferred from an *ndhF* strict consensus tree redrawn from Fuentes-Soriano and Al-Shehbaz (2013). Support values above branches are reported (from left to right) for parsimony bootstrap, likelihood bootstrap, and Bayesian posterior probability. The genus *Physaria* is the sister genus to *Paysonia*.

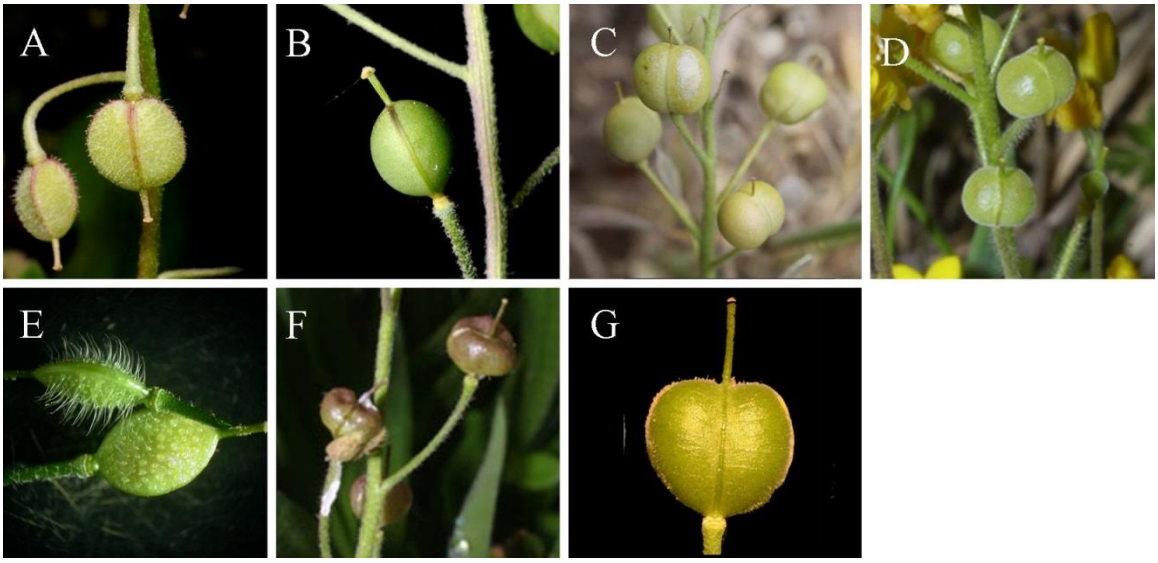


Figure 1.4: Diversity of fruit morphology in *Paysonia.* A) *Paysonia lasiocarpa* (Texas); B) *Paysonia grandiflora* (Texas); C) *Paysonia auriculata* (Oklahoma); D) *Paysonia densiplia* (Tennessee); E) *Paysonia lescurii* (cultivated, seed from wild population in Tennessee); F) *Paysonia perforata* (Tennessee); G) *Paysonia stonensis* (cultivated, seed from wild population in Tennessee)

Table 1.1: Morphological variation among *Paysonia* species

Species	Fruits			Exterior Valves		Interior Valves		Septum	Number of Ovules	Seeds	Flower	Vegetative Trichomes
				Indument	Trichomes	Indument	Trichomes		(per ovary)		Petal Color	
Paysonia lasiocarpa	compressed	angustisep-tate	sessile	densely pubescent	branched, simple large	glabrous	none	complete	14-32	suborb-icular	light yellow	branched or mixed with simple large
Paysonia grandiflora	round	globose, subglobose	sessile, subsessile	glabrous	none	glabrous	none	complete	16-28	orbicu-lar	yellow	stellate
Paysonia auriculata	round	globose, subglobose	subsessile	glabrous	none	glabrous	none	complete	12-16	suborb-icular	yellow	long simple mixed with small branched
Paysonia densipila	round	subglobose	subsessile	densely pubescent	simple or branched	glabrous	none	complete	4-8	orbicu-lar or suborb-icular	yellow	simple mixed with smaller branched
Paysonia lescurii	compressed	latiseptate	sessile	densely pubescent	bulbous-based simple, branched	sparsely pubescent	branched	complete	4-8	suborb-icular	yellow	simple with smaller branched
Paysonia lyrata	round	subglobose	sessile	glabrous	none	glabrous	none	complete	4-8	oval to suborb-icular	yellow	simple proximally, simple with smaller branched, distally
Paysonia stonensis	round	subglobose	subsessile	densely hirsute	simple	glabrous	none	perforate or complete	8-12	oval	white	simple proximally, simple or mixed simple, forked, and slightly branched distally
Paysonia perforata	round	obovoid, subpyriform	subsessile	sparsely hirsute/ glabrate	simple or branched	densely pubescent	branched (dendritic)	perforate, nearly absent	4-12	orbicu-lar or suborb-icular	white, pale lavend-er	simple with smaller, branched

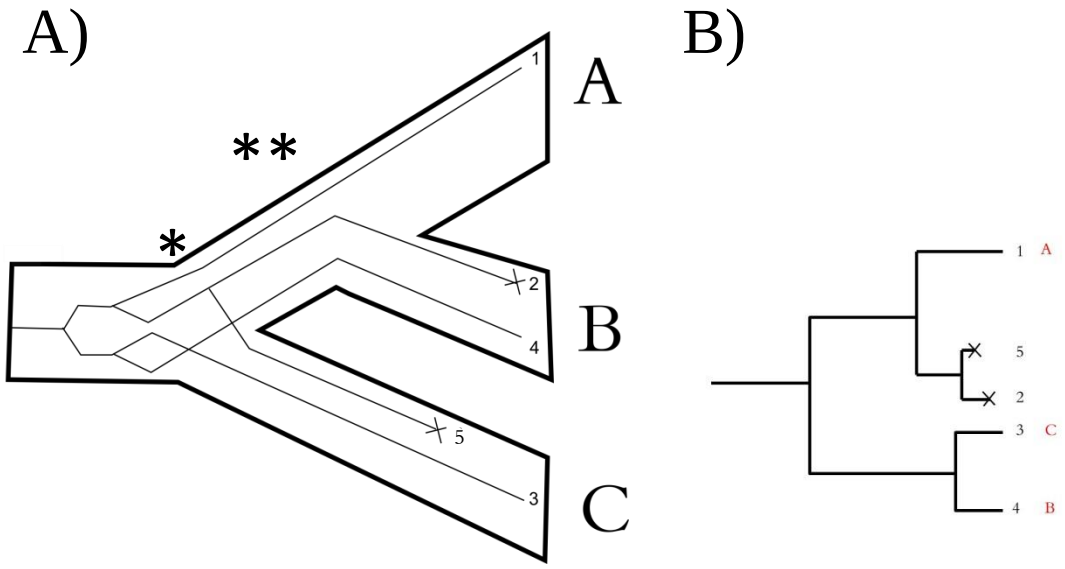


Figure 1.5: Incomplete lineage sorting. In (A) the species tree is depicted with the bold outline and the gene tree is shown within it. In this diagram, Species A and B are sister species, and thus shared a common ancestral population more recently with each other than with Species C. Suppose that at one time, the ancestral population to all three species had four alleles, with two alleles more closely related to one another than the other two gene forms (B). A random event causes the population to initially split (*, Fig 1.5.A), and it so happens that by chance, Allele 1, 2 and 4 are retained in the A,B lineage and only Alleles 3 and a more recently formed Allele 5 are retained in the C lineage. Another speciation event occurs (, Fig 1.5.A) that gives rise to Species A and B., and the same case may arise where not all the genetic diversity is transferred into each of the two populations.**

The extant populations of Species A contain Allele 1, Species B only has had Allele 4 as Allele 2 has gone extinct, and Species C has retained only Allele 3 because Allele 5 went extinct. The gene tree of the sampled Alleles 1, 3, and 4 would be discordant with the species tree, as it would show A(B,C) rather than C(A,B).

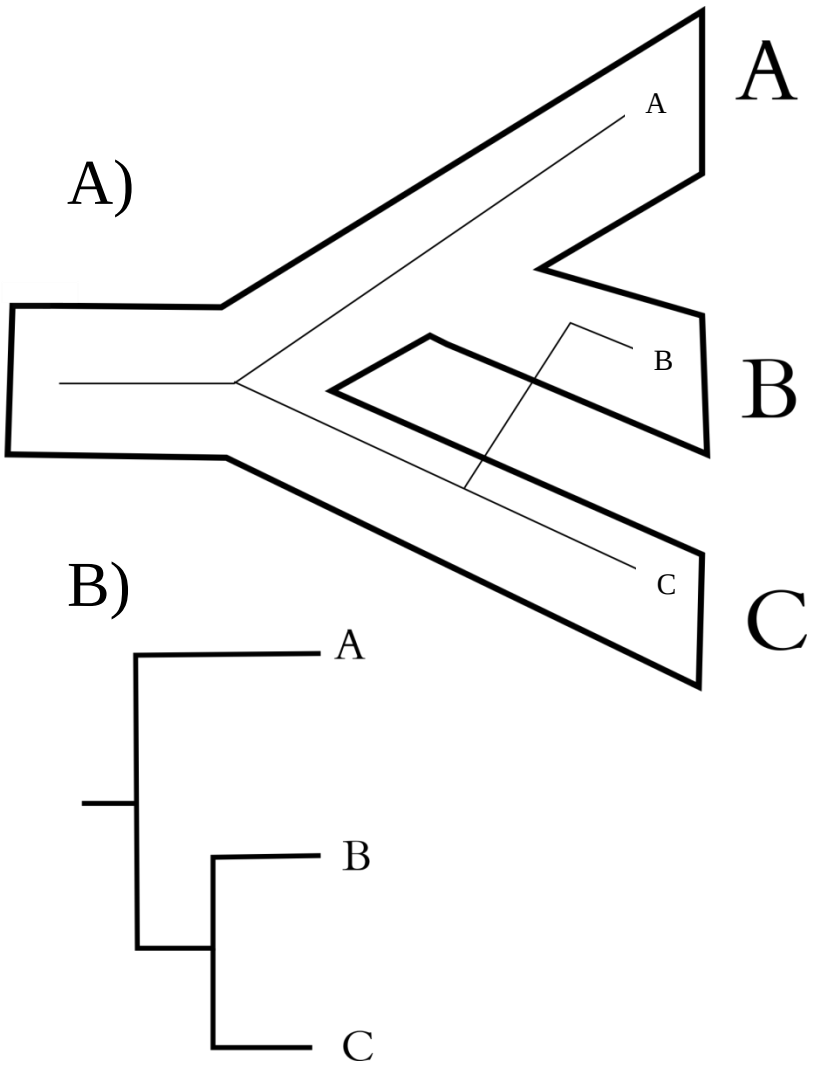


Figure 1.6: Hybridization and gene tree discordance. In A) the species tree is outlined in bold and the gene tree is outlined within the species tree. Hybridization between B and C causes the gene tree B) to show A(B,C), when the species tree shows the true relationship of C(A,B).

CHAPTER II

RECONSTRUCTION OF PHYLOGENETIC RELATIONSHIPS IN PAYSONIA

2.1. INTRODUCTION:

Paysonia O'Kane and Al-Shehbaz is an herbaceous genus in the Brassicaceae comprised of eight species with a disjunct range across the southern United States; one group of three species occurs mainly in Texas, Oklahoma, and Mexico, and the other group of five species occurs predominantly around the Central Basin of Tennessee. All eight species have morphological differences that allow them to be unambiguously identified (Chapter 1, Table 1.1). The genus includes *Paysonia auriculata* (Engelm. & A. Gray) O'Kane and Al-Shehbaz, *P. densipila* (Rollins) O'Kane and Al-Shehbaz, *P. grandiflora* (Hook.) O'Kane and Al-Shehbaz, *P. lasiocarpa* (Hook. ex A. Gray) O'Kane and Al-Shehbaz, *P. lescurii* (A. Gray) O'Kane and Al-Shehbaz, *P. lyrata* (Rollins) O'Kane and Al-Shehbaz, *P. perforata* (Rollins) O'Kane and Al-Shehbaz, and *P. stonensis* (Rollins) O'Kane and Al-Shehbaz. Previously, these eight species were classified with the much larger *Lesquerella* S. Watson, but were formally transferred to a new genus, *Paysonia*, in 2002 (O'Kane and Al-Shehbaz 2002).

All eight species are united by the presence of auriculate, cauline leaves and lateral outgrowths resembling wings on the seeds. The fruits are round siliques that are either inflated or compressed. Within the genus, there is variation in the number of ovules per ovary, ranging

from as few as four in *P. densipila*, *P. lescurii*, *P. lyrata*, and *P. perforata* to as many as 28 in *P. grandiflora*. Most species have a chromosome number of eight with the exception of *P. grandiflora* and *P. lasiocarpa* with n=9 and n=7, respectively. Rollins (1955) hypothesized that the genus originated in the Southwest and wondered from where the "Tennessee *Lesquerellas*" came from and "by what migratory route did they get there?" He believed *P. lyrata* to be the evolutionary link between the Tennessee and southwestern species and that it was perhaps a remnant of a more continuous distribution that connected the two groups of species at one time.

As it stands, a completely resolved phylogeny of the genus does not exist. Fuentes-Soriano and Al-Shebaz (2013) recently published a chloroplast phylogeny of the tribe Physarieae utilizing the *ndhF* marker. The tribal phylogeny included all species of *Paysonia*, with one accession per species represented. The backbone of the tree was resolved with strong support values, and it supports a Texas origin to the genus, as Rollins hypothesized. All of the five Central Basin/Northern Alabama species, however, fall within an unresolved clade. Due to the geological and climatic history of the area (Chapter 1), it is probable that *Paysonia* underwent a recent radiation in that region, and thus, the species have not had sufficient time to become distinct at the genetic level. Biological processes, such as hybridization and incomplete lineage sorting (ILS), could further be confounding inference of the species phylogeny.

With recently diverged lineages, gene sequences may not be sufficiently variable enough for phylogenetic inference. The plastid gene used by Fuentes-Soriano and Al-Shehbaz (2013), *ndhF*, has frequently been successfully used to estimate plant phylogenies, but most of these studies have been at the generic level or above (Bremer et al. 2002; Beilstein et al. 2008). Although this gene might prove variable enough to use in higher-level taxonomic studies, a coding gene may be too highly conserved to use at the interspecific level, due to higher levels of functional constraints. Non-coding regions in the chloroplast have been explored that have the potential of being more informative for use in lower level studies (Shaw et al. 2005; Shaw et al. 2007) and

several studies have employed these non-coding regions to resolve species-level phylogenies (Beck et al. 2006; Granados Mendoza et al. 2013).

For *Paysonia*, a comprehensive phylogenetic study incorporating multiple accessions of each species and multiple genes may provide sufficient data to resolve the phylogeny. It is likely that ILS may produce conflict between the gene trees and the true species relationships in *Paysonia*. Maddison and Knowles (2006) found that incorporating multiple genes and individuals improved the accuracy of inferred species trees, even in the presence of ILS, since some phylogenetic signal can still be derived from the gene sequences. ILS can be modeled using coalescent approaches, such as *BEAST (Heled and Drummond 2010), a Bayesian program for multi-locus data that takes into account the coalescence. This allows the gene to evolve unconstrained within a species tree framework, and requires multi-locus data from multiple individuals per species. This type of methodology has not been employed previously to estimate species relationships in *Paysonia*.

Although the circumscription of *Paysonia* appears well defined, relationships within the genus are not, particularly those of the southeastern species. The main objective of this chapter is to provide an in-depth analysis of the species relationships and evolutionary history of *Paysonia* by sampling multiple individuals from natural populations, incorporating non-coding chloroplast regions (Shaw et al. 2005; Shaw et al. 2007) and the variable internal transcribed spacer region (ITS) of nuclear ribosomal DNA (White et al. 1990), and implementing the multi-species coalescent analysis.

2.2. MATERIALS AND METHODS:

2.2.1. Taxonomic Sampling

A total of 30 accessions were sampled representing eight species of *Paysonia* (Table 2.1). A minimum of three different individuals for each of the eight species in the genus was sampled, as Maddison and Knowles (2006) showed that the accuracy of inferred rooted species tree to be

higher than 50% when using at least three individuals and three loci. Figure 2.1 shows the geographic distribution of the taxa sampled. *Physaria gordonii* and *P. lindheimeri* were chosen as outgroup species based on nuclear and chloroplast sequences from previous studies of the Physarieae that shows *Physaria* sister to *Paysonia* (Al-Shehbaz and O'Kane 2002; O'Kane and Al-Shehbaz 2002; Fuentes-Soriano and Al-Shehbaz 2013).

2.2.2. DNA Extraction

Leaf material for DNA extraction was obtained from cauline leaves on plants collected in the field and preserved in silica gel; specimen information is listed in Table 2.1. Total genomic DNA was extracted from 12-15 mg of silica-dried leaf material using the hexadecyltrimethyl-ammonium bromide (CTAB) extraction protocol (Doyle & Doyle 1987), with slight modifications. These modifications include using zirconia beads and a bead mill to initially grind the leaf tissue before adding the CTAB and β-mercaptoethanol mixture. Nuclease-free water was used to re-suspend the DNA, rather than TE buffer.

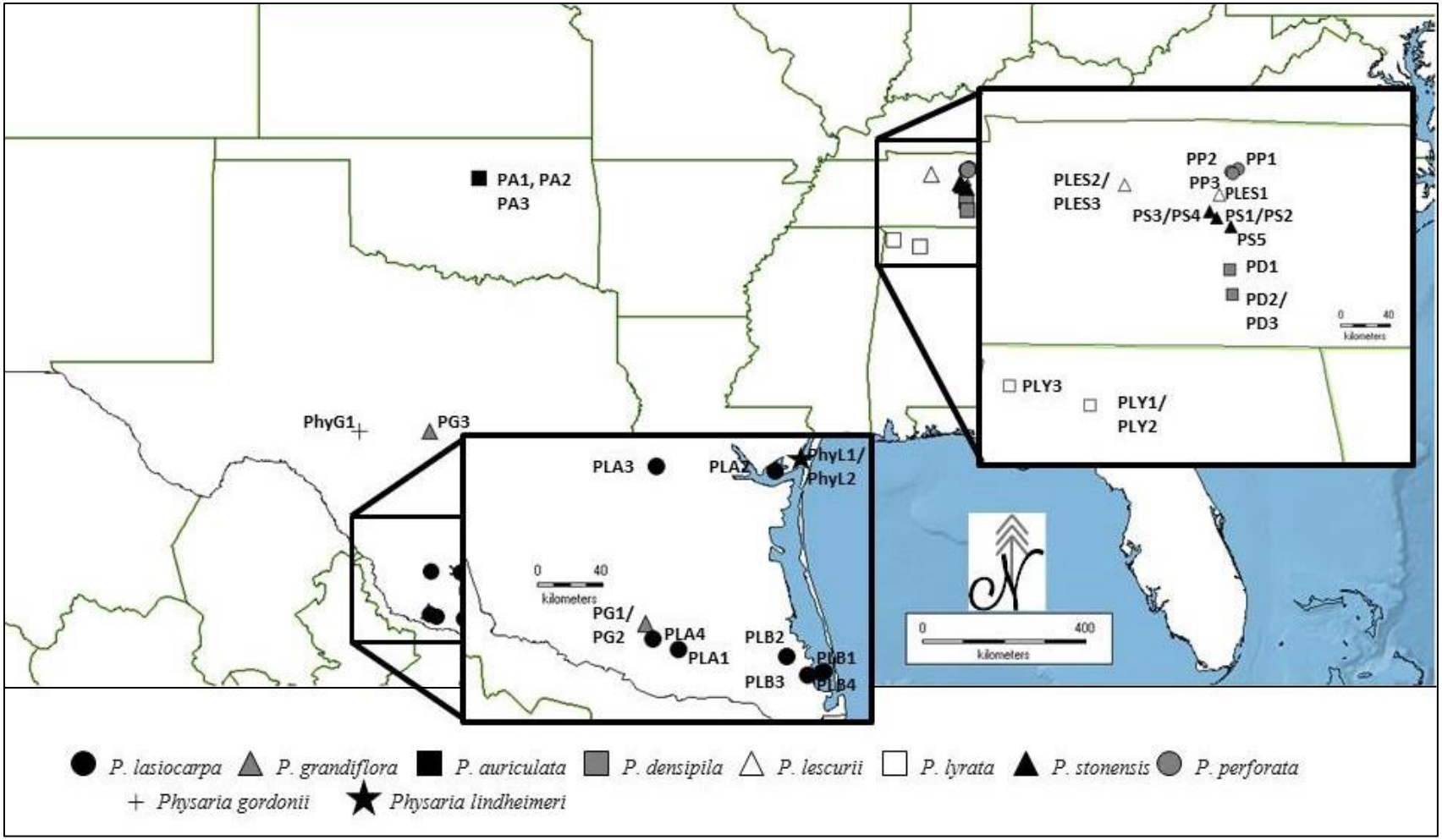


Figure 2.1: Phylogeny Sample Localities: Map of the sites where species included in phylogeny were collected. Boxes enclose zoomed portions of the map southern Texas and middle Tennessee.

Table 2.1: Phylogeny Specimens: The species included in this study and geographic location of the sampled specimens (see also Fig. 2.1)

Species	ID	County	State	Date	Collectors	Latitude	Longitude	Accession	Collecton #
Paysonia auriculata (Engelmann & A. Gray)	PA1	Payne	OK	13-Apr-10	Andrew Doust	36.07	-97.13	BP11	DOUST 2713
	PA2	Payne	OK	13-Apr-10	Andrew Doust	36.07	-97.13	BP12	DOUST 2724
	PA3	Payne	OK	13-Apr-10	Andrew Doust	36.07	-97.13	BQ3	DOUST 2715
Paysonia densipila (Rollins)	PD1	Bedford	TN	9-Apr-08	Andrew Doust	35.56	-86.28	1742	DOUST 1742
	PD2	Coffee	TN	9-Apr-08	Andrew Doust	35.38	-86.26	1757	DOUST 1757
	PD3	Coffee	TN	9-Apr-08	Andrew Doust	35.38	-86.26	1761	DOUST 1761
Paysonia grandiflora (Hooker)	PG1	Hidalgo	TX	1-Apr-10	Andrew Doust	26.47	-98.25	BP7	DOUST 2700
	PG2	Hidalgo	TX	1-Apr-10	Andrew Doust	26.47	-98.25	BP8	DOUST 2702
	PG3	Burnet	TX	10-Mar-10	James Borrone	30.45	-98.23	BP13	DOUST 2773
Paysonia lasiocarpa (Hooker ex A. Gary)	PLA1	Hidalgo	TX	1-Apr-10	Andrew Doust	26.33	-98.07	BP1	DOUST 2691
	PLA2	Kleberg	TX	31-Mar-10	Andrew Doust	27.32	-97.53	BP2	DOUST 2638
	PLA3	Jim Wells	TX	2-Apr-10	Andrew Doust	27.35	-98.19	BP3	DOUST 2711
	PLA4	Hidalgo	TX	1-Apr-10	Andrew Doust	26.39	-98.21	BQ4	DOUST 2697
Paysonia lasiocarpa subsp. *berlandieri* (A. Gray)	PLB1	Cameron	TX	1-Apr-10	Andrew Doust	26.21	-97.26	BP4	DOUST 2653
	PLB2	Rio Hondo	TX	1-Apr-10	Andrew Doust	26.29	-97.47	BP5	DOUST 2684
	PLB3	Cameron	TX	1-Apr-10	Andrew Doust	26.19	-97.35	BP6	DOUST 2680
	PLB4	Cameron	TX	1-Apr-10	Andrew Doust	26.20	-97.27	BQ5	DOUST 2658
Paysonia lescurii (A. Gray)	PLES1	Wilson	TN	9-Apr-08	Andrew Doust	36.11	-86.36	1587	DOUST 1587
	PLES2	Cheatham	TN	10-Apr-08	Andrew Doust	36.18	-87.05	1823	DOUST 1823
	PLES3	Cheatham	TN	11-Apr-08	Andrew Doust	36.18	-87.05	1825	DOUST 1825
*Paysonia lyrata (*Rollins)	PLY1	Lawrence	AL	10-Apr-11	David H. Webb	34.57	-87.30	BC1	WEBB
	PLY2	Lawrence	AL	10-Apr-11	David H. Webb	34.57	-87.30	BC2	WEBB
	PLY3	Colbert	AL	5-Apr-11	David H. Webb	34.71	-87.89	BC3	WEBB
Paysonia perforata (Rollins)	PP1	Wilson	TN	1-Apr-09	Andrew Doust	36.30	-86.22	PpCC1	DOUST 2454
	PP2	Wilson	TN	1-Apr-09	Andrew Doust	36.28	-86.27	PpSC1	DOUST 1966
	PP3	Wilson	TN	2-Apr-09	Andrew Doust	36.27	-86.26	PpSC2	DOUST 2087
Paysonia stonensis (Rollins)	PS1	Rutherford	TN	9-Apr-08	Andrew Doust	35.94	-86.38	1701	DOUST 1701
	PS2	Rutherford	TN	9-Apr-08	Andrew Doust	35.94	-86.38	1709	DOUST 1709
	PS3	Rutherford	TN	9-Apr-08	Andrew Doust	35.99	-86.43	1732	DOUST 1732
	PS4	Rutherford	TN	9-Apr-08	Andrew Doust	35.99	-86.43	1735	DOUST 1735
	PS5	Rutherford	TN	9-Apr-08	Andrew Doust	35.88	-86.27	BQ6	DOUST 2741
Physaria gordonii	PhyG1	Kimble	TX	23-Apr-10	Mark Fishbein	30.47	-99.78	BQ1	FISHBEIN 6484
Physaria lindheimeri	PhyL1	Nueces	TX	3-Apr-10	Andrew Doust	27.34	-97.37	BP10	DOUST 2708
	PhyL2	Nueces	TX	3-Apr-10	Andrew Doust	27.34	-97.37	BP9	DOUST 2709

2.2.3. PCR Amplification

Sequences from four regions were obtained for examining relationships among taxa. These include the nuclear ITS region, and three non-coding chloroplast regions, *trnD*^{GUC}-*trnT*^{GGU} (Shaw et al. 2005), *ndhF-rpl32* (Shaw et al. 2007), and *psbD-trnT*^{GGU} (Shaw et al. 2007). These chloroplast regions were chosen because they had previously been shown to be variable in *Paysonia* (Doust, pers. comm.).

Polymerase chain reaction (PCR) was performed with an Eppendorf Mastercycler® pro (Westbury, New York, USA). The plastid regions were PCR-amplified using the primers specified by Shaw et al. (2005, 2007). The primers, ITS1 and ITS4 (White et al. 1990), were used to amplify the ITS regions. PCR amplifications for the plastid regions was carried out in 25 µL volumes using these final concentrations of the following components: 1 µL template DNA (~10-100 ng), 1X Go*Taq* Reaction Buffer (Promega, Madison, Wisconsin, USA), 1 U of Go*Taq* Flexi DNA polymerase (Promega), 0.2 mM of each deoxynucleioside triphosphate, 2.25 mM of $MgCl_2$, 0.4 pmol/µL of each primer, and 0.2 mg/mL bovine serum albumin. PCR amplifications for ITS utilized the following components: 1X Phusion® HF buffer (New England Biolabs, Inc., Ipswich, Massachusetts, USA), 0.2 mM dNTPs, 0.5 µM of each primer, and 1 U of Phusion® High-Fidelity DNA polymerase in 20 µL volumes. Phusion® High-Fidelity DNA Polymerase was used for ITS because it has a greater than 50-fold lower error-rate than *Taq* polymerase during replication, which is important since the ITS product will be cloned and subjected to further rounds of PCR (see below).

The PCR program used for the *trnD*^{GUC}-*trnT*^{GGU} region is the cycle program outlined by Shaw & al. (2007). A slightly modified cycle program was used for the *ndhF-rpl32* and *psbD-trnT*^{GGU} regions which differed in the annealing temperature. This protocol consists of an initial template denaturation step at 80°C for 5 minutes followed by 30 cycles each of denaturation at 95°C for 1

minute, annealing at 50°C for 1 minute followed by a ramp from 0.3 °C/s to 67°C, and elongation at 67°C for 4 minutes; ending with a final elongation step at 67°C for 5 minutes. The thermocycler program used to amplify ITS began with an initial template denaturation 98°C for 3 min, followed by 36 cycles of a 98°C denaturation for 15 seconds, 50°C annealing for 1 min, and 72°C extension for 1.5 min, ending with a final extension for 20 minutes at 72°C. PCR products were checked on 1% agarose gels and cleaned using the Promega Wizard® SV Gel and PCR Clean-up System (Madison, Wisconsin, USA) or the Epoch Life Science Gencatch™ PCR Purification Kit (Missouri City, Texas, USA).

2.2.4. Cloning of ITS

ITS was cloned using the Invitrogen TOPO® TA Cloning® Kit with the pCR™ 4-TOPO® Vector (Carlsbad, California, USA). Before cloning ITS, single adenosine tails were added to the column purified PCR template to permit the insertion of the blunt PCR products into the T-vector. This was accomplished by adding 5 µg/µL of Go*Taq* Flexi DNA polymerase (Promega), 1.5 mM of $MgCl_2$, 1X Go*Taq* Reaction Buffer and 0.2 mM of deoxyadenosine to 1-4 µL of blunt PCR template. Nuclease-free water was added to bring the final volume to 10 µL and the reaction was then incubated at 70° C for 30 minutes. This was used without further cleanup for ligation into the vector. Half-reactions adapted from the protocol in the Invitrogen Cloning Kit were used for the ligation and transformation reactions. Transformed cells were streaked onto 100 µg/mL Ampicillin agar plates and incubated overnight at 37°C. Colonies were picked and suspended in 12 µL of water and then placed into the thermocycler for an incubation period at 95°C for 5 minutes to lyse the cells. An M13 PCR was used to check for the insert and used 1.5 µl of the DNA template and the M13 (-20) Forward (5´-GTAAAACGACGGCCAG-3´) and M13 Reverse (5´-CAGGAAACAGCTATGAC-3´) primers (Invitrogen). PCR was carried out in 20 µL volumes using the following reagents: 1X Phusion® HF buffer, 0.2 mM dNTPs, 0.5 µM of each

primer, 1 U of Phusion® High-Fidelity DNA polymerase. The following PCR program was employed: an initial denaturation at 98°C for 5 minutes, followed by 30 cycles comprised of denaturation for 30 seconds at 98°C, primer annealing at 52°C for 1 minute, and an extension at 72°C for 2 minutes; ending with a final extension at 72°C for 10 minutes. Products were checked on 1% agarose gels.

2.2.5. Sequencing

Prior to sequencing, PCR products were quantified using the NanoDrop 1000 v3.7 spectrophotometer (ThermoScientific,Waltham, Massachusetts, USA). An attempt was made to sequence eight clones per individual in order to detect paralogous copies of ITS. Sequencing reactions were carried out in 10 µL volumes using 0.5 µl Big Dye v3.1 Terminator Ready Reaction Mix (Applied Biosystems, Foster City, California, USA), 1 µL of 1.6 pmol/µL of the forward or reverse primers listed previously for the plastid regions and the universal T3 (5´-ATTAACCCTCACTAAAGGGA-3´) and T7 (5´-TAATACGACTCACTATAGGG-3´) promoter primers for the cloned ITS products, 0.5 µL of 5X Sequencing Buffer (Applied Biosystems), and around 20-50 ng of DNA. The following profile was used to carry out the reactions: initial template denaturation at 96°C for 30 sec; followed by 35 cycles of denaturation at 96°C for 10 sec, primer annealing at 50°C for 10 sec, and extension at 60°C for 4 minutes. Extension products were cleaned using the Applied Biosystems Ethanol/EDTA precipitation method for Big Dye v3.1 chemistry. Afterwards, 20 µL of nuclease-free water was added to the precipitated product and spun at 3500 RPM for 3 minutes before running on an ABI Prism 3130 or 3100 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA). Contigs were assembled automatically and manually edited with Geneious v5.6.4 (http://www.geneious.com/;Biomatters Ltd., Auckland, New Zealand).

2.2.6. Phylogenetic analysis of sequence data

Two data sets were assembled: one comprised of the cloned nuclear ITS sequences, and the other containing concatenated sequences of the three chloroplast markers. Complete linkage and lack of recombination in the chloroplast genome allows for the concatenation of these non-coding chloroplast regions. Identical sequences from the same individual in the ITS dataset were removed before analysis. Sequences were aligned with Geneious v5.6.4 (Biomatters Ltd.) to automatically determine the correct direction of the sequences and then the alignments were stripped of gaps and realigned using the default settings in the MUSCLE (Edgar 2004) plug-in. Alignments were visually checked and edited using MacClade v4.08 (Maddison and Maddison 2005).

Maximum likelihood (ML) and Bayesian inference (BI) methods were used to estimate the evolutionary relationships among the taxa for each dataset. The ML analyses were performed using the graphical user interface of PhyML v3.0 (Guindon et al. 2010) in Geneious v5.6.4 (Biomatters Ltd.). Nucleotide models of substitution for the concatenated chloroplast data set and the nuclear region were determined by implementing the Akaike Information Criterion (AIC; Akaike 1974) in jModeltest (Posada 2008). Eighty-eight candidate models were tested: these models encompassed 11 different substitution schemes, equal or unequal base frequencies and those with or without a proportion of invariable sites and rate variation among sites. "ML optimization" was selected for the "base tree for the likelihood calculations option". This allowed the program to conduct a tree search for the maximum likelihood topology separately for each model. PhyML analyses were customized to employ the model parameters selected in jModeltest. A nonparametric bootstrap method (Efron 1979; Felsenstein 1985) with 1000 bootstrap replicates was employed to assess the reliability of the internal branches on the phylogeny. The topology, length, and rate of the tree for each bootstrap replicate were optimized

using the nearest neighbor interchange (NNI) heuristic corresponding to the original PhyML algorithm (Guindon and Gascuel 2003).

The BI analyses were performed on each data set using MrBayes v3.2.1 (Huelsenbeck and Ronquist 2001). Optimal models of nucleotide substitution in MrBayes were estimated by applying the AIC as implemented in MrModeltest2 v2.3 (Nylander 2004). To investigate model heterogeneity in the plastid genome, optimal models of evolution for each cpDNA locus were obtained in MrModeltest2 v2.3 and employed in a separate partitioned analysis. In addition, the original two datasets were combined and a partitioned analysis of the nuclear and chloroplast data were performed in MrBayes. PAUP* v4.0 (Swofford 1998) was used to conduct a partitioned homogeneity test prior to combining the datasets to test for topological incongruence that might preclude concatenation. Data and topological congruence were tested with the Shimodaira-Hasegawa nonparametric (SH) test (Shimodaira and Hasegawa 1999) in PAUP version 4.04b. The test was run using full optimization and 1,000 bootstrap replicates. For each analysis, Metropolis-coupled Markov chain Monte Carlo (MCMCMC) simulations were run with eight linked chains (seven heated and one cold). Posterior probabilities were calculated using parameter values and trees were sampled every 1000 generations from the stationary distribution. Two independent runs of 10×10^6 generations were compared to assess convergence to a stationary distribution. These analyses were performed at the OSU High Performance Computing Center (OSUHPCC) at Oklahoma State University on the Cowboy HPC cluster supercomputer. Tracer v1.5 (Drummond and Rambaut 2007) was used to assess the stability of the runs and convergence of the MCMCMC; adequate chain mixing was determined when effective sample size (ESS) values were greater than 200 for each parameter and when the plot of log likelihood values against generations was constant. The average squared deviation of split frequencies was also used as a convergence diagnostic method; convergence was indicated when the average squared deviation of split frequences was less than 0.001. A conservative percentage (10%) of

the posterior samples from each Markov chain was disregarded in order to analyze only data that were within the stationary distribution.

Additionally, a multispecies coalescent approach was used to analyze the datasets. Methods that take into account gene heterogeneity are ideal when using multilocus data sets (Edwards 2009; Kubatko 2009; Liu et al. 2009; Meng and Kubatko 2009). Multispecies/multilocus coalescent-based approaches are believed to be a better estimator for species tree topology than other methods, such as concatenation (Heled and Drummond 2010). To estimate the most probable species tree, *BEAST (Heled and Drummond 2010) was employed in BEAST v1.7.5 (Drummond and Rambaut 2007). The input file for *BEAST was generated using BEAUti, an application provided with the BEAST v1.7.5 package. The site models employed for the individual partitions, the concatenated chloroplast alignment, and the ITS alignment were those estimated previously in MrModelTest2 v2.3 (Nylander 2004). A posterior distribution of phylogenies was produced in *BEAST using a relaxed uncorrelated lognormal molecular clock with the continuous time Markov chain rate reference prior (Ferreira and Suchard 2008) and the Yule tree prior. The continuous time Markov chain reference prior is used to obtain proper posterior distributions when the exact parameters for the elapsed time prior are unknown, making it a good default option. All remaining priors were set to the defaults. The starting tree was randomly generated, and two replicate runs of 50 million generations were performed, sampling trees and parameter estimates every 5000 generations. Convergence was analyzed using Tracer v1.5 (Drummond and Rambaut 2007), and LogCombiner v1.7.5 was used to combine the log and tree files from the independent runs, deleting the first five million generations as determined by average squared deviations less than 0.001 and visual inspection of likelihood traces in Tracer. The maximum clade credibility tree was generated with TreeAnnotator v1.7.5, also part of the BEAST package. Adequate mixing of the chains was determined when ESS values were greater

than 200. A conservative percentage of each run was omitted as burn-in depending on when convergence was achieved.

2.3. RESULTS

2.3.1. Phylogenetic analyses of the nuclear ITS sequence data

The inclusion of gaps to accommodate for indels resulted in an ITS alignment of 696 base pairs (bp) in length. The total variation within the ingroup taxa was 27%, but only 8% within the southeastern taxa (Table 2.2). Nucleotide models of evolution selected were SYM+G in MrModelTest2 and TPM3uf+G in jModeltest. The maximum clade credibility tree after 25% burnin was obtained from the BI, and this tree did not differ significantly in topology in comparison to the ML tree, according to a Shimodaira-Hasegawa test (Shimodaira and Hasegawa 1999). The log likelihood value of the ML tree was -3111.86. The deeper clades on the tree have strongly supported nodes (Fig. 2.2). *Paysonia lasiocarpa* and *P. grandiflora* accessions were each monophyletic with a posterior probability (PP) of 1 and bootstrap support (BS) of 88 and a PP of 1 and BS of 99, respectively. The phylogeny resulted in the placement of *P. lasiocarpa* as sister to a clade consisting of the remainder of *Paysonia* species. *Paysonia auriculata* is sister to the clade comprised of all the southeastern species. There is little support for relationships within the southeastern clade. There are three significantly supported clusters of clones, two of *P. lescurii* and one of *P. lyrata*. Only one of these clusters contains clones from different accessions, PLES1 and PLES2. However, these accessions also contain other clone copies that fall out in other places, but without high support.

Paysonia grandiflora has two significant clusters within it, one of all the PG3 and another containing both PG1 and PG2 clones. *Paysonia auriculata* has two clones from different accessions that do not group with the highly supported cluster of the remaining *P. auriculata* accessions or with other clone copies from the same accessions.

Table 2.2: Nuleotide models of evolution, variabilty, and length of each alignment utilized in the analyses

Marker		Number of characters				
		Total	Variable[a]	Variable[b]	Mutation Models[c,d]	
1	ITS	696	189	55	SYM+G,	TPM3uf+G
2	ndhF-rpl32	921	67	3	GTR+G,	TVM+I
3	trnD	1513	55	6	GTR+G,	TVM
4	psbD	1193	43	8	GTR+I,	TPM1uf+I
	Combined cpDNA	3627	159	17	GTR+G,	TVM+G

Note: Variable[a] = number of variable sites within *Paysonia*; Variable[b] = number of variable sites among the southeastern species in *Paysonia*; Mutation models estimated in c), MrModelTest and d), Jmodeltest.

2.3.2. Phylogenetic analyses of the combined chloroplast sequence data

Combined chloroplast loci yielded an alignment of 3627 bp. The total variation within the ingroup taxa is 4% while that between the southeastern taxa is only 0.5%. Each of the southwestern taxa is strongly supported as monophyletic (Fig. 2.3). The ML tree had a log likelihood value of -6611.42, and has similar topologies to the BI tree. In contrast to the relationships observed with the nuclear ITS, *P. lasiocarpa* and *P. grandiflora* are strongly supported as sister to each other (PP=0.99; BS=98). Both are in a clade that is sister to *P. auriculata* and the southeastern clade. Similar to the ITS, there is little resolution within the southeastern clade, although all the *P. lyrata* accessions form a monophyletic group with the inclusion of one *P. perforata* accession, PP1. This is supported by a high PP value of 0.96 but a moderate BS value of 66.

2.3.3. Phylogenetic analyses of the combined nuclear ITS and chloroplast dataset

Two different methods were used to analyze the nuclear and chloroplast data together. The first was a partitioned analysis, where one of the partitions was the concatenated chloroplast dataset and the other was the nuclear ITS dataset. A maximum clade credibility tree from a BI analysis was calculated after a 25% burnin (Fig 2.4). There is little structure in the southeastern clade, except for some *P. perforata* accessions being sister to *P. lyrata.* One cluster of *P. lescurii* shows an allele shared between PLES1 and PLES2. The second species tree was derived from the *BEAST analysis (Fig. 2.5). *Paysonia lasiocarpa* and *P. grandiflora* placements were only

weakly supported, but the placement of *P. auriculata* as sister to the clade of southeastern species was strongly supported. The southeastern species are placed in a strongly supported clade and within it are two unsupported clades. One is comprised of *P. densipila*, *P. stonensis*, and *P. lescurii*, with *P. stonensis* and *P. lescurii* shown as sister taxa with weakly supported posterior probability of 0.64. The other clade is comprised of *P. perforata* and *P. lyrata.* *Paysonia auriculata* is strongly supported as sister to the southeastern clade.

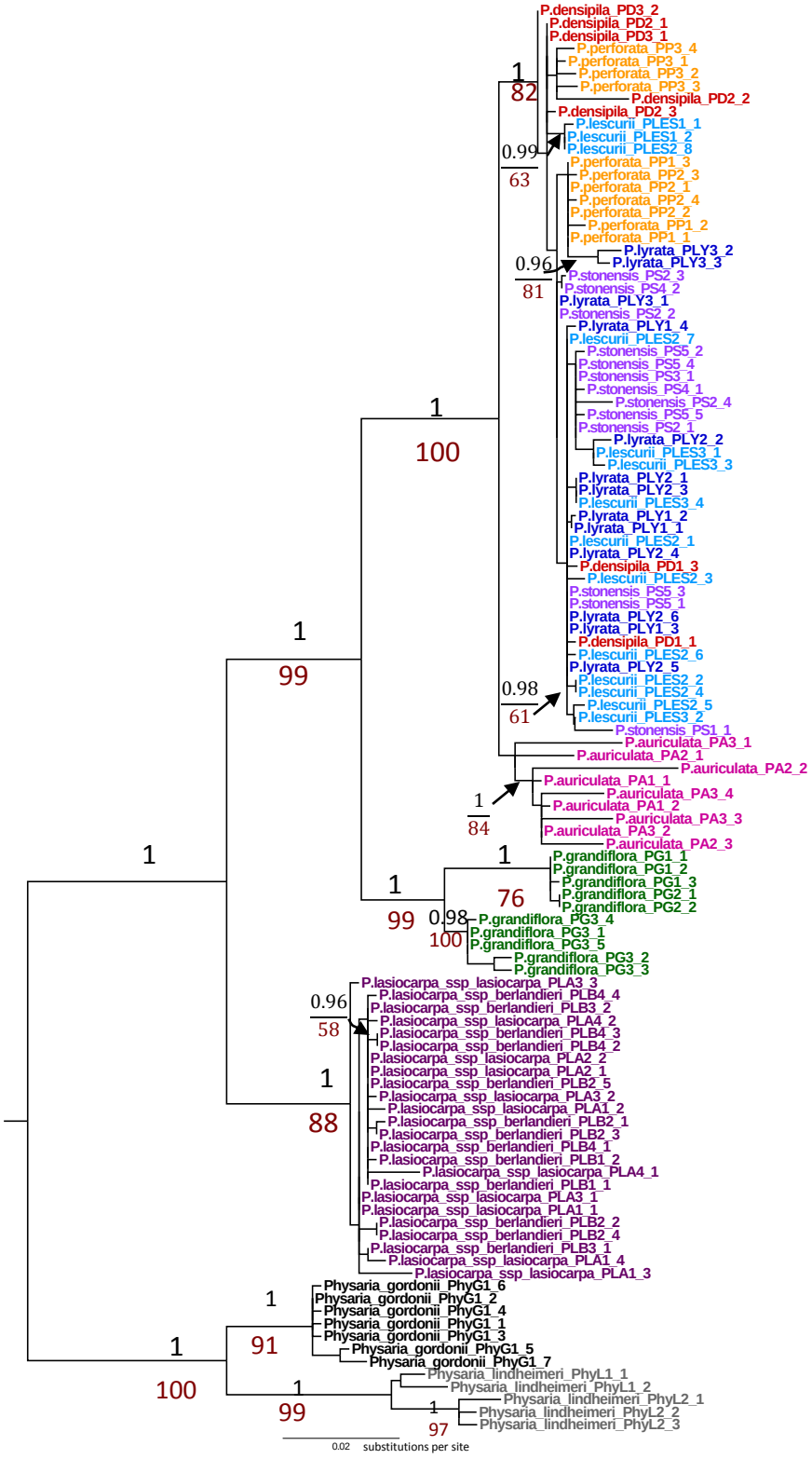


Figure 2.2: Maximum-likelihood tree of the ITS dataset. Numbers above the branches indicate posterior probabilities (PP) for that node. Only PP values ≥0.95 are displayed. Corresponding bootstrap support values (BS) are listed below the branch and arrows point to supported nodes.

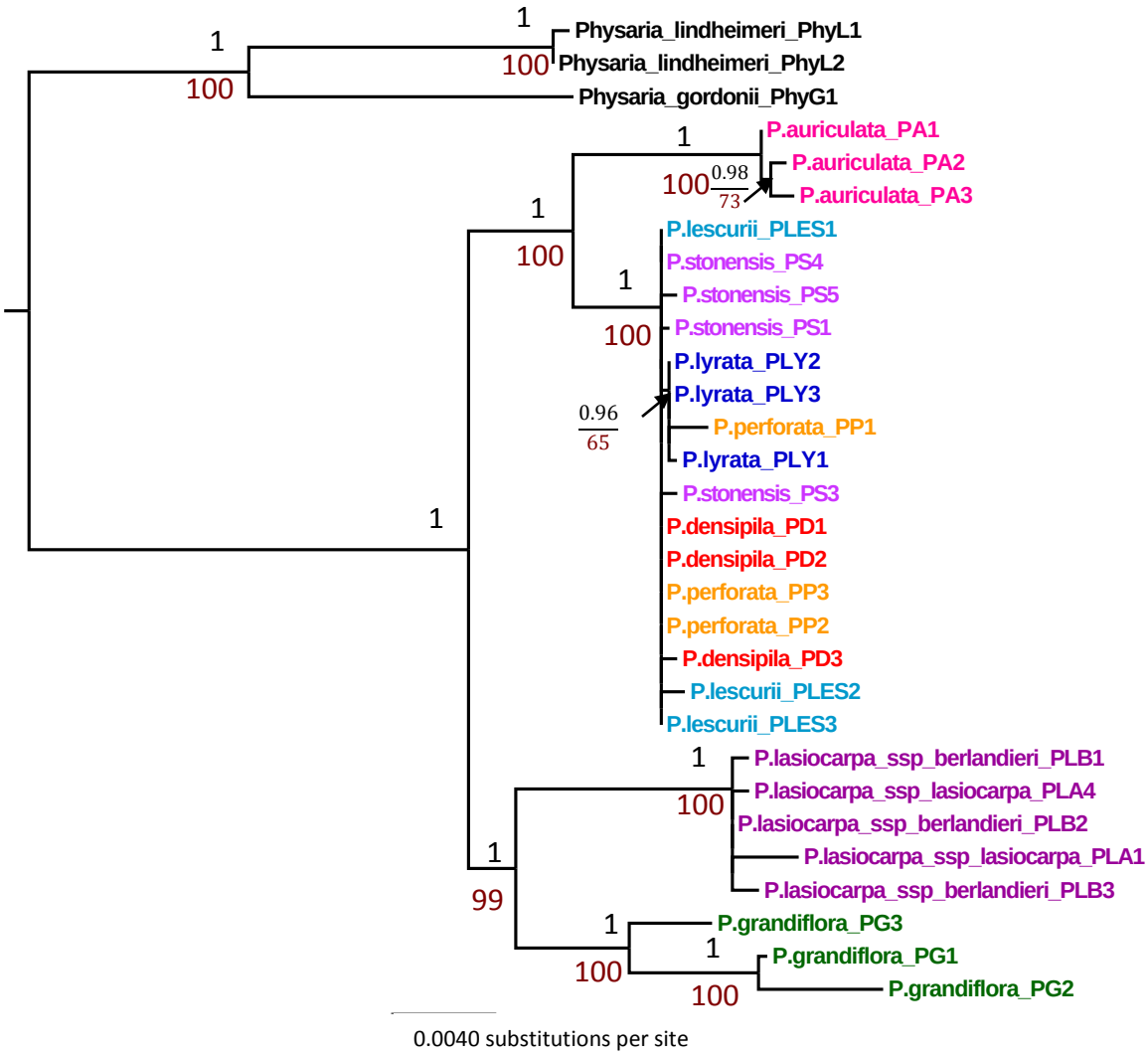


Figure 2.3: Maximum likelihood tree of the concatenated chloroplast alignment. Posterior probability values are listed above the branches, while those numbers below indicate the bootstrap percentage values. Arrows point to supported nodes.

Figure 2.4: Nuclear and chloroplast maximum clade credibility tree generated from the BI. Posterior probabilities above 0.95 are indicated in front of each node, with arrows pointing to supporting nodes.

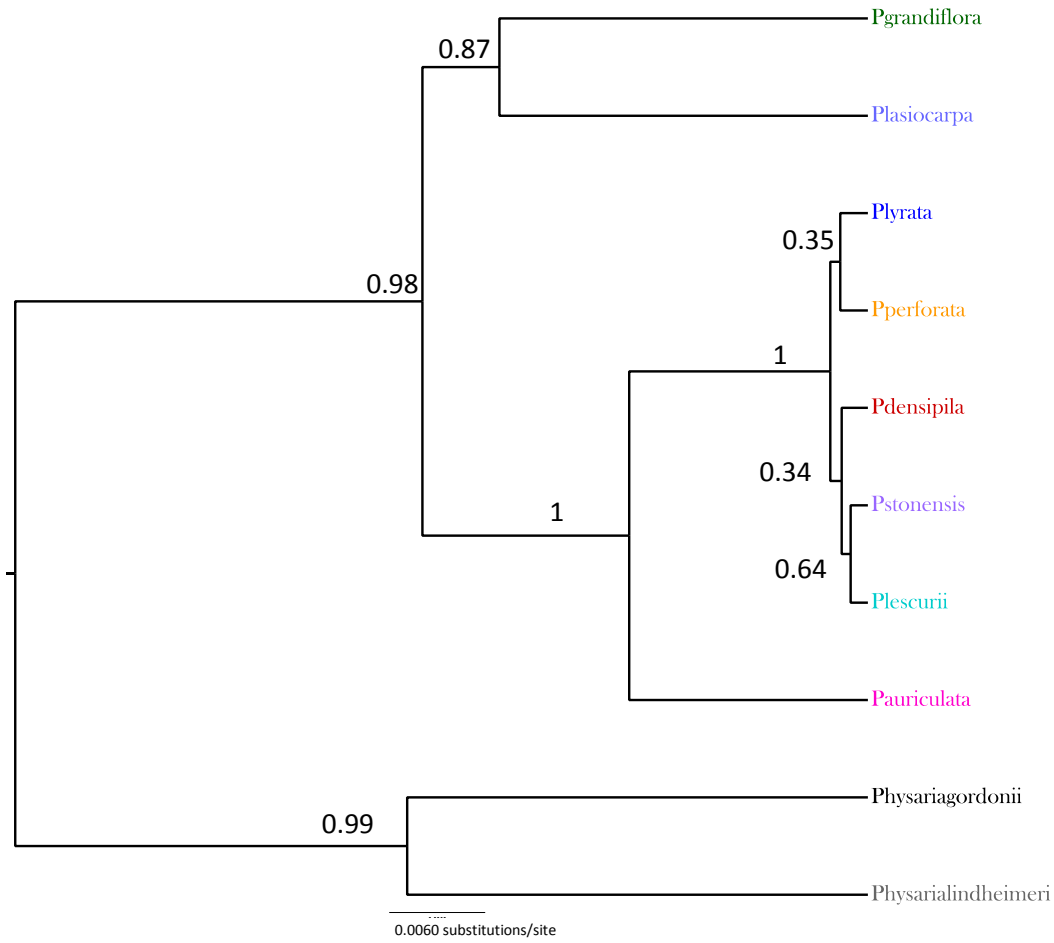


Figure 2.5: Maximum clade credibility tree generated from the *BEAST species-tree inference output for the 4 loci (nuclear ITS and 3 non-coding chloroplast regions). The posterior probabilities for each clade are depicted in front of each node.

2.4. DISCUSSION

Each gene tree (Fig. 2.2 and Fig. 2.3) shows that each of the southwestern species is monophyletic. The relationships between *P. grandiflora* and *P. lasiocarpa*, however, differ depending on what genome was analyzed and the method of analysis. *Paysonia auriculata*, from Oklahoma, is shown to have an intermediate position between the western species and the Tennessee/Alabama species in all analyses. Maximum likelihood and Bayesian gene trees do not resolve the relationships among the southeastern *Paysonia* species, but do support their monophyly. The partitioned Bayesian analyses of ITS and the concatenated chloroplast dataset (Fig. 2.4) does not provide clear insight into the relationships of that group of species within the

southeastern clade either. The species tree derived using *BEAST (Fig. 2.5), however, suggests the presence of two clades within the southeast, albeit not supported. *Paysonia lescurii* and *P. stonensis* are moderately supported as sister taxa within a clade that also includes *P. densipila*. All three species are known to form hybrid populations in the wild, which suggests gene flow may be affecting the phylogenetic signal in the southeastern clade. Another possibility that might explain the lack of support is not having sufficiently variable data and not enough loci.

2.4.1. ITS phylogeny

The relationships of the southwestern taxa are resolved and strongly supported in the ITS tree (Fig. 2.2). *Paysonia lasiocarpa* accessions are shown to be sister to the remainder of the genus. In addition, the data show some structure within *P. grandiflora*, with two strongly supported clades; one containing the two *P. grandiflora* accessions from Hidalgo county and one from Burnet county, located further north in Texas. The topology of the southeastern clade remains unresolved, most likely due to the lack of divergence between gene sequences.

The southeastern clade is presumably a young lineage, but due to the lack of fossils for this genus (as well as scarcity for the Brassicaceae as a whole); the exact age cannot be ascertained. There are also some drawbacks to using ITS that have been identified which might further explain the lack of divergence, such as concerted evolution, which is believed to be fairly common in ITS (Álvarez and Wendel 2003). Concerted evolution would lead to the homogenization of the tandem ITS copies within a species, even for paralogous ITS copies, through homologous recombination events that could convert the gene by copying a particular sequence and overwriting the homologous region of the other with that sequence (Álvarez and Wendel 2003). If gene flow is occurring between species, homogenization of ITS sequences could occur across species boundaries as well; hence the lack of resolution between those species could also be due to hybridization.

2.4.2. The evolution of the plastid genome in *Paysonia*

The chloroplast genome, in comparison to the biparentally-inherited nuclear genome, is maternally inherited in most angiosperms, including *Paysonia.* In addition, the plastid genome, due to its circular nature, does not undergo recombination (Soltis and Soltis 1998). The small genome size and single copy genes of the chloroplast make it valuable for phylogenetic reconstruction. Although the matrilineal inheritance of the chloroplast genome only permits the reconstruction of the maternal lineage, this has the advantage of not being confounded by reticulate evolution that could plague nuclear loci such as recombination of alleles from the different parents.

The chloroplast topology shows little resolution in the southeastern *Paysonia* clade. One clade with strong support consists of three *P. lyrata* accessions clustered together with the one *P. perforata* accession. All the other accessions in the southeastern clade are unresolved. *Paysonia lyrata*, in northern Alabama, is geographically situated relatively far from the other Central Basin species of Tennessee and is restricted to only three localities, so it makes sense that this species is differentiated from the rest. An unanticipated result is that one *P. perforata* accession is placed in the same cluster with *P. lyrata*. One likely explanation is that ILS has caused that particular *P. perforata* individual to retain ancestral polymorphisms shared with *P. lyrata.* Another explanation could be introgression. The short branches of this southeastern complex suggest that this clade is relatively young.

The chloroplast phylogeny clearly delimits the southwestern species, *P. lasiocarpa*, *P. grandiflora*, and *P. auriculata*. However, the plastid genome places *P. lasiocarpa* and *P. grandiflora* as sister taxa and is in conflict with the ITS results, which place *P. lasiocarpa* basal to *P. grandiflora*. This placement by the chloroplast markers, however, matches up with the *ndhF* results from Fuentes-Soriano and Al-Shehbaz (2013).

2.4.3. Incongruence of nuclear and plastid loci

Results from the phylogenetic analyses of *Paysonia* show conflicting relationships between the nuclear and chloroplast data sets. The relationships of *P. lasiocarpa* and *P. grandiflora* differed depending on the marker used. The chloroplast data place *P. grandiflora* and *P. lasiocarpa* as sister species. The nuclear ITS data place *P. lasiocarpa* as basal to the entire genus. The ITS phylogeny's placement of *P. lasiocarpa* and *P. grandiflora* agrees with dissertation work by Fuentes-Soriano (2010) that utilized the nuclear marker, LUMINIDEPENDENS. The Bayesian partitioned analyses of the concatenated data set show a strongly supported tree with *P. lasiocarpa* basal to the genus; but the *BEAST species tree shows *P. lasiocarpa* and *P. grandiflora* as sister, although the relationship has a moderate PP value of 0.86. The partitioned homogeneity test resulted in a highly significant difference between the ITS and chloroplast data set (p-value=0.01), so the partitioned Bayesian analysis should be interpreted with some caution. More nuclear data need to be explored in order to determine the relationship of *P. lasiocarpa* to *P. grandiflora.*

The incongruence between the nuclear and chloroplast trees for the relationships between *P. grandiflora* and *P. lasiocarpa* is unlikely due to ILS. Long branches on every tree for those two particular species suggest sufficient generations after speciation for sorting to have occurred. An alternative explanation for the incongruence is the phenomenon of chloroplast capture. Through some introgression/hybridization event, the *P. lasiocarpa* and *P. grandiflora* lineages may have shared a chloroplast. Phylogenies are known to be affected by this phenomenon, which is thought to be a major reason behind many nuclear and chloroplast gene tree discordances (Rieseberg and Soltis 1991). This has been seen in many plant groups (Soltis and Kuzoff 1995; Mort et al. 2002; Tsitrone et al. 2003), especially in groups noted for hybridization. Complete capture of a chloroplast can theoretically occur relatively quickly in a population, due to the small effective population size of the chloroplast genome (Tsitrone et al. 2003). Hybridization between

species followed by intensive backcrossing to the parental species is thought to be the cause for plastid capture. 'Spontaneous' androgenesis has been proposed as a plausible explanation for chloroplast capture, and has been observed to occur with some frequency in other angiosperms, including the related crop species, *Brassica napa* (Hedtke and Hillis 2011). What might occur is that through some mechanistic failure in meiosis, a sperm cell would not be reduced. When this unreduced sperm of one species unites with a reduced egg from a different species, this would lead to the replacement of the egg nucleus by the sperm nucleus instead of the two nuclei uniting. This would then produce offspring with a paternal nuclear genome, but a maternal plastid genome. Further research is still needed to establish whether chloroplast capture may explain the inferred phylogenetic patterns. Nonetheless, these data show that solely relying on chloroplast data to determine species relationships in *Paysonia* may yield misleading results and thus, examining additional nuclear loci is needed to make conclusive statements about the relationships of *P. lasiocarpa* to *P. grandiflora*.

2.4.4. Recent radiation in the southeastern U.S.

In both the ITS and chloroplast gene trees, little resolution is in the clade containing the southeastern species. The chloroplast dataset exhibited low levels of variation (0.5%) within the southeastern clade, although ITS was more variable at 8%. Strong basal support for the Texas and Oklahoma species and short branches for the species in the southeastern clade suggests that the southeastern species had differentiated more recently. Although *BEAST resolved relationships among the southeastern species, none are strongly supported. Two weakly supported clades are found within the southeast, and two species, *P. lescurii* and *P. stonensis*, have moderate support suggesting that they may have shared a more recent common ancestor. The partitioned Bayesian analysis which incorporated both data sets showed some significant clades (Plescurii_1823, Plescurii_1587, Pstonensis_BQ6, Plyrata_BC1, Pperforata_PpCC1), but they are comprised of clones from the same accessions, except for a set of Plescurii_1587 clones

grouped with one copy from Plescurii_1823. More than likely, this just means that the two different accessions shared a similar copy of ITS, although they are from different populations in different counties.

Geological history may have played a role in the recent speciation of the southeastern group of *Paysonia*. Multiple orogenic events during the Paleozoic and Mesozoic caused an uplift of the Nashville dome, so that the sea, which previously covered the region, eventually drained. Subsequently, erosion greatly altered the uncovered fractured strata, exposing an older limestone layer formed during the Precambrian (Safford 1869; Wilson 1962). The continued erosion eventually led to the modern characteristic bowl-like structure of the Central Basin, with tributaries of the Cumberland and Tennessee rivers running extensively throughout this region.

During the Last Glacial Maxima (LGM), 24,000 to 12,000 years before the present, it is believed that suitable cedar glade habitats could only be found south of the 34° N latitude due to unfavorable climatic conditions that affected the Central Basin (Delcourt et al. 1986). Northern Alabama may have provided refugia for *Paysonia* (Hewitt 2000). Once temperatures began to warm and soils became drier and less water logged (8500-4500 BP), the cedar glade habitat expanded north into Tennessee (Delcourt et al. 1986). Postglacial speciation may have occurred when a more widespread cedar glade ancestral species expanded northwards. Although the southeastern *Paysonia* are endemic to the Central Basin (with the exception of *P. lyrata* in Alabama), only *P. densipila* and *P. lyrata* are considered true cedar glade endemics (Baskin and Baskin 2003), and both are the most southern species. Cedar glade species tend to have weedy tendencies (Baskin and Baskin 2003), so it is probable that the common ancestor to the southeastern clade may have easily moved from a cedar glade habitat and differentiated along the different drainage systems as the ability to occupy a more northern habitat occurred.

Another possibility is that during the LGM, the distribution of *Paysonia* may have been fragmented into areas with suitable microclimates for cedar glade endemics. The species we now observe might be derived from fragmented populations that survived the LGM. With the postglacial warm-up, these fragmented populations may have differentiated and then expanded their range.

Known hybrid populations exist between some of the southeastern species; naturally occurring hybrid populations of *P.densipila* x *P. lescurii*, *P.densipila* x *P. stonensis*, and *P. lescurii* x *P. stonensis* can be found. While coalescent methodology, such as *BEAST, is a common method used to attempt to resolve species relationships, it assumes any conflict in the data is due to ILS (Degnan and Rosenberg 2009). It disregards hybridization as a process that can hamper species tree inference and distinguishing between ILS and hybridization at the phylogenetic level is not always easy. However, the lack of divergence between sequences of the species in the southeastern clade makes it difficult to disentangle the cause as being due to hybridization or ILS. With the exception of *P. lyrata*, the lack of support for the monophylyl of the species in the southeastern clade is likely due to the relatively young age of the species, and perhaps hybridization. A more complete history of the species can only be inferred by including additional nuclear loci that are more informative, along with examining present day distribution and population processes to ascertain the evidence for interspecific hybridization via introgression.

Morphological characters in the Brassicaceae have been found to lead to incorrect assumptions about species relationships due to convergent evolution (Al-Shehbaz 2011). Nonetheless, if one were to hypothesize relationships among the southeastern species based on morphology, it appears that *P. lyrata* and *P. densipila* are more closely related to one another than to the other species. *Paysonia. lescurii* has the most distinctive fruits, while *P. stonensis* and *P. perforata* are the only two white-flowered species and were thought by Rollins (1955) to be very closely

related. The data, however, do not support this. In the partitioned Bayesian analyses, the *BEAST maximum clade credibility tree (Fig. 2.5) and the plastid phylogenetic tree (Fig. 2.3), *P. lyrata* appears to be more closely related to *P. perforata* rather than to *P. densipila*. The *BEAST maximum clade credibility tree also suggests that *P. stonensis* is sister to *P. lescurii*, instead of to the other white flowered species, *P. perforata*. This may be an example of convergent evolution in morphological characters leading to incorrect assumptions of species relationships in *Paysonia*.

CHAPTER III

POPULATION GENETICS OF SOUTHEASTERN PAYSONIA

3.1. INTRODUCTION

The Central Basin of Tennessee houses species diversity for *Paysonia*, where five out of the eight *Paysonia* species grow in close proximity. In the previous chapter, the phylogenetic relationships of *Paysonia* were explored. These analyses clearly showed that while each species in the southwestern group (*P. auriculata, P. lasiocarpa,* and *P. grandiflora*) is reciprocally monophyletic, this is not the case for the southeastern species. The short branch lengths in the reconstructed phylogeny of *Paysonia* suggests that species in the Central Basin may have diverged relatively recently, and that the genes chosen do not evolve rapidly enough to track these recent divergences and thus adequately distinguish among species. It may also indicate active and ongoing gene exchange between species. The southeastern species include *P. lescurii, P. densipila, P. stonensis, P. perforata,* and *P. lyrata,* which are characterized by differences in several morphological traits (fruit shape, petal color, trichome shape and density). This complex of species can freely interbreed with one another (Rollins 1988).

Each *Paysonia* species in the Central Basin is associated with a particular watershed; *P. densipila* occurs in the Duck River watershed, *P. lescurii* in the Cumberland River watershed, *P. perforata* with Spring, Barton, and Cedar Creeks, and *P. stonensis* with the Stones River. It is suspected

that the southeastern species have not developed strong reproductive barriers because historically their distributions were relatively isolated along specific watersheds (Rollins 1954). However, their present day ranges, especially those of *P. densipila* and *P. lescurii* on the larger rivers, encroach upon different regions where they historically were not found. Rollins (1954) posited that the anthropogenic disturbance has caused the species to spread throughout the region, mainly because of agricultural practices that facilitated the establishment of populations. The migration of *Paysonia* species within the region, and their ability to interbreed, increases the likelihood of interspecific gene flow and introgression where different species come into contact.

The extent to which gene flow via hybridization has played a role in this region is not well known, and it is possible that the lack of differentiation in molecular markers among species in this region may be due to current gene flow in addition to recent divergence (Chapter 2). Naturally occurring hybrid populations of *P. densipila* x *P. lescurii*, *P. densipila* x *P. stonensis*, and *P. stonensis* x *P. lescurii* were studied by Rollins (1952, 1954,1955, 1988) and are known to still exist. *Paysonia densipila* x *P. lescurii* hybrids are found at and downstream from the junction of Arrington Creek and the Harpeth River; *P. densipila* x *P. stonensis* hybrids are found at and downstream from the junction of the east and west forks of the Stones River; and *P. lescurii* x *P. stonensis* hybrids are known from Rutherford County, Tennessee (Rollins 1952; Rollins and Shaw 1973; Rollins 1988). Rollins observed that most hybrid populations show great variation in their morphological traits and most likely represent a variety of additive genotypes, since they resembled F2 and subsequent generation hybrids (Rollins 1957; Rollins and Solbrig 1973). However, during some years, Rollins and Solbrig (1973) observed that the upstream hybrid populations (which are closer to the junction of different watersheds and to the parental populations) resembled one parent more than the other. This is explained by the uneven flooding patterns that can occur, which would introduce an influx of seeds from mostly one parent if its associated river flooded. In the more stable populations, no readily observable evolutionary

morphological trends are noted, suggesting that these hybrid populations are relatively young. Rollins (1954) hypothesized that through unusual circumstances an initial hybrid population formed downstream of populations of the parental species and that the majority of subsequent hybrid populations were established from viable hybrid seeds from the initial hybrid population, and that these populations maintain themselves.

Hybridization between the populations of 'pure' species in Tennessee has not yet been thoroughly explored at the molecular level. With the existence of natural hybrid populations and the expansion of the distribution of species, it is possible that gene flow through introgression is also occurring between the otherwise distinct populations of the different species. Gene flow is considered to be a homogenizing force that counteracts speciation (Freeman and Herron 2001). The low genetic distinctness at the phylogenetic level might be due to this admixture. In this study, the main objective is to use rapidly evolving markers to explore whether morphologically distinct populations of species are also genetically distinct. Because the genes used in the phylogenetic analyses showed little variation, we used microsatellites to infer population genetic structure and identify divergent gene pools at the spatial level. Microsatellite regions evolve at a much faster rate than DNA sequence evolution of genes, which is mostly due to DNA replication slippage (Schlötterer 2000), and are commonly used to infer population structure within species. Using microsatellites can help determine whether gene flow is occurring between morphologically distinct populations, and can add further insight into speciation.

3.2 MATERIALS AND METHODS

3.2.1. Population-level sampling

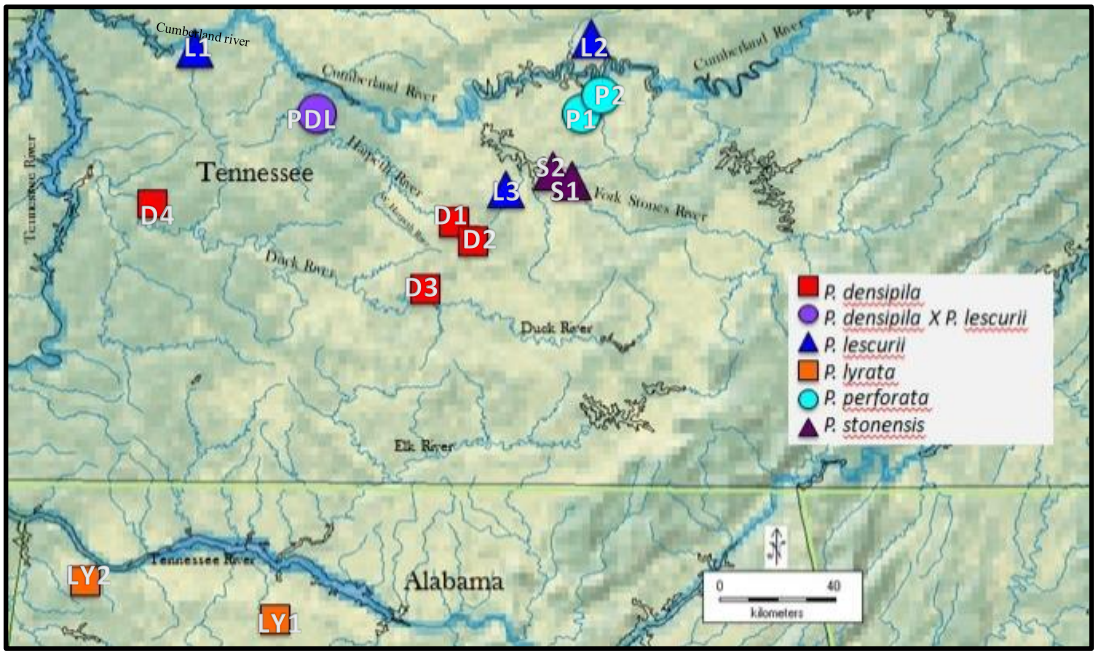
Multiple populations of each southeastern species were located and sampled during the spring of 2008, 2010 and 2011. Fourteen populations were genotyped for this study including three populations of *P. lescurii*, four of *P. densipila*, and two of each *P. lyrata*, *P. perforata*, and *P.*

stonensis. One *P. lescurii x P. densipila* hybrid population was also sampled (Table 3.1; Figure 3.1). Individuals used in the phylogenetic analysis were also incorporated into the microsatellite analysis, although some individuals from the phylogeny could not be incorporated into one of the 14 more intensively sampled populations because they were collected from other sites. Each population was randomly sampled along a randomly placed transect that spanned the length of the population. Only single branches were taken from each sampled plant, rather than uprooting entire individuals. The number of individuals sampled varied with population size. GPS coordinates were recorded and geographic distances calculated between populations using the software GenAlEx v6.5 (Peakall and Smouse 2006).

Species	Pop. ID	County	State	Latitude (N)	Longitude (W)	No. Sampled	Collector/Collection #
Paysonia lescurii	L1	Stewart	TN	36.397	-87.537	12	BORJA 1000.1-1000.12
	L2	Montgomery	TN	36.414	-86.285	12	BORJA 1001.101-1001.112
	L3	Rutherford	TN	35.931	-86.575	11	BORJA 1005.351; 1005.353-1005.362
	L4*	Wilson	TN	36.110	-86.365	1	DOUST 1587
	L5*	Cheatham	TN	36.178	-87.049	2	DOUST 1823, 1825
Paysonia densipila	D1	Williamson	TN	35.829	-86.698	12	BORJA 1006.480-1006.483; 1006.486-1006.493
	D2	Williamson	TN	35.807	-86.662	12	BORJA 1008.580-1008.583; 1005.588-1005.595
	D3	Maury	TN	35.622	-86.805	12	BORJA 1015.700-1015.703; 1015.708—1015.715
	D4	Hickman	TN	35.883	-87.688	12	BORJA 1017.870-1017.881
	D5*	Bedford	TN	35.564	-86.281	1	DOUST 1742*
	D6*	Coffee	TN	35.380	-86.258	2	DOUST 1757*, 1761*
Paysonia lyrata	Ly1*	Colbert	AL	34.712	-87.894	13	WEBB B15-B26; BC3*
	Ly2*	Lawrence	AL	34.567	-87.302	14	WEBB A28-A39; BC1*-BC2*
Paysonia perforata	P1	Wilson	TN	36.221	-86.313	12	DOUST 2179-2182; 2193; 2202; 2211; 2229; 2239; 2249
	P2*	Wilson	TN	36.276	-86.271	14	DOUST 1964; 1966; 1974; 1983; 1992; 2000; 2007; 2019; 2029; 2041-2044; 2087
	P3*	Wilson	TN	36.300	-86.219	1	DOUST 2454*
Paysonia stonensis	S1*	Rutherford	TN	35.940	-86.378	19	DOUST1700-1709; 1711-1717; 1721-1722
	S2*	Rutherford	TN	35.988	-86.426	5	DOUST 1731-1735
	S3*	Rutherford	TN	35.881	-86.274	1	DOUST 2741*
Paysonia densipila x P. lescurii	PDL	Dickson	TN	36.190	-87.167	12	BORJA 1011.251-1011.262

Table 3.1: Description and sample size of natural populations sampled: *Paysonia lescurii, P. densipila, P. densipila x P. lescurii, P. lyrata, P. perforata*, and *P. stonensis* populations. Asterisks indicate populations that incorporate the individuals from the phylogeny. Populations of sample size ≤2 are populations from the phylogeny that could not be incorporated into the larger, more intensively sampled populations and were left out of genetic diversity estimates.

Figure 3.1: Map of the Population Localities: Tennessee and Alabama populations included in population structure analyses and the drainage systems on which they are found



3.2.2. DNA extraction and amplification of the microsatellite regions

For the 14 populations sampled in this study, DNA was extracted from at least 12 individuals per population using the modified CTAB protocol (Doyle and Doyle 1987) described in the previous chapter. Nine nuclear microsatellite regions were screened for these populations. Two of the nine, BF18 and B07, were designed for *Boechera* (Brassicaceae) (Song et al. 2006; Schranz et al. 2007) and had been tested on *P. perforata* (J Borrone, pers. comm.). The remaining seven primer pairs were specifically designed for *Paysonia* using 454-transcriptome data obtained previously from *P. lescurii* and *P. stonensis* fruit, analyzed through the Cotton Microsatellite Database (CMD) http://www.cottonssr.org (Borrone unpublished data). The new markers are R3c01703; R3c00750; R3c06172; R3c00852; R3c07671; R3c00234; and R3c00555-2.

PCR amplification and fluorescent-tagging was carried out in a volume of 10 µl, containing 1 X Go*Taq* Reaction Buffer (Promega, Madison, Wisconsin, USA), 2.25 mM $MgCl_2$, 0.2 mM dNTPs, 0.1 pmol/µl tailed primer, 0.3 pmol/µl un-tailed primer, and 0.2 pmol/µl of fluorescent primer

(NED, PET, VIC), 1 U of Go*Taq* Flexi DNA polymerase (Promega), and 1 µl template DNA (~10-100 ng). Amplification was conducted in an Eppendorf Mastercycler® pro (Westbury, New York, USA). During the initial cycles of the amplification process, the tailed and un-tailed forward and reverse primers amplify the microsatellite. The amplified products are labeled with the fluorescent primer as it incorporates itself onto the tailed primer, as described by Schuelke (2000). The thermocycler profile began with a 94°C initial denaturation for 2 minutes, followed by 38 cycles each consisting of a 94°C denaturation for 45 seconds, 55°C or 59°C primer annealing for 45 seconds, and a 72°C extension for 30 seconds, ending with a final extension for 1 minute at 72°C. Primer sequences and annealing temperatures used are listed in Table 3.2.

Fluorescently-tagged microsatellite products were run on a 3130 Applied Biosystems Genetic Analyzer (Foster City, USA) using the GeneScan™ - 600 LIZ® Size Standard v2.0 (Applied Biosystems). Markers were pooled into several groups: R3c01703, R3c06172, and R3c00750; R3c07671, R3c00852, and R3c00555-2; and B07 and R3c00234. Marker BF18 was run on its own. Peak data were analyzed in GeneMapper v4.0 (Applied Biosystems, Foster City, USA).

Table 3.2: Microsatellite primer pairs, annealing temperatures, and bp length

Name/ID	Forward	Reverse	Annealing Temp.	bp
R3c01703	(VIC)CGTGGGAAAATTCTCCTGAA	TTCTGCATTTAGCATTTGTCA	55° C	225-233
R3c00750	(NED)TCTGATTCATCCTCGGTCGTTG	CAGATTTAATTTTCGTTTCCTTCC	55° C	250-259
R3c06172	(PET)TAGACCCRACCAAAGGACCA	CCTTCACCAAAGCTTCTTGC	59° C	142-184
R3c00852	(PET)CCTTCCTCTATTTCCCCTCG	GTGTTTACCACCTGAGACATATCCA	59° C	97-100
R3c00234	TGGCCGTCTTGTCTCTTAGGTC	(PET)TCAGATCCAAACCC	59° C	261-288
R3c00555-2	AAGCCAAAAGGGTGTTTTGA	(NED)GTCACCAATACGTCAAAGTCCG	55° C	120-128
R3c07671	(VIC)GTGGGATGTTTGCTGGACTT	CCTCACAGATGGTTCACTGG	55° C	283-308
B07	(NED)CGGGAAGATTCAGCAGGTAA	TCCTTTCCTCTCTTTATCCATCA	55° C	148-152
BF18	AACCTCCCAAGATTCGCTTC	(PET)TTCGCCATTGTTGTGATTTG	55° C	114-138

3.2.3. Analysis of the microsatellite data

Nine polymorphic nuclear loci were used to assess population structure among the 14 populations using a variety of methods. Population statistics and diversity indices were calculated for each population and for each species (*P. lescurii*, *P. densipila*, *P. lyrata*, *P. perforata*, *P. stonensis*). Nei's (1978) pairwise genetic distance among the populations and genetic similarities between

the individuals (Smouse and Peakall 1999) were used to build a population-level microsatellite dendrogram, and also served as the source for Principle Coordinates Analysis (PCO) and for the modal clustering program, PCO-MC (Reeves and Richards 2009). Population structure was additionally examined in a Bayesian framework using the program STRUCTURE (Pritchard et al. 2000).

GenAlEx v6.5 (Peakall and Smouse 2006; Peakall and Smouse 2012) was used to estimate population statistics and diversity indices. Mean observed number of alleles across loci (n_a), the effective number of alleles (n_e), the average observed heterozygosity (H_O) and the average expected heterozygosity (H_e) across populations, the total number of private alleles (Pr_a) and the proportion of polymorphic loci (p) were calculated for each population and for each species. Weir and Cockerham's (1984) inbreeding coefficient, f, which estimates the average departure of genotype frequencies from Hardy-Weinberg expectations within populations, was estimated using the program FSTAT (Goudet 1995) and assessed for significance with 120000 randomizations. In addition, GenAlEx v6.5 was used to conduct Mantel tests (999 permutations, significance level $p<0.01$) to test for isolation by distance among the species and all populations by comparing the genetic distance with geographic distance along waterways. Waterway distance was measured because gene flow is most likely occurring along rivers between populations, either by seed dispersal or by pollinators traveling along waterways. To calculate distances along the waterways, Fig. 3.1 was imported into imaging processing software ImageJ (http://rsbweb.nih.gov/ij/; Abràmoff et al. 2004). The scale was calibrated and set to kilometers and the distances between populations were measured by tracing along the waterways. This was done three times and averaged for each distance measured. The online version of GENEPOP (http://genepop.curtin.edu.au/; Raymond and Rousset 1995; Rousset 2008) was used to test for linkage disequilibrium across all loci. The Markov chain algorithm of

Raymond and Rousset (1995) was employed using default parameters to conduct probability tests across all pairs of loci.

A hierarchical visualization of population relationships was achieved by constructing a population-level dendrogram using pairwise differences between populations. The initial genetic distance matrix, calculated with Nei's (1978) standard genetic distance, was computed in GenAlEx v6.5, followed by calculation of bootstrap support values for the population tree, using Seqboot and Gendist in Phylip package v.3.6 (Felsenstein 2005) and 1000 bootstrapped data sets. The program, Neighbor (Felsenstein 2005), was used to estimate the trees and a majority-rule tree was derived using Consense (Felsenstein 2005) and edited in FigTree v1.4 (http://tree.bio.ed.ac.uk/software/figtree/). Phylogenetic analyses (Chapter 2) showed *P. auriculata* as an outgroup to the southeastern group of species, therefore one *P. auriculata* population was also included. Bootstrap values from 50-74% were considered as providing weak to moderate support, and over 75% as strong support.

To visualize the patterns of genetic variation shared by the individuals and the populations, the data were additionally subjected to a principal coordinates analysis (PCO) using population- and individual-level genetic distance matrices. PCO is a method that is used to reduce the dimensionality of multivariate datasets, condensing the variation observed in the dataset to a smaller number of orthogonal (uncorrelated) axes. This method is useful in identifying groups that share similar patterns of variation. Pairwise genetic distances between the individuals and Nei's genetic distance (Nei 1978) among the 14 populations were computed in Genalex v6.5. The PCO was also conducted in GenAlEx v6.5. The population-level matrix and a shared band similarity matrix (Lynch 1990) for the individuals that was computed in NTSYSpc v2.1 (Rohlf 2000) were used as input for the modal clustering program, PCO-MC (Reeves and Richards 2009). The shared band similarity matrix is essentially computed using the average fraction of shared allele fragments between individuals. PCO-MC is a statistically rigorous method that can

identify significant clusters of data subjected to ordination (Reeves and Richards 2009). It has been found to be more effective at identifying cryptic substructure that would otherwise be overlooked using model-based approaches (Reeves and Richards 2009). PCO-MC utilizes kernel density estimation to generate a multidimensional density landscape from which the principal coordinates have been sampled. PCO-MC can simultaneously analyze data from all principal coordinate axes to determine the number of subpopulations and the individuals that have membership within them. Each group is then assigned a "stability" value, which is the percentage measure of informative R-space that the group occupies, where informative R-space is the subset of density landscapes that yield clusters more than one and less than the number of points. A stability value greater than 15% is considered significant (Reeves and Richards 2009).

Population structure among all populations and within species was additionally investigated in Structure v2.3.4 (Pritchard et al. 2000; Falush et al. 2007). This program uses a Bayesian approach to identify and assign clusters (*K*) of related individuals in a data set (Pritchard et al. 2000; Falush et al. 2003, 2007). To find the optimal *K* value, 10 iterations of 1 to 20 *K* population clusters were run with an initial burn-in of 10,000 replicates and a Markov chain Monte Carlo (MCMC) run length of 1×10^6 generations. The ancestry model used excluded any *a priori* assumptions of populations and utilized the admixture model default parameters. Allele frequencies were assumed to be independent and the admixture alpha value was inferred from the data. The optimal *K* value was selected using the Delta K method as described by Evanno et al. (2005) and implemented in Structure Harvester (Earl 2012); http://taylor0.biology.ucla.edu/structureHarvester/). The results from the replicate runs were aligned in CLUMPP v1.1.2 (Jakobsson and Rosenberg 2007) and then Distruct v1.1 (Rosenberg 2007) was used to obtain an optimal visualization of the cluster groups.

3.3 RESULTS

3.3.1. Microsatellite variation

All nine primer sets amplified products across all species. After applying sequential Bonferroni corrections (Rice 1989), all loci were found to be unlinked. The population-level estimates of H_e (Table 3.3) vary between 0.470-0.528 in *P. lescurii,* 0.396-0.472 in *P. densipila,* 0.378-0.489 in *P. lyrata*, 0.482-0.537 in *P. perforata,* and 0.419-0.446 in *P. stonensis.* Mean per-population estimates of the inbreeding coefficient *f* for each taxa were *f* = 0.193 (*P. lescurii*), *f* = 0.170 (*P. densipila*), *f* = 0.051 (*P. lyrata*), *f* = 0.204 (*P. perforata*) and *f* = 0.097 (*P. stonensis*). The hybrid population, *P. densipila* x *P. lescurii,* had an inbreeding coefficient of 0.167. *Paysonia perforata* and *P. lescurii* have the highest level of inbreeding relative to the other taxa. Only two populations, one *P. lescurii* and one *P. perforata,* have a significant deficit of heterozygotes after applying Bonferroni corrections (Rice 1989), where the adjusted nominal alpha-value (5%) was 0.0004.

All populations of *P. lyrata, P. lescurii,* and *P. perforata* harbored private alleles, with the highest mean in *P. lescurii* of 2.67. The western-most population of *P. lescurii* contained five private alleles, the highest of all populations; this population was also the furthest downstream on the Cumberland River that was sampled in this study. In *P. densipila*, however, only two of the four populations had private alleles and only one of the two *P. stonensis* populations had private alleles. The two *P. densipila* populations that had private alleles were from the two different drainage systems that were sampled, the Harpeth River and the Duck River. The *P. densipila* x *P. lescurii* population had two private alleles.

The mean proportion of polymorphic loci ranged from 0.86-1.00 across all taxa. The highest proportion of polymorphic loci was found in *P. perforata* where all populations had 100%. The

P. densipila x *P. lescurii* hybrid population also had 100% polymorphic loci. The lowest average percentage of polymorphic loci is in *P. densipila* (86%).

Table 3.3: Population genetic statistics for all five species and the putative hybrid population derived from nine microsatellite loci. *N*, total number of individuals sampled; *n*, mean number of individuals sampled per locus; n_a, mean number of alleles per locus; *f*, Weir and Cockerham's inbreeding coefficient F_{IS}; *p*, proportion of polymorphic loci; Pr_a, total number of private alleles; *H*o, observed heterozygosity; and H_e, expected heterozygosity. *Significant deficit of heterozygotes, based on 126,000 randomizations where the indicative adjusted nominal level for multiple comparisons to maintain a table-wide error rate of 5% was alpha = 0.0004.

Population	*N*	*n*	n_a	n_e	*f*	*p*	Pr_a	H_o	H_e
Paysonia lescurii									
L1	12	11.333	3.778	2.332	0.200	1.00	5	0.397	0.470
L2	12	11.000	4.222	3.022	0.079	0.78	2	0.511	0.528
L3	11	11.000	3.778	2.418	0.299*	0.89	1	0.384	0.515
Mean	**11.67**	**11.11**	**3.93**	**2.59**	**0.193**	**0.89**	**2.67**	**0.431**	**0.505**
SE	**0.33**	**0.11**	**0.15**	**0.22**	**0.064**	**0.06**	**1.20**	**0.040**	**0.018**
Paysonia densipila									
D1	12	11.444	2.889	1.920	0.187	0.78	3	0.351	0.409
D2	12	11.556	3.667	2.362	0.220	0.89	0	0.388	0.472
D3	12	10.889	2.778	2.013	0.141	0.89	1	0.383	0.421
D4	11	10.778	2.556	1.869	0.131	0.89	0	0.364	0.396
Mean	**11.75**	**11.17**	**2.97**	**2.04**	**0.170**	**0.86**	**1.00**	**0.371**	**0.425**
SE	**0.25**	**0.20**	**0.24**	**0.11**	**0.021**	**0.03**	**0.71**	**0.009**	**0.017**
Paysonia lyrata									
LY1	13	12.333	3.222	2.179	0.224	0.89	3	0.397	0.486
LY2	14	12.111	2.444	1.775	-0.122	0.89	1	0.441	0.378
Mean	**13.50**	**12.22**	**2.83**	**1.98**	**0.051**	**0.89**	**2.00**	**0.419**	**0.432**
SE	**0.50**	**0.11**	**0.39**	**0.20**	**0.173**	**0.00**	**1.00**	**0.022**	**0.054**
Paysonia perforata									
P1	12	11.222	3.556	2.491	0.280*	1.00	3	0.410	0.537
P2	14	12.667	3.778	2.210	0.127	1.00	2	0.441	0.482
Mean	**13.00**	**11.94**	**3.67**	**2.35**	**0.204**	**1.00**	**2.50**	**0.426**	**0.509**
SE	**1.00**	**0.72**	**0.11**	**0.14**	**0.076**	**0.00**	**0.50**	**0.016**	**0.028**
Paysonia stonensis									
S1	19	17.556	3.111	2.015	0.191	1.00	3	0.374	0.446
S2	5	4.778	2.556	1.964	0.002	0.89	0	0.467	0.419
Mean	**12.00**	**11.17**	**2.83**	**1.99**	**0.097**	**0.95**	**1.50**	**0.420**	**0.433**
SE	**7**	**6.39**	**0.28**	**0.03**	**0.095**	**0.06**	**1.50**	**0.047**	**0.014**
P. densipla x P. lescurii									
PDL	**12**	**10.889**	**4.222**	**2.415**	**0.167**	**1.00**	**2**	**0.424**	**0.481**

3.3.2. Population Structure

The optimal number of clusters identified in the STRUCTURE analysis using the Evanno et al. (2005) method was K=5. The clusters corresponded to a group each of *P. lyrata*, *P. perforata*, and *P. stonensis*, and two separate clusters corresponding to *P. densipila* (Fig. 3.2). *Paysonia lescurii* populations and the putative *P. densipila x P. lescurii* population did not segregate into distinct clusters and displayed the highest levels of admixture.

In the STRUCTURE population classification, populations D3 and D4, the two *P. densipila* populations that are differentiated from the other two *P. densipila*, D1 and D2, are both found along the Duck River, while D1 and D2 are found in the Harpeth River watershed (Fig 3.3).

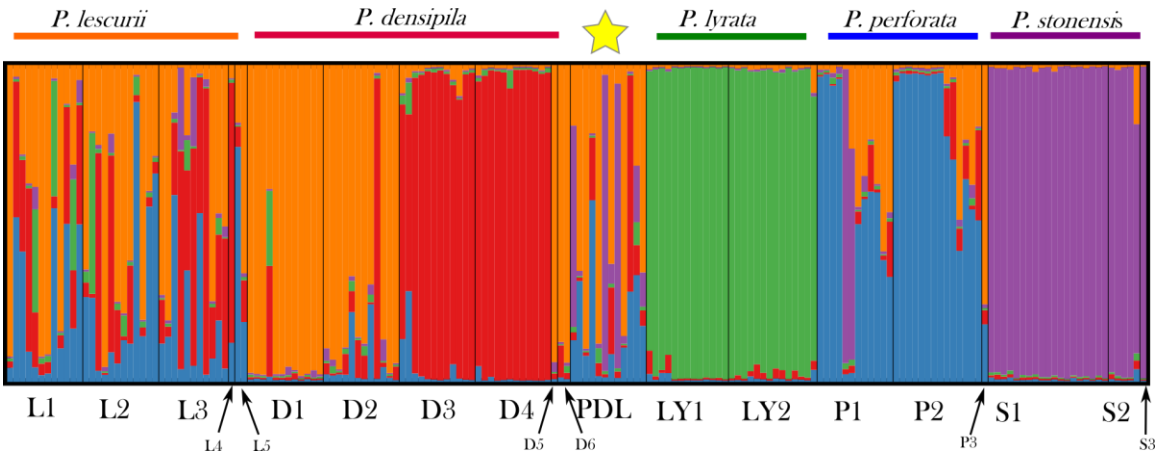


Figure 3.2: Population structure as revealed by STRUCTURE, at the optimal K value 5. Each line in the plot corresponds to an individual's genotype which is segregated into their respective populations where black lines indicate the boundaries of each group. The population identifiers are listed below the bar plot, while taxon groups are designated above the plot. The star is used to designate the *P. lescurii x P. densipila* hybrid population.

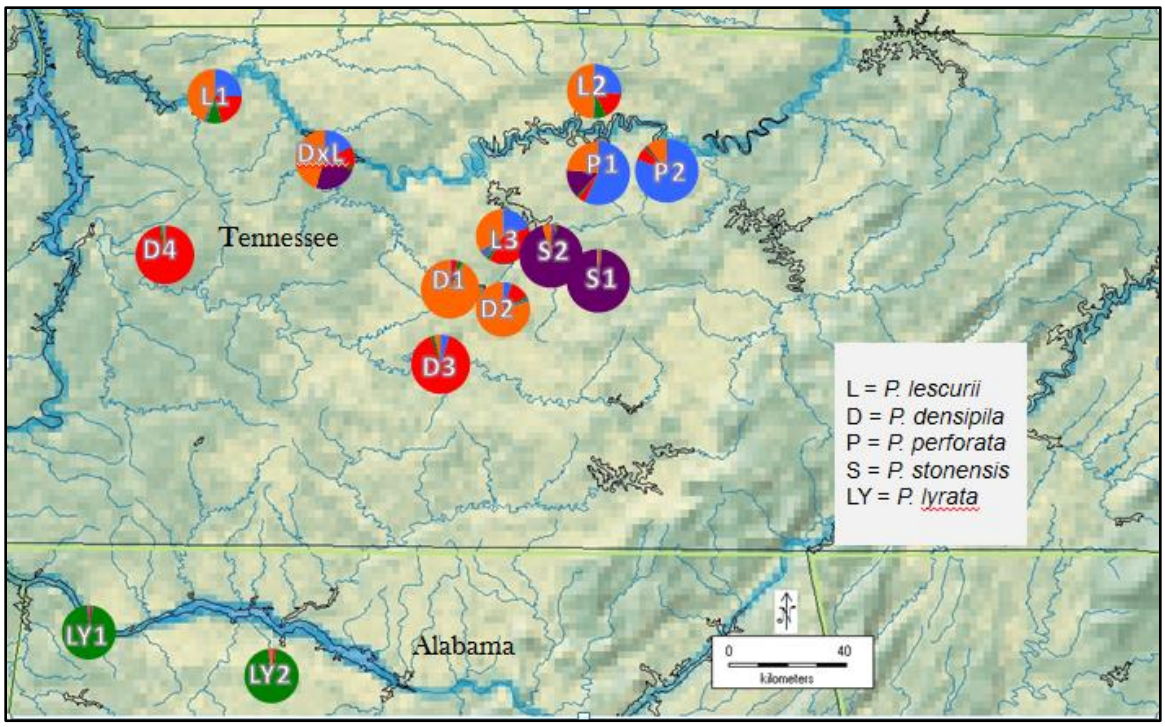


Figure 3.3: Map showing the genotypic classification of each population in relation to its geographic position. *Paysonia densipila* populations along the Duck River differ from those in the Harpeth River watershed.

3.3.3. Population-level Distance Dendrogram

The population-level dendrogram bears similarities to the assigned clusters of the STRUCTURE analysis (Fig. 3.4). The dendrogram shows the two Duck River *P. densipila* populations, D4 and D3, as distinct from the two Harpeth River *P. densipila* populations, D1 and D2. The D1/D2 cluster has a moderate support value of 67% and the D3/D4 cluster has a moderate support value of 60%. The two *P. perforata* populations, P1 and P2, are only weakly supported in a cluster of 48% bootstrap support. However, *P. lyrata* and *P. stonensis* populations form strongly supported clusters with 80% and 100% bootstrap support, respectively.

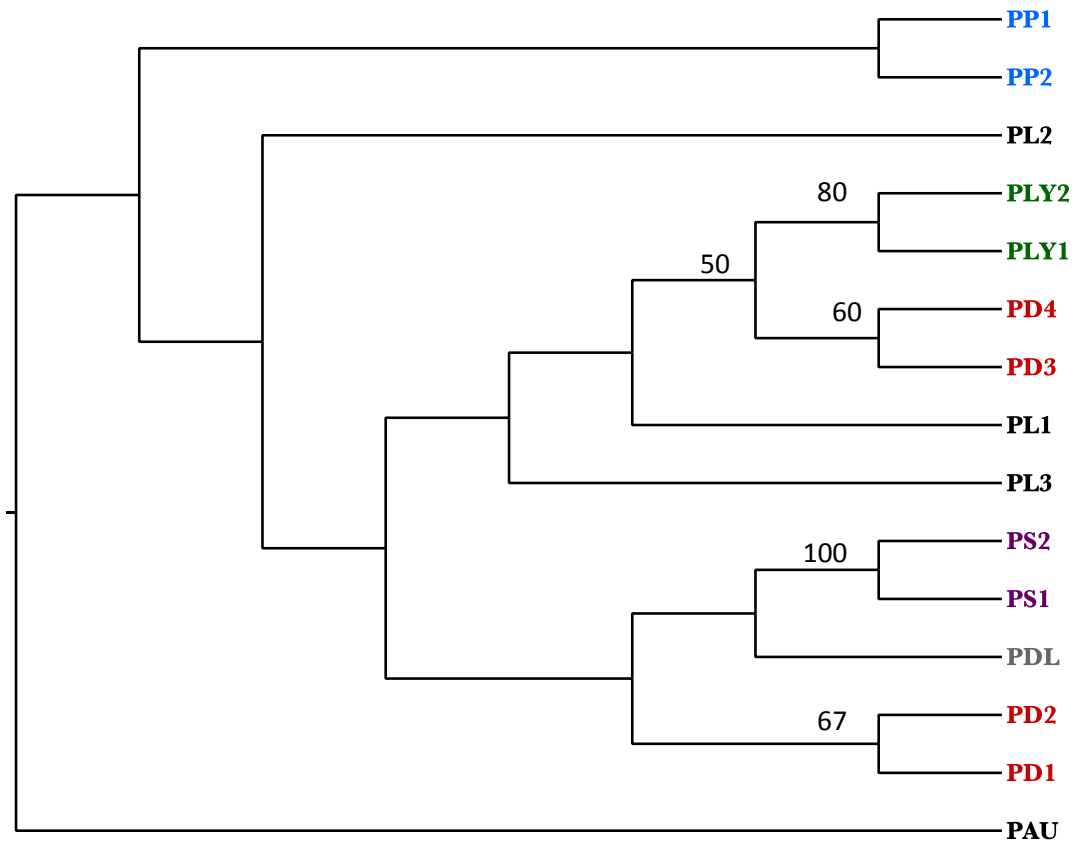


Figure 3.4: Neighbor-joining tree constructed from the Nei distance matrix of the microsatellite data. Bootstrap support values ≥50% are displayed on corresponding branches. PAU= *P. auriculata*; PD=*P. densipila*; PL=*P. lescurii*; PLY=*P. lyrata*; PP=*P. perforata*; and PS=*P. stonensis*.

3.3.4. Principal Coordinate Analysis

In the individual-level PCO, the first axis accounts for about 24% of the variation, the second axis accounts for 18%, and the third accounts for 17% of the variation (Fig.3.5). At the individual-level there is strong separation on axis 1. Coordinates 1 x 2 show that most *P. stonensis* individuals separated from the middle cluster where there are mostly *P. densipila*, *P. lescurii*, and *P. perforata* individuals together. Coordinates 1 x 2 also show that most *P. lyrata* and *P. densipila* individuals from D3 and D4 populations grouped together. The PCO-MC analyses, however, does not identify any significant clusters with over 15% stability at the individual level.

At the population-level, the first axis represents around 54% of the variation, the second around 17%, and the third accounts for 13% of the observed variation (Fig. 3.6). In Fig. 3.6.B, *P. lyrata*, *P. stonensis* and *P. perforata* populations separate from a cluster of the *P. lescurii* and *P. densipila* populations, although in Fig. 3.6.A, D3 and D4 appear somewhat separated while the *P. perforata* populations do not. The PCO-MC analysis detects two distinct, statistically significant clusters with over 15% stability (Fig. 3.6.A), with stability values of 43% (green cluster) and 28% (blue cluster). The green cluster is comprised of all *P. lescurii*, *P. perforata,* and the two Harpeth watershed *P. densipila* populations. The blue group includes all the populations except for the *P. stonensis* populations.

3.3.5. Isolation by distance

Taking into account all of the populations of all the species, the null hypothesis, that there is no relationship between geography and genetic distance, was rejected (Table 3.4). Geographic distance also appears to be correlated with genetic distance within *P. densipila, P. lyrata,* and *P. perforata* populations, but not within *P. lescurii* and *P. stonensis.*

Table 3.4: Mantel test of isolation by distance (correlation of genetic distance with geographic distance). Double asterisks indicate significance at p<0.01.

	R^2	P-value
All populations	0.0589	0.001**
P. densipila populations	0.1776	0.010**
P. lescurii populations	0.0090	0.018
P. lyrata populations	0.1456	0.001**
P. perforata populations	0.0294	0.002**
P. stonensis populations	0.0023	0.345

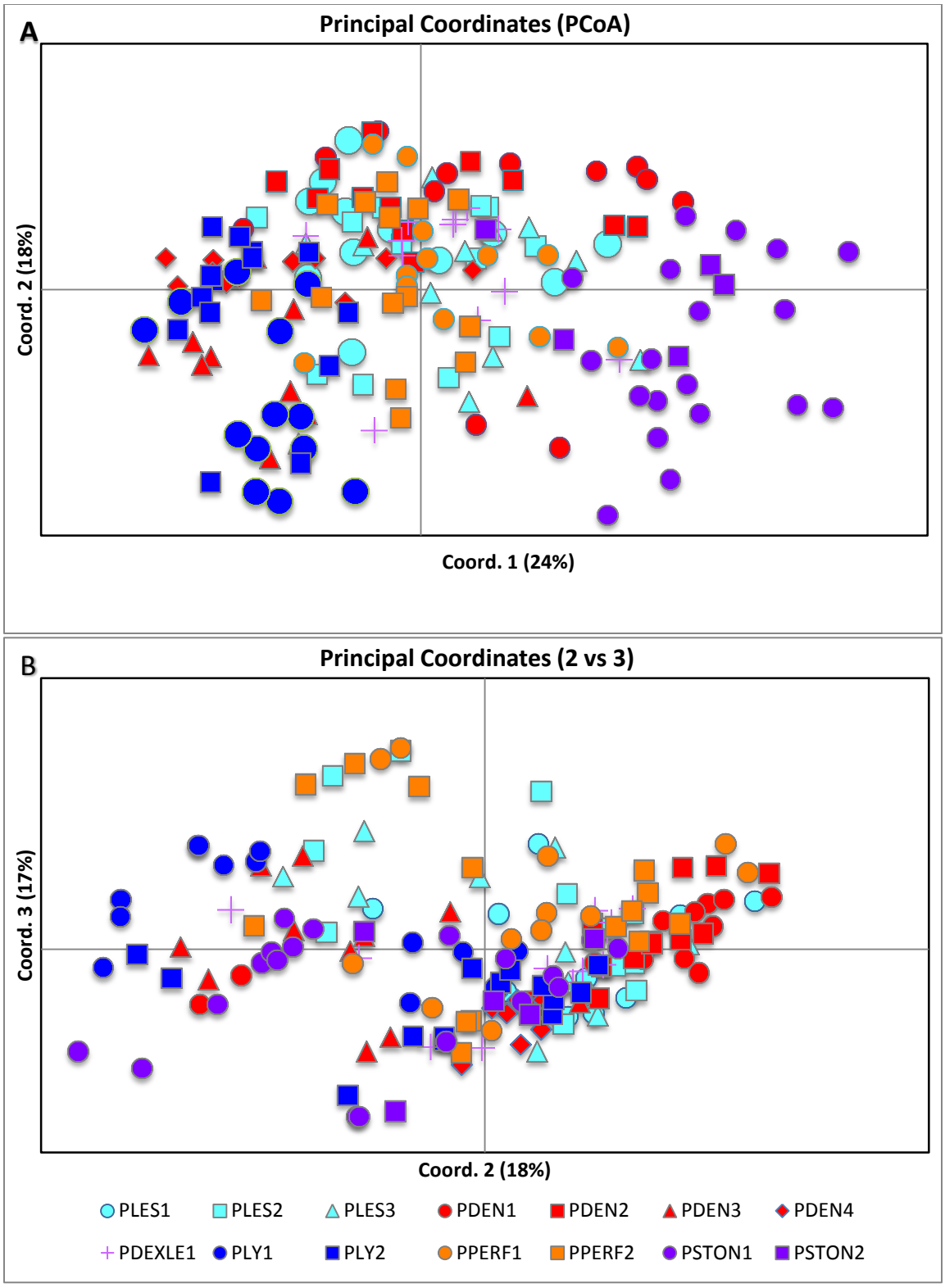


Figure 3.5: PCO of the microsatellite dataset of 171 individuals. Colors (see legend) correspond to species designation while shape pertains to population. (A) Plot of scores on principal coordinates 1 and 2. (B) Plots of scores on principal coordinates 2 and 3.

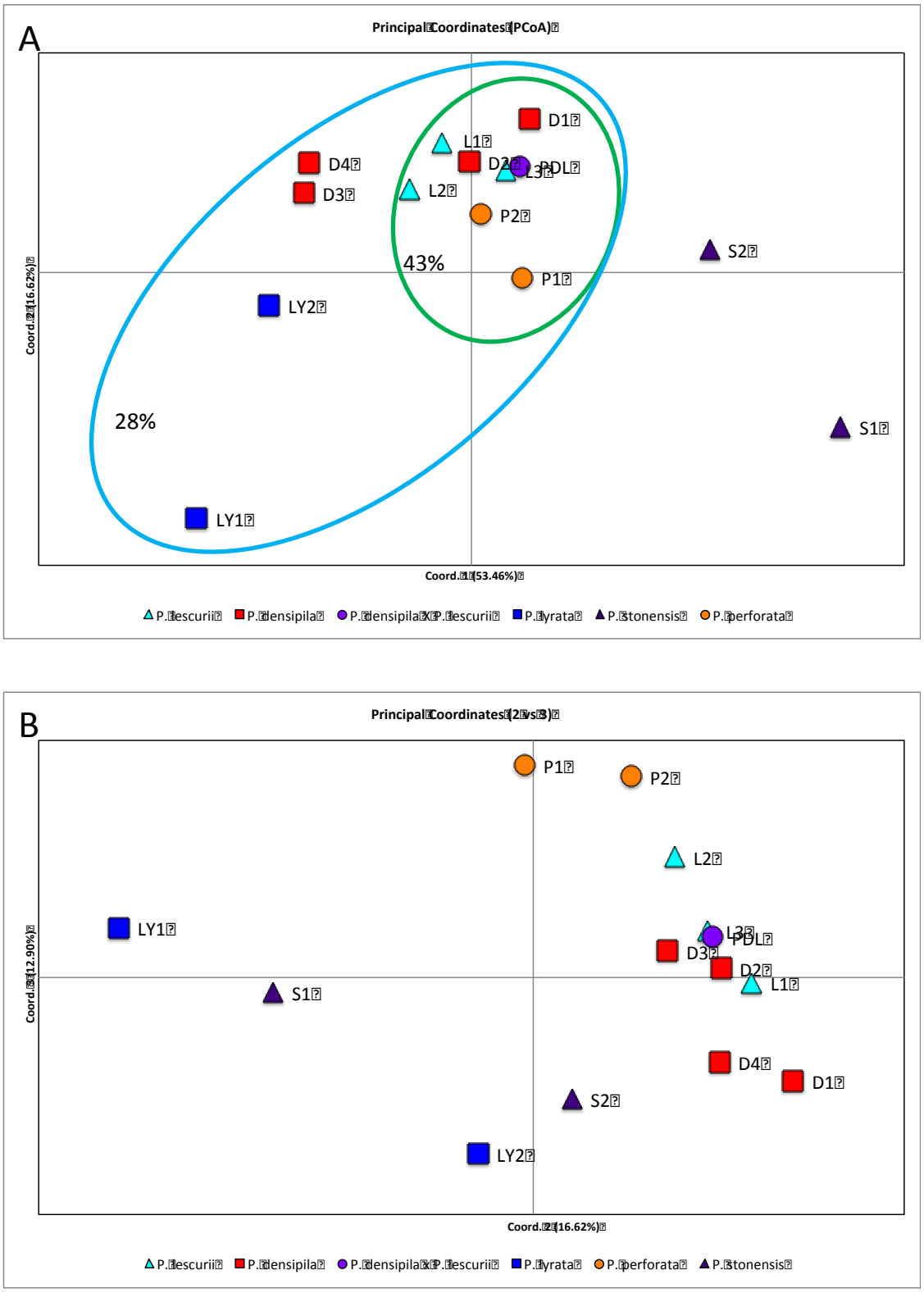


Figure 3.6: PCO of the microsatellite dataset using population genetic distances. (A) Plot of scores on principal coordinates 1 and 2. PCO-MC identified clusters outlined in blue or green. (B) Plots of scores on principal coordinate 2 and 3.

3.4 DISCUSSION

Evolutionary patterns inferred by the microsatellite analyses identify five divergent gene pools in the STRUCTURE analysis. Three of the five morphologically defined species are genetically distinct and include *P. lyrata*, *P. perforata*, and *P. stonensis*. However, *P. lescurii*, has populations that appear to contain a mixture of alleles present in other species, and *P. densipila*, contains two genetically distinct groups of populations. All taxa exhibit private alleles. The PCO-MC results are not conclusive at the individual-level, because no significant clusters were identified using the individual-distance matrix. However, Reeves and Richards (2009) explain that this program works best when using more than 100 loci whereas STRUCTURE can outperform PCO-MC for analyses using less than 10 loci. At the population-level, the PCO-MC analysis indicated two main clusters, but not at a very fine resolution. *Paysonia lescurii, P. perforata*, and the two *P. densipila* Harpeth River populations have overlapping variation at 43% stability, while the other group at 28% stability is comprised of all the populations except the two *P. stonensis* populations.

The five genetic clusters identified in STRUCTURE correspond to *P. lyrata, P. stonensis, P. perforata*, and two clusters of *P. densipila.* *Paysonia lescurii* individuals are a mixture of all 5 clusters as well as the *P. densipila* x *P. lescurii* population. The admixture could represent interspecific gene flow with other species in the morphologically defined *P. lescurii*, or it could represent shared ancestral polymorphisms with the other taxa. The microsatellite-dendrogram (Fig 3.4) shows *P. lyrata* and *P. stonensis* as separate groups, and both population and individual-level PCOs (Fig 3.5, Fig 3.6) show genetic distinctiveness. The PCO-MC results suggest that out of all the populations, the *P. stonensis* populations are the most genetically distinct since they are the only populations that do not fall into one of the two PCO-MC clusters. *Paysonia densipila* populations from the Duck River watershed, as well as the two *P. lyrata* populations in northern Alabama, appear to be distinct from the core cluster of 43% stability. This may be because these

are the most southerly populations, and are situated along different drainage systems. Grouping of the *P. perforata* populations is weakly supported in the population dendrogram, while STRUCTURE results show some level of admixture.

The STRUCTURE analysis also corroborates the *P. densipila* sub-structuring seen in the PCO (Fig. 3.6.A). Although the Duck River *P. densipila* populations appear segregated from all other populations, the two populations in the Harpeth watershed appear to be more genetically similar to *P. lescurii* in the population-level PCO than to the Duck River *P. densipila* populations. The results from the Mantel test of *P. densipila* populations suggest that geographic distance may be positively correlated with genetic distances of the individuals, which are separated along different drainage systems. The fragmentation of *P. densipila* and the genetic separation of populations along these drainage systems also point to the possibility of cryptic speciation within *P. densipila*.

Microsatellite studies can be used to ask questions about groups of species and their genetic structure. Since these markers evolve quickly, one can better judge if populations of species are distinguishable by their genetic makeup. Interspecific gene flow can also be identified and more detailed studies can estimate migration rates and patterns (Beerli and Palczewski 2010). Out of the five southeastern species, *P. lescurii* exhibits higher levels of genetic diversity (highest average of private alleles, second highest expected heterozygosity, highest number of observed alleles, and a relatively high percentage of polymorphic loci). The STRUCTURE analysis shows that *P. lescurii* contains high levels of admixture, since its individuals share genotypes found in the other taxa. The population-level dendrogram does not place the *P. lescurii* populations within a cluster; instead, they occur in several places throughout the dendrogram.

The high genetic diversity observed in *P. lescurii* may also be due to its widespread distribution and overlapping range with other species, where interspecific gene flow could be the cause for the incorporation of a variety of alleles. These populations are also downstream from other

populations and seasonal flooding could be introducing seeds from upstream populations. This species is known to form natural hybrids with *P. densipila* and *P. stonensis* (Rollins 1955; 1988). Many *P. lescurii* populations have been found to be situated closer to the hybrid populations than to the other parental species (Rollins 1957), so the possibility of introgression from hybrid populations into pure *P. lescurii* sites should not be discounted. In the STRUCTURE analysis, its individuals cannot be readily distinguished from the individuals in the sampled *P. lescurii* x *P. densipila* population.

Although *P. densipila* is widely distributed, it has the lowest mean expected heterozygosity as well as the lowest average number of private alleles when compared to the other taxa. STRUCTURE results and the microsatellite dendrogram show that its populations are distinct from populations of the other taxa, with significant sub-structuring. These sub-populations correlate to the different watershed systems on which they are found (Fig 3.3), suggesting that the fragmentation of populations along different drainage systems is influencing the genetic composition of the populations. It is unclear whether the genetic distinctiveness is due to an ancient divergence. Increasing the sampling of population sites will help to identify whether this substructure is consistently observed across all populations along the different drainage systems. The Harpeth River *P. densipila* populations in the STRUCTURE analysis shows some admixture.

Rollins (1955) believed that *Paysonia lyrata* is possibly a relic of a more continuous distribution of *Paysonia* and an intermediate evolutionary link connecting the southeastern to the southwestern species. Both populations of *P. lyrata* is strongly genetically distinct from populations of the other species, and all accessions cluster together in the population-level dendrogram, supported by the STRUCTURE and the PCO-MC results. The phylogenetic tree (Fig. 2.3; Chapter 2) based on plastid markers shows that this species is distinct as well. Homozygosity is not as high as in other species, which is surprising for an endangered species with such a restricted distribution. Seed banks, however, are known to preserve genetic diversity

and could be functioning as a reservoir for an assortment of alleles (Ellstrand and Elam 1993). These populations appear to be evolving separately from the other Central Basin species.

Paysonia perforata populations are only moderately distinct from all other populations, and the population-level dendrogram groups the two populations with a weak value of 48% bootstrap. The STRUCTURE results indicate these two *P. perforata* populations have some levels of admixture in the individuals. In the phylogenetic analysis (Fig. 2.3; Chapter 2), an unexpected result is that one of the accessions shares a similar plastid genome with *P. lyrata*. There are several possible causes for this result, such as ILS, introgression, human error, or lack of sufficiently variable markers. That particular individual was sampled and included in the STRUCTURE analysis (Labeled P3 on the lower axis in Fig. 3.2), and its genotype is a mixture of *P. perforata* and Harpeth River *P. densipila*. It is also possible that the admixture observed in the *P. perforata* STRUCTURE diagram comes from ancestral polymorphisms shared with the other taxa, and thus, that particular individual's plastid genome has retained a *P. lyrata*-like chloroplast, but its nuclear genome might share polymorphisms found in *P. densipila.* Like *P. lescurii*, however, this species harbors the second highest number of unique, private alleles, suggesting that these populations have followed or are following a distinct evolutionary path, although insufficient sampling of the other populations could explain the relatively higher number of private alleles. In addition, these populations also average the highest level of homozygosity; in particular population P1, which may be of concern for this endangered species, as its distribution is also restricted.

The two *P. stonensis* populations are strongly distinct from the populations of the other taxa. Both STRUCTURE and the population-level trees cluster these two as distinct entities, with the population-level tree grouping these together with 100% bootstrap support. Although this species is known to hybridize with *P. lescurii* and *P. densipila*, the populations sampled do not occur near hybrid populations or to populations of other species. The population sampling of this species did

not include all extant populations, and it would be of interest to sample others throughout Rutherford County, Tennessee to compare to those that are situated near the hybrid populations.

Most populations of these *Paysonia* species tend to grow along major waterways, and may have used these to expand their range. Presumably, the connection by waterways would have allowed for dispersal from the south along the Tennessee River as these seeds traveled downstream. Founder effects are known to lead to a loss of genetic diversity (Hewitt 2000), however these data do not consistently show this to be the case with downstream populations. Downstream populations for each species along the waterways have a similar genetic makeup (Fig. 3.3) and only one population of *P. lescurii* (L3) and one of *P. perforata* (P1) have a significantly higher inbreeding coefficient, but this may be due to sampling error. It appears likely that these populations are not newly established and have been around sufficiently long for evolutionary processes, such as mutation and gene flow, to account for the genetic diversity observed within the populations.

Present-day distribution of plant biodiversity in the southeastern United States has been greatly influenced by climatic changes during the past three million years (Hewitt 2000). It is surmised that current genetic structure of populations and species was shaped by the Last Glacial Maxima (LGM) (Hewitt 2000). In Tennessee, it is hypothesized that suitable temperate regions were found further south in northern Alabama, and that most of the flora present today resulted from post-glacial expansion from southern refugia (Baskin and Baskin 2003). This post-glacial expansion could have occurred relatively quickly for some species as they expanded into more suitable habitats. Although many phylogeographic studies on animals show colonization northwards after the glacial retreat, the spread from south to north for plants in the southeastern United States remains poorly understood as it has not been well-studied (Hewitt 2000).

The distinctive morphological characters of each species of *Paysonia*, however, indicate the possibility of another scenario during the LGM. Rather than *Paysonia* species surviving only in northern Alabama during the colder climate and water-logged conditions caused by the LGM and then traveling northward, it is possible that fragmentation of a more continuously distributed ancestral species occurred. Population genetic studies in other plants, such as in *Astralagus tennesseensis* and in *Dalea foliosa* (Fabaceae) (Edwards et al. 2004), suggest that some plant populations may have existed northward of 34°N during the last glacial maxima. Baskin and Baskin (2003) also note that some species with a primarily Central Basin distribution can be found in similar habitats in the northern United States and Canada. Therefore, it is possible that a few populations were able to survive in microrefugia in middle Tennessee, whereby genetic drift of these small, isolated populations or natural selection on certain traits led to the fixation of the alleles responsible for the current phenotypes observed today. The fact that the southeastern *Paysonia* species do not exhibit strong reproductive barriers suggests that their populations were historically isolated geographically. In addition, the data show that restricted populations that have had limited contact with populations of other species, such as *P. lyrata* and *P. stonensis*, have followed divergent evolutionary paths. It is also likely that southeastern *Paysonia* are in an early stage of divergence.

The biology of these plants can make these types of analyses difficult to understand, and thus the results from the analyses should be interpreted with some caution. These winter annuals require the proper type of disturbance for germination and establishment of the populations (Baskin and Baskin 1990; Fitch et al. 2007). If for some reason the conditions of a particular season are not optimal for germination at a site, the population may be nonexistent for that season. The presence of a persistent seed bank, however, can allow for the plants to establish in a different year. This is stochastic, however, and can violate expected population genetic parameters, making it difficult to trace migration routes and effective population sizes (Stacey et al. 1997). It could also mean

that the admixture observed might be due to seed banks, where ancestral polymorphisms can be retained longer in the populations and thus, are appearing as shared alleles.

Another limitation to this study is that the number of sites sampled per species is not representative of all the populations that exist for each species. Additional populations from along each of the drainage systems should be included to make better conclusions about the directionality of allele sharing that is occurring along the waterways, if it is, in fact, occurring. This is a large drawback to making broad statements of the genetic history of the southeastern *Paysonia* species because it is uncertain whether the patterns observed in two to four populations per species will be observed in all. In addition, the number of individuals sampled might not be representative of the populations. Increasing the sample size of each population would increase the probabilities that more alleles have been sampled and would lead to a more comprehensive understanding of within population dynamics (Hale et al. 2012).

Regardless of the limitations of the study, this study has offered some level of insight into understanding the genetic makeup of the southeastern *Payonia* species. Five distinct gene pools were identified that do not all correspond to each of the five species. Only one species, *P. lescurii*, shows evidence of extensive admixture, while populations of the other four southeastern *Paysonia* species do not show much evidence of widespread gene flow, although the PCO results do suggest that many populations have great overlap in genetic variation. In addition, *P. densipila* actually displays two evolutionarily distinct groups of populations. These distinct groups each occupy different drainage systems, suggesting that the effect of geography on population history is substantial and might have played a significant role in the speciation of the southeastern *Paysonia* species.

CHAPTER IV

CONCLUSION

Analyses utilizing both phylogenetics and population genetics have provided an understanding of the evolutionary history of *Paysonia*. The three southwestern species (*P. auriculata, P. grandiflora,* and *P. lasiocarpa*) are all shown to be monophyletic in the phylogenetic. Plastid and nuclear data, however, show somewhat conflicting topologies for *P. grandiflora* and *P. lasiocarpa,* where *P. lasiocarpa* is sister to the remaining species or alternatively is sister to *P. grandiflora*, where the two species together are sister to the remaining six species. *Paysonia auriculata,* in both plastid and nuclear trees, is sister to a clade composed of all the southeastern species of *Paysonia* (*P. densipila, P. lescurii, P. lyrata, P. perforata,* and *P. stonensis*).

Phylogenetic analyses suggest that the southeastern species have recently undergone, are presently undergoing speciation, or diverged long ago, but continue to exchange genes. Although most relationships among the southeastern *Paysonia* species remain unresolved, the chloroplast gene tree (Fig. 2.3; Chapter 2) provides some support for the monophyly of *P. lyrata*. The species tree (Fig. 2.5: Chapter 2) derived using coalescent methodology shows two weakly supported clades in the southeast. One is composed of *P. perforata* and *P. lyrata,* and although this is weakly supported, corroborates the gene tree analyses (Fig. 2.2 and Fig. 2.3; Chapter 2). The sister clade is composed of *P. stonensis, P. lescurii*, and *P. densipila,* which all form hybrid populations in Tennessee. Within this clade, *P. stonensis* and *P. lescurii* are moderately supported as sister taxa, while *P. densipila* is basal to these two species. The relationships shown

within this clade suggest the possibility of present-day gene flow affecting our understanding of species relationships in this clade. Geology, climatic, and anthropogenic history of the region also need to be considered in evaluating how these species have diverged.

Population genetic analyses using rapidly evolving microsatellites show allele sharing among taxa in the southeastern clade of species, but it also shows genetic distinctiveness among most of the species that was not clearly observed in the phylogenetic analysis. The exception to this observation is *P. lescurii*. However, *P. lescurii* is morphologically one of the most distinctive species of the southeastern group with its compressed fruits and combination of bulbous-based and branched trichomes. The northern Alabama species, *P. lyrata*, in both the phylogenetic and population genetic analyses, seems to be the most genetically distinct of the southeastern *Paysonia*, which is perhaps not surprising due to its geography. *Paysonia stonensis*, which is also geographically restricted, appears genetically distinct from the rest of the southeastern species. An unexpected result is that the morphologically distinct *P. densipila* is not genetically uniform across populations, with substantial sub-structure indicated in the data, correlated with the different drainage systems.

Historically, all eight species in the genus have been delimited by their morphological variation, but the variability of genetic distinctiveness for the southeastern species, at the phylogenetic and population-level, calls into question whether these southeastern species are indeed species. No one species concept applies perfectly to each species, but De Queiroz's 'unified species concept' can be used to justify why that is and why the southeastern *Paysonia* should be considered species. Species are continually evolving and the acquisitions of the properties that are used to define species happen at different stages and in a random order. Although morphologically distinct, the results from this study show that the southeastern *Paysonia* have not all acquired phylogenetic distinctiveness, while all the southwestern *Paysonia* are monophyletic, in addition to having acquired reproductive barriers. The southeastern *Paysonia* lack ample gene sequence

divergence in the chosen genes for species delimitation (with the exception of *P. lyrata*) and it is possible that other markers and methods, such as next-generation sequencing (Straub et al. 2012), would help provide a clearer phylogenetic signal by providing additional data. Considering the data from both the phylogenetic and population genetic studies for the southeastern *Paysonia, P. lyrata* is the most genetically distinct, supported by both phylogenetic and population genetic data. Population genetic data shows *P. stonensis* as strongly differentiated, while *P. perforata* is only moderately differentiated. *Paysonia lescurii*, although morphologically distinct, shows high levels of admixture and molecular data does not show it as distinct, which brings to question the age of this lineage. *Paysonia densipila* may have two lineages in the process of differentiating along two different drainage systems, according to population genetic data, possibly evidence of cryptic speciation.

Therefore, this study, which has integrated population genetics and phylogenetics, has provided better insight into understanding the evolution of *Paysonia*, especially into the evolution of its southeastern species. The different levels of studies have shown that each species is separately evolving, although some lineages are at different stages of evolution. Nonetheless, this project provides a firm foundation on which to continue building our knowledge of the evolution of this genus.

REFERENCES

Abràmoff, M. D., P. J. Magalhães, and S. J. Ram. 2004. Image processing with ImageJ. Biophotonics international 11:36-42.

Al-Shehbaz, I. A. 2011. Brassicaceae (Mustard Family). eLS. John Wiley & Sons, Ltd.

Al-Shehbaz, I. A., M. A. Beilstein, and E. A. Kellogg. 2006. Systematics and phylogeny of the Brassicaceae (Cruciferae): an overview. Plant Syst. Evol. 259:89-120.

Al-Shehbaz, I. A., and S. L. O'Kane, Jr. 2002. Lesquerella Is United with Physaria (Brassicaceae). Novon 12:319-329.

Álvarez, I., and J. F. Wendel. 2003. Ribosomal ITS sequences and plant phylogenetic inference. Mol. Phylogenet. Evol. 29:417-434.

Avise, J. C. 1994. Molecular markers: natural history and evolution. Springer.

Bailey, C. D., M. A. Koch, M. Mayer, K. Mummenhoff, S. L. O'Kane, S. I. Warwick, M. D. Windham, and I. A. Al-Shehbaz. 2006. Toward a Global Phylogeny of the Brassicaceae. Molecular Biology and Evolution 23:2142-2160.

Baskauf, C. J. 2002. The Population Genetics of Two Endemic Species of Lesquerella (Brassicaceae): L. perforata (Spring Creek Bladderpod) and L. stonensis (Stones River Bladderpod). Tennesse Department of Environment and Conservation, Division of Natural Heritage, Nashville.

Baskin, J. M., and C. C. Baskin. 1990. Seed Germination Biology of the Narrowly Endemic Species Lesquerella stonensis (Brassicaceae). Plant Species Biology 5:205-213.

Baskin, J. M., and C. C. Baskin. 2003. The Vascular Flora of Cedar Glades of the Southeastern United States and Its Phytogeographical Relationships. Journal of the Torrey Botanical Society 130:101-118.

Baskin, J. M., C. C. Baskin, and P. J. Lawless. 2007. Calcareous rock outcrop vegetation of eastern North America (Exclusive of the Nashville Basin), with particular reference to use of the term "cedar glades". Bot. Rev. 73:303-325.

Baskin, J. M. a. C. C. B. 1992. Seed Dormancy Pattern and Seed Reserves as Adaptations of the Endemic Winter Annual Lesquerella lescurii (Brassicaceae) to its Floodplain Habitat. Natural Areas Journala 12:184-190.

Baskin, J. M. a. C. C. B. 2000. Seed germination ecology of Lesquerella lyrata Rollins (Brassicaceae) a federally threatened winter annual. Natural Areas Journala 20:159-165.

Beck, J. B., I. A. Al-Shehbaz, and B. A. Schaal. 2006. Leavenworthia (Brassicaceae) revisited: Testing classic systematic and mating system hypotheses. Systematic Botany 31:151-159.

Beerli, P., and M. Palczewski. 2010. Unified Framework to Evaluate Panmixia and Migration Direction Among Multiple Sampling Locations. Genetics 185:313-326

Beilstein, M. A., I. A. Al-Shehbaz, S. Mathews, and E. A. Kellogg. 2008. Brassicaceae phylogeny inferred from Phytochrome A and ndhF sequence data: tribes and trichomes revisted. Am. J. Bot. 95:1307-1327.

Blanco-Pastor, J. L., P. Vargas, and B. E. Pfeil. 2012. Coalescent simulations reveal hybridization and incomplete lineage sorting in Mediterranean Linaria. PLoS ONE 7:e39089.

Bremer, B., K. Bremer, N. Heidari, P. Erixon, R. G. Olmstead, A. A. Anderberg, M. Källersjö, and E. Barkhordarian. 2002. Phylogenetics of asterids based on 3 coding and 3 non-coding chloroplast DNA markers and the utility of non-coding DNA at higher taxonomic levels. Mol. Phylogenet. Evol. 24:274-301.

Buckley, T. R., M. Cordeiro, D. C. Marshall, and C. Simon. 2006. Differentiating between hypotheses of lineage sorting and introgression in New Zealand alpine cicadas (Maoricicada Dugdale). Syst. Biol. 55:411-425.

Camargo, A., L. J. Avila, M. Morando, and J. W. Sites. 2012. Accuracy and precision of species trees: effects of locus, individual, and base pair sampling on inference of species trees in lizards of the Liolaemus darwinii group (Squamata, Liolaemidae). Syst. Biol. 61:272-288.

Carling, M. D., and R. T. Brumfield. 2008. Integrating Phylogenetic and Population Genetic Analyses of Multiple Loci to Test Species Divergence Hypotheses in Passerina Buntings. Genetics 178:363-377.

Carstens, B. C., and L. L. Knowles. 2007. Estimating species phylogeny from gene-tree probabilities despite incomplete lineage sorting: an example from Melanoplus grasshoppers. Syst. Biol. 56:400-411.

Chung, Y., and C. Ané. 2011. Comparing Two Bayesian Methods for Gene Tree/Species Tree Reconstruction: Simulations with Incomplete Lineage Sorting and Horizontal Gene Transfer. Syst. Biol. 60:261-275.

Conner, J. K., and D. L. Hartl. 2004. A Primer of Ecological Genetics. Sinauer Associates, Sunderland, MA

Couvreur, T. L. P., A. Franzke, I. A. Al-Shehbaz, F. T. Bakker, M. A. Koch, and K. Mummenhoff. 2010. Molecular Phylogenetics, Temporal Diversification, and Principles of Evolution in the Mustard Family (Brassicaceae). Molecular Biology and Evolution 27:55-71.

De Queiroz, A., and J. Gatesy. 2007. The supermatrix approach to systematics. Trends in Ecology & Evolution 22:34-41.

De Queiroz, K. 2007. Species Concepts and Species Delimitation. Syst. Biol. 56:879-886.

De Villiers, M. J., M. D. Pirie, M. Hughes, M. Moeller, T. J. Edwards, and D. U. Bellstedt. 2013. An approach to identify putative hybrids in the 'coalescent stochasticity zone', as exemplified in the African plant genus Streptocarpus (Gesneriaceae). New Phytologist 198:284-300.

Degnan, J. H., and N. A. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution 24:332-340.

Delcourt, H. R., P. A. Delcourt, G. Wilkins, and E. N. Smith Jr. 1986. Vegetational history of the cedar glades regions of Tennessee, Kentucky, and Missouri during the past 30,000 years. Association of Southeastern Biologists Bulletin 33:128-137.

Dobzhansky, T. G. 1951. Genetics and the Origin of Species. Columbia University Press, New York.

Drummond, A., and A. Rambaut. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. BMC Evolutionary Biology 7:214.

Earl, D. A. a. v., Bridgett M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359-361

Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32:1792-1797.

Edwards, A. L., B. Wiltshire, and D. L. Nickrent. 2004. Genetic diversity in Astragalus tennesseensis and the federal endangered Dalea foliosa (Fabaceae). Journal of the Torrey Botanical Society:279-291.

Edwards, S. V. 2009. Is a new and general theory of molecular systematics emerging? Evolution 63:1-19.

Edwards, S. V., L. Liu, and D. K. Pearl. 2007. High-resolution species trees without concatenation. Proceedings of the National Academy of Sciences 104:5936-5941.

Efron, B. 1979. Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics 7:1-26.

Ellstrand, N. C., and D. R. Elam. 1993. Population genetic consequences of small population size: implications for plant conservation. Annual review of ecology and systematics:217-242.

Estill, J. C., and M. B. Cruzan. 2001. Phytogeography of rare plant species endemic to the southeastern United States. Castanea:3-23.

Evanno, G., S. Regnaut, and J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620.

Falush, D., M. Stephens, and J. K. Pritchard. 2003. Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. Genetics 164:1567-1587.

Falush, D., M. Stephens, and J. K. Pritchard. 2007. Inference of population structure using multilocus genotype data: dominant markers and null alleles. Molecular Ecology:895-908.

Felsenstein, J. 1985. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. Evolution 39:783-791.

Felsenstein, J. 2004. Inferring phytogenies. Sunderland, Massachusetts: Sinauer Associates 4.

Felsenstein, J. 2005. PHYLIP (phylogeny inference package) version 3.6, 2004. Distributed by the author.

Ferreira, M. A. R., and M. A. Suchard. 2008. Bayesian analysis of elapsed times in continuous-time Markov chains. Canadian Journal of Statistics 36:355-368.

Fitch, E. A., J. L. Walck, and S. N. Hidayati. 2007. Agroecosystem management for rare species of Paysonia (Brassicaceae): Integrating their seed ecology and life cycle with cropping regimens in a changing climate. Am. J. Bot. 94:102-110.

Freeman, S., and J. C. Herron. 2001. Evolutionary Analysis. Prentice-Hall, Inc., New Jersey.

Fuentes-Soriano, S., and I. Al-Shehbaz. 2013. Phylogenetic Relationships of Mustards with Multiaperturate Pollen (Physarieae, Brassicaceae) Based on the Plastid ndhF Gene: Implications for Morphological Diversification. Systematic Botany 38:178-191.

Gadagkar, S. R., M. S. Rosenberg, and S. Kumar. 2005. Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. Journal Of Experimental Zoology. Part B, Molecular And Developmental Evolution 304:64-74.

Gerard, D., H. L. Gibbs, and L. Kubatko. 2011. Estimating hybridization in the presence of coalescence using phylogenetic intraspecific sampling. BMC Evolutionary Biology 11:291.

Goudet, J. 1995. FSTAT (Version 1.2): A Computer Program to Calculate F-Statistics. Journal of Heredity 86:485-486.

Granados Mendoza, C., S. Wanke, K. Salomo, P. Goetghebeur, and M. S. Samain. 2013. Application of the phylogenetic informativeness method to chloroplast markers: A test case of closely related species in tribe Hydrangeeae (Hydrangeaceae). Mol. Phylogenet. Evol. 66:233-242.

Gray, A. Manual of the botany of the northern United States, including the district east of the Mississippi and north of North Carolina and Tennessee, arranged according to the natural system, 5 ed. 1867, Ivison, Phinney, Blakemann & Co.: New York. p. 72.

Guindon, S., J. F. Dufayard, V. Lefort, M. Anisimova, W. Hordijk, and O. Gascuel. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst. Biol. 59:307-321.

Guindon, S., and O. Gascuel. 2003. A Simple, Fast, and Accurate Algorithm to Estimate Large Phylogenies by Maximum Likelihood. Syst. Biol. 52:696.

Gurushidze, M., R. M. Fritsch, and F. R. Blattner. 2010. Species-level phylogeny of Allium subgenus Melanocrommyum: Incomplete lineage sorting, hybridization and trnF gene duplication. Taxon 59:829-840.

Hale, M. L., T. M. Burg, and T. E. Steeves. 2012. Sampling for Microsatellite-Based Population Genetic Studies: 25 to 30 Individuals per Population Is Enough to Accurately Estimate Allele Frequencies. PLoS ONE 7:e45170.

Harrington, R. C., and T. J. Near. 2012. Phylogenetic and Coalescent Strategies of Species Delimitation in Snubnose Darters (Percidae: Etheostoma). Syst. Biol. 61:63-79.

Hedtke, S. M., and D. M. Hillis. 2011. The Potential Role of Androgenesis in Cytoplasmic-Nuclear Phylogenetic Discordance. Syst. Biol. 60:87-U137.

Heled, J., and A. J. Drummond. 2010. Bayesian Inference of Species Trees from Multilocus Data. Molecular Biology & Evolution 27:570.

Hewitt, G. 2000. The genetic legacy of the Quarternary ice ages. Nature 405:907-913.

Hudson, R. R. 1990. Gene genealogies and the coalescent process. Oxford surveys in evolutionary biology 7:44.

Huelsenbeck, J., and F. Ronquist. 2001. MrBayes: Bayesian inference of phylogenetic trees. Bioinformatics 17:754 - 755.

Hughes, M., M. MÖLler, D. U. Bellstedt, T. J. Edwards, and M. De Villiers. 2005. Refugia, dispersal and divergence in a forest archipelago: a study of Streptocarpus in eastern South Africa. Molecular Ecology 14:4415-4426.

Jakobsson, M., and 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.

Joly, S., P. A. McLenachan, and P. J. Lockhart. 2009. A statistical approach for distinguishing hybridization and incomplete lineage sorting. The American Naturalist 174:E54-E70.

Kingman, J. 1982. The coalescent. Stochastic Processes and Their Applications 13:235 - 248.

Kubatko, L. S. 2009. Identifying Hybridization Events in the Presence of Coalescence via Model Selection. Syst. Biol. 58:478-488.

Kubatko, L. S., B. C. Carstens, and L. L. Knowles. 2009. STEM: species tree estimation using maximum likelihood for gene trees under coalescence. Bioinformatics 25:971-973.

Kubatko, L. S., and J. H. Degnan. 2007. Inconsistency of Phylogenetic Estimates from Concatenated Data under Coalescence. Syst. Biol. 56:17-24.

Liu, L. 2008. BEST: Bayesian estimation of species trees under the coalescent model. Bioinformatics 24:2542-2543.

Liu, L., L. Yu, L. Kubatko, D. K. Pearl, and S. V. Edwards. 2009. Coalescent methods for estimating phylogenetic trees. Mol. Phylogenet. Evol. 53:320-328.

Lynch, M. 1990. The similarity index and DNA fingerprinting. Molecular Biology and Evolution 7:478-484.

Maddison, D. R., and W. P. Maddison. 2005. MacClade 4: Analysis of phylogeny and character evolution. Sinauer Associates, Sunderland, MA.

Maddison, W. P. 1997. Gene trees in species trees. Syst. Biol. 46:523-536.

Maddison, W. P., and L. L. Knowles. 2006. Inferring phylogeny despite incomplete lineage sorting. Syst. Biol. 55:21-30.

Maureira-Butler, I. J., B. E. Pfeil, A. Muangprom, T. C. Osborn, and J. J. Doyle. 2008. The reticulate history of Medicago (Fabaceae). Syst. Biol. 57:466-482.

Mayden, R. L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. Systematics Association Special Volume 54:381-424.

Mayr, E. 1942. Systematics and the origin of species. Columbia Univer sity Press, New York.

Meng, C., and L. S. Kubatko. 2009. Detecting hybrid speciation in the presence of incomplete lineage sorting using gene tree incongruence: A model. Theoretical Population Biology 75:35-45.

Mort, M. E., D. E. Soltis, P. S. Soltis, J. Francisco-Ortega, and A. Santos-Guerra. 2002. Phylogenetics and evolution of the Macaronesian clade of Crassulaceae inferred from nuclear and chloroplast sequence data. Systematic Botany 27:271-288.

NatureServe. 2013. NatureServe Explorer: An online encyclopedia of life [web application]. NatureServe, Arlington, Virginia.

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

Neigel, J. E., and J. C. Avise. 1986. Phylogenetic relationships of mitochondrial DNA under various demographic models of speciation in E. Nevo, and S. Karlin, eds. Evolutionary Processes and Theory. Academic Press, Inc, New York.

Nylander, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.

O'Kane, S.L., Jr., 2010. *Paysonia.* In: Flora of North America Editorial Committee, e. 2010. Flora of North America North of Mexico. Vol. 7, pp. 611-614, New York and Oxford.

O'Kane, S. L., and I. A. Al-Shehbaz. 2002. Paysonia, a new genus segregated from Lesquerella (Brassicaceae). Novon 12:379-381.

Payson, E. B. 1922. A Monograph of the Genus Lesquerella.

Peakall, R., and P. E. Smouse. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research—an update. Bioinformatics 28:2537-2539.

Peakall, R., and P. E. Smouse. 2006. GenAlEx 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288-295.

Pinzón, J. H., and T. LaJeunesse. 2011. Species delimitation of common reef corals in the genus Pocillopora using nucleotide sequence phylogenies, population genetics and symbiosis ecology. Molecular Ecology 20:311-325.

Pollard, D. A., V. N. Iyer, A. M. Moses, and M. B. Eisen. 2006. Widespread discordance of gene trees with species tree in Drosophila: evidence for incomplete lineage sorting. PLoS genetics 2:e173.

Posada, D. 2008. jModelTest: Phylogenetic model averaging. Molecular Biology and Evolution 25:1253-1256.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of Population Structure Using Multilocus Genotype Data. Genetics 155:945-959.

Quarterman, E. 1950. Ecology of Cedar Glades. I. Distribution of Glade Flora in Tennessee. Bulletin of the Torrey Botanical Club 77:1-9.

Raymond, M., and F. Rousset. 1995. GENEPOP (Version 1.2): Population Genetics Software for Exact Tests and Ecumenicism. Journal of Heredity 86:248-249.

Reeves, P. A., and C. M. Richards. 2009. Accurate Inference of Subtle Population Structure (and Other Genetic Discontinuities) Using Principal Coordinates. PLoS ONE 4:e4269.

Rieseberg, L. H., S. M. Beckstrom-Sternberg, A. Liston, and D. M. Arias. 1991. Phylogenetic and systematic inferences from chloroplast DNA and isozyme variation in Helianthus sect. Helianthus (Asteraceae). Systematic Botany:50-76.

Rieseberg, L. H., and D. E. Soltis. 1991. Phylogenetic consequences of cytoplasmic gene flow in plants. Evolutionary Trends in Plants 5:65-84.

Rokas, A., B. L. Williams, N. King, and S. B. Carroll. 2003. Genome-scale approaches to resolving incongruence in molecular phylogenies. Nature 425:798-804.

Roland, M. H. 1926. The Cedar Glades of Middle Tennessee. Ecology 7:48-54.

Rollins, R. C. 1952. Some Cruciferae of the Nashville Basin, Tennessee. Rhodora v.54:182-192.

Rollins, R. C. 1954. Interspecific hybridization and its role in plant evolution. Pp. 172–180. VIII International Botanical Congress: Reports and Communications Section Paris.

Rollins, R. C. 1955. The auriculate-leaved species of Lesquerella (Cruciferae). Rhodora 57:241-264.

Rollins, R. C. 1957. Interspecific hybridization in Lesquerella (Cruciferae). Contributions from the Gray Herbarium of Harvard University 181:3-40.

Rollins, R. C. 1988. A Population of Interspecific Hybrids of Lesquerella (Cruciferae). Systematic Botany 13:60-63.

Rollins, R. C., and E. A. Shaw. 1973. The Genus Lesquerella (Cruciferae) in North America. Pp. 1-35. Harvard University Press, Cambridge, Massachusetts.

Rollins, R. C., and O. T. Solbrig. 1973. Interspecific hybridization in Lesquerella. Contributions from the Gray Herbarium of Harvard University 203:1-48.

Rosenberg, N. A. 2007. Distruct: a program for the graphical display of population structure. Molecular Ecology Notes 4:137-138.

Rosenberg, N. A., and M. Nordborg. 2002. Genealogical trees, coalescent theory and the analysis of genetic polymorphisms. Nat. Rev. Genet. 3:380-390.

Rousset, F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8:103-106.

Schlötterer, C. 2000. Evolutionary dynamics of microsatellite DNA. Chromosoma 109:365-371.

Schranz, M. E., A. J. Windsor, B.-h. Song, A. Lawton-Rauh, and T. Mitchell-Olds. 2007. Comparative Genetic Mapping in Boechera stricta, a Close Relative of Arabidopsis. Plant Physiology 144:286-298.

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

Service, U. S. F. a. W. 1996. Recovery Plan for the Lyrate Bladderpod (Lesquerella lyrata Rollins). Pp. 27, Atlanta Georgia.

Shaw, J., E. B. Lickey, J. T. Beck, S. B. Farmer, W. S. Liu, J. Miller, K. C. Siripun, C. T. Winder, E. E. Schilling, and R. L. Small. 2005. The tortoise and the hare II: Relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. Am. J. Bot. 92:142-166.

Shaw, J., E. B. Lickey, E. E. Schilling, and R. L. Small. 2007. Comparison of whole chloroplast genome sequences to choose noncoding regions for phylogenetic studies in angiosperms: the tortoise and the hare III. Am. J. Bot. 94:275-288.

Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Molecular Biology and Evolution 16:1114-1116.

Smouse, P. E., and R. Peakall. 1999. Spatial autocorrelation analysis of individual multiallele and multilocus genetic structure. Heredity 82:561-573.

Soltis, D. E., and R. K. Kuzoff. 1995. Discordance between Nuclear and Chloroplast Phylogenies in the Heuchera Group (Saxifragaceae). Evolution 49:727-742.

Soltis, D. E., and P. S. Soltis. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. Pp. 1-42. Molecular systematics of plants II. Springer.

Song, B.-H., M. J. Clauss, A. Pepper, and T. Mitchell-Olds. 2006. Geographic patterns of microsatellite variation in Boechera stricta, a close relative of Arabidopsis. Molecular Ecology 15:357-369.

Stacey, P. B., M. L. Taper, and V. A. Johnson. 1997. 12 - Migration within Metapopulations: The Impact upon Local Population Dynamics. Pp. 267-291 in H. Ilkka, and E. G. Michael, eds. Metapopulation Biology. Academic Press, San Diego.

Starnes, W. C., and D. A. Etnier. 1986. Drainage evolution and fish biogeography of the Tennessee and Cumberland Rivers drainage realm. Pp. 325-361 in C. H. Hocutt, and E. O. Wiley, eds. The zoogeography of North American freshwater fishes. John Wiley & Sons, New York.

Straub, S. C. K., M. Parks, K. Weitemier, M. Fishbein, R. C. Cronn, and A. Liston. 2012. Navigating the tip of the genomic iceberg: Next-generation sequencing for plant systematics. Am. J. Bot.

Swofford, D. L. 1998. PAUP*. Phylogenetic Analysis using Parsimony (*and other methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Tsitrone, A., M. Kirkpatrick, and D. A. Levin. 2003. A Model for Chloroplast Capture. Evolution 57:1776-1782.

U.S. Fish and Wildlife Service. 1990. Determination of Endangered Status for Lesquerella lyrata (Lyrate bladder-pod). Pp. 39864-39868 Federal Register.

U.S. Fish and Wildlife Service. 1996. Determination of Endangered Status for Lesquerella Perforata (Spring Creek Bladderpod). Federal Register 61:67493-67497.

U.S. Fish and Wildlife Service. 1999. Endangered and threatened wildlife and plants; notice of reclassification of Lesquerella stonensis (Stones River bladderpod). Federal Register 64:25216.

U.S. Fish and Wildlife Service. 2006. Recovery Plan for Lesquerella peforata (Spring Creek Bladderpod). Pp. 23, Atlanta, Georgia.

Warwick, S., K. Mummenhoff, C. Sauder, M. Koch, and I. Al-Shehbaz. 2010. Closing the gaps: phylogenetic relationships in the Brassicaceae based on DNA sequence data of nuclear ribosomal ITS region. Plant Syst. Evol. 285:209-232.

White, T. J., T. Bruns, and S. T. Lee. 1990. PCR protocols: A guide to methods and applications. Pp. 315-322 in D. H. G. M. A. Innis, and J. J. S. T. J. White, eds. Amplification and Direct Sequencing of Fungal Ribosomal RNA Genes for Phylogenetics. Academic Press, San Diego.

Wilson, C. W. 1962. Stratigraphy and Geologic History of Middle Ordovician Rocks of Central Tennessee. Geological Society of America Bulletin 73:481-504.

Wright, S. 1931. Evolution in Mendelian Populations. Genetics 16:97-159.

Yu, Y., J. H. Degnan, and L. Nakhleh. 2012. The Probability of a Gene Tree Topology within a Phylogenetic Network with Applications to Hybridization Detection. PLoS Genet 8:e1002660.

Yu, Y., C. Than, J. H. Degnan, and L. Nakhleh. 2011. Coalescent Histories on Phylogenetic Networks and Detection of Hybridization Despite Incomplete Lineage Sorting. Syst. Biol. 60:138-149.

VITA

Guadalupe Borja

Candidate for the Degree of

Master of Science

Thesis: THE EVOLUTIONARY HISTORY OF PAYSONIA (BRASSICACEAE)

Major Field: Botany

Biographical:

Education:

Completed the requirements for the Master of Science in Botany at Oklahoma State University, Stillwater, Oklahoma in July, 2013.

Completed the requirements for the Bachelor of Science in Botany at University of Oklahoma, Norman, Oklahoma in December 2008.

Experience:

Graduate Teaching Assistant, Botany
Graduate Teaching Assistant, General Biology
Research Assistant

Professional Memberships:

Botanical Society of America
American Society of Plant Taxonomists
Systematic Botany