POPULATION GENETICS, SPECIATION, AND HYBRIDIZATION IN *DICERANDRA* (LAMIACEAE), AN ENDEMIC OF THE SOUTHEASTERN USA

By

ADAM PAYTON

A THESIS PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE

UNIVERSITY OF FLORIDA

2012

To my Parents

ACKNOWLEDGMENTS

I thank my advisor Douglas Soltis for giving me the opportunity to further my education in a field I am passionate about. His guidance and support have been critical for the completion of this project. I also thank the members of my advisory committee: Pam Soltis, Matt Gitzendanner, and Walter Judd who have significantly contributed to my ability to advance this study.

I thank Robin Huck for her involvement in initially shaping the direction of this project and her lifelong dedication to the study of the diversity and evolution of this group of unique plants. I am grateful to Tom Patrick of the Georgia Division of Natural Resources and Alison McGee of the Nature Conservancy for their assistance with *Dicerandra radfordiana*. I thank the staff of the University of Florida Herbarium, particularly Kent Perkins, and the members of the U.F. Biology Department for their knowledge, encouragement, and support. Finally, I thank my family and friends who have been patient and endlessly supporting throughout my graduate career, without which, this would not have been possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

°C	Degrees Celsius
2*n*	Diploid chromosome number
AMOVA	Analysis of Molecular Variance
BLAST	Basic local alignment search tool
bp	Basepairs
ca	Circa
CHR	*Dicerandra christmanii*
cm	Centimeter
COR	*Dicerandra cornutissima*
CTAB	Hexadecyltrimethylammonium
D	Jost's *D*
D.	*Dicerandra*
DAPI	4',6-diamidino-2-phenylindole
DEN	*Dicerandra densiflora*
D_{EST}	Estimate of Jost's *D*
diH2O	Deionized water
dm	Decimeter
DNA	Deoxyribonucleic acid
DNR	Division of Natural Resources
dNTP	Deoxyribonucleotide triphosphate
EDTA	Ethylenediaminetetraacetic acid
Eff.	Effective
et al.	Et alia
F_{IS}	Inbreeding coefficient

FLAS	University of Florida Herbarium
FNAI	Florida Natural Areas Inventory
FRU	*Dicerandra frutescens*
F_{ST}	Fixation index
FUM	*Dicerandra fumella*
GC	Guanine and cytosine
GPB	General purpose buffer
H_e	Expected heterozygosity
H_o	Observed heterozygoisty
HWE	Hardy-Weinberg equilibrium
HYB	Hybrid
IMM	*Dicerandra immaculata* variety *immaculata*
ITS	Internal transcribed spacer
K	Clusters
km	Kilometer
LD	Linkage disequilibrium
LIN	*Dicerandra linearifolia* variety *linearifolia*
m	Meter
M	Molar
matK	Maturase K gene
MID	Multiplex identifier
mL	Milliliter
mm	Millimeter
mM	Millimolar
MOD	*Dicerandra modesta*

NAD83	North American datum 1983
ng	Nanogram
Num.	Number
ODO	*Dicerandra odoratissima*
PCR	Polymerase chain reaction
POP	Population
psi	Pounds per square inch
RAD	*Dicerandra radfordiana*
ROB	*Dicerandra linearifolia* variety *robustior*
s.l.	Sensu lato
SAV	*Dicerandra immaculata* variety *savannarum*
sp.	Species
Std.	Standard
THI	*Dicerandra thinicola*
trnL	Transfer ribonucleic acid L gene
trnT	Transfer ribonucleic acid T gene
USFWS	United States Fish and Wildlife Service
UTM	Universal Transverse Mercator
v.	Version
var.	Variety
µg	Microgram
µL	Microliter
µM	Micromolar

Abstract of Thesis Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Master of Science

POPULATION GENETICS, SPECIATION, AND HYBRIDIZATION IN *DICERANDRA* (LAMIACEAE), AN ENDEMIC OF THE SOUTHEASTERN USA

By

Adam Payton

August 2012

Chair: Douglas Soltis
Major: Botany

Endemic organisms are valuable biologic and intrinsic components of ecosystems and the study of endemics can address a variety of evolutionary questions. The southeastern United States harbors a high number of endemics including the eleven species of the genus *Dicerandra*. Recent phylogenetic analysis revealed accessions of the same species of *D. linearifolia* var. *linearifolia* and *D. odoratissima* from different geographic regions were not sister to each other. These phylogenetic discontinuities raised questions about the evolutionary history, relationships, and taxonomic circumscription within these species. This study used microsatellite based population level analyses of 42 populations from four of the annual species (*D. linearifolia* var. *linearifolia*, *D. linearifolia* var. *robustior*, *D. fumella*, *D. odoratissima*, and *D. radfordiana*), to address the questions of population structure, gene flow, and hybridization. Strong support was found for the species level recognition of the recently described *D. fumella* from the Florida panhandle with no evidence for hybridization with the proximate populations of *D. linearifolia* var. *robustior*. *Dicerandra linearifolia* var. *linearifolia* showed some regional cohesion of populations, seen in STRUCTURE analyses and neighbor-joining trees based on genetic distance, but there was no consistent

geographic pattern to the clustering of populations for the recognition of regional populations at the species level. *Dicerandra radfordiana* showed consistent clustering in STRUCTURE analyses and distance trees with proximate populations of *D. odoratissima*. Since *D. radfordiana* is found at the southeastern extreme of the *D. odoratissima* range these populations may represent the beginning of speciation by isolation. While there are morphological distinctions between *D. odoratissima* and *D. radfordiana* there is no molecular support for a distinct *D. radfordiana* entity. This may result from incomplete lineage sorting of alleles early in the speciation process. As a result the taxonomic circumscription of *D. radfordiana* at the species rank is in question. Overall, there is significant genetic diversity found at the populations level for all *Dicerandra* annuals, 11.39% of the total genetic diversity is found among populations within species. This value is similar to the proportion of genetic variation found between species, 12.49%. The isolated nature of many populations, limited dispersal capability, and frequent visitation from generalist pollinators produce conditions where drift can cause significant differentiation between populations, even within the same geographic area. However the obligate outcrossing nature of these plants and the potential for multi-generational populations resulting from recruitment from the seed bank are likely stabilizing factors maintaining heterozygosity within populations.

CHAPTER 1
INTRODUCTION

Study System

Endemic organisms are valuable biologic and intrinsic components of ecosystems. The study of endemics can address a variety of evolutionary questions. For example, analyses of endemics can provide insights into the recent evolutionary history of a region, as many endemics are the result of recent speciation and thus represent the leading edge of the evolutionary process. In contrast, endemics can be among the last remaining members of ancient communities and thus provide a glimpse into the past, showing what forms of life previously occupied a region.

Endemics, especially those that are rare, threatened, or endangered, receive a relatively high degree of public interest and scientific study. Such studies guide conservation efforts, which must begin with a strong understanding of the life-history of an organism. Ecology and demography are often the first aspects studied as they can reveal how an organism fits into and functions within a community. Much can be learned about the niche that a given endemic fills by utilizing these approaches. Equally as important, but typically less studied, is the pattern and role of genetic diversity and population structure. Such studies provide insights into overall genetic differentiation, the degree of migration or admixture between populations, and the evolutionary history of an organism. Understanding genetic diversity and how that diversity is distributed across a landscape provides fundamentally different insights than ecological studies, but is equally important.

The southeastern United States contains a large number of endemic plant species (Estill and Cruzan 2001). Many of these species occur in specific habitats, which are

also rare, such as scrub, sandhill, and coastal dune communities (Myers and Ewel 1990). Often studies addressing endemics focus on individual species that are members of genera/clades containing widely distributed species along with a few narrow endemics. This approach can yield insights into the underlying genetic factors such as hybridization, polyploidy, and introgression, which can be major factors driving the evolution of rare species. We have chosen to investigate a clade composed solely of southeastern endemic plants. This approach facilitates the comparison of multiple closely related species to investigate patterns of genetic diversity and structure, which in turn provides new insights into common factors or mechanisms that lead to the formation and maintenance of clades containing rare species. Understanding these speciation patterns and processes forms the foundation for understanding the origin of organismal diversity, and ultimately the biotic diversity of a region.

We chose for investigation the annual clade of the southeastern endemic mint genus *Dicerandra* Benth. (Huck 1987, Oliveria 2007). We used a microsatellite based population genetic approach to investigate genetic variation, population structure, evidence for hybridization, and degree of differentiation between members of this annual clade.

Study Organisms

Distribution

Dicerandra (Lamiaceae) is endemic to the southeastern United States. First described by Bentham (1848), *Dicerandra* is characterized by spurs on the anther sacs. From these spurs, the genus derives its name meaning “two-horned male”. This character remains a significant distinguishing synapomorphy of all species in the genus.

Dicerandra currently comprises 11 species, of which two have varietal designations (Huck 1987, 2010). There are five annual species [*D. densiflora*, *D. fumella*, *D. linearifolia* (composed of var. *linearifolia* and var. *robustior*), *D. odoratissima*, and *D. radfordiana*] and six perennial species [*D. christmanii*, *D. cornutissima*, *D. frutescens*, *D. immaculata* (composed of var. *immaculata* and var. *savannarum*), *D. modesta*, and *D. thinicola*].

The perennial species are restricted to ancient dune ridges in central peninsular Florida and inland Atlantic coastal dunes. Many perennials have very narrow ranges (Figure 1-1), small population sizes, and have experienced decrease and fragmentation of critical habitat. As a result, all of the perennials except *D. thinicola* are listed as federally endangered (USFWS 2007).

The combined range of the annual species of *Dicerandra* is considerably wider than that of the perennials. They occur from the Florida panhandle through southeastern Alabama and southern Georgia. Their northern limit is at the Fall Line of Georgia, a geologic barrier separating the hard crystalline rock of the Piedmont from the sedimentary deposits of the Coastal Plains. All annual species, with the exception of *D. radfordiana*, have wide distributions occupying hundreds of square miles; however, populations are often isolated from one another creating a sparse local distribution. *Dicerandra radfordiana* has only two known populations in McIntosh Co., Georgia. In eastern Georgia the ranges of *D. linearifolia* var. *linearifolia* and *D. odoratissima* overlap, and rare hybrids have been described from this contact zone where populations of the two species are adjacent or intermixed. In the Florida panhandle *D. fumella*, *D. linearifolia* var. *robustior*, and *D. densiflora* have parapatric distributions with

a proposed hybridization zone between *D. fumella* and *D. linearifolia* var. *robustior* in the central panhandle (Figure 1-1) (Huck 2010).

Ecology

All *Dicerandra* species are found on deep, well-draining, sandy soils, typical of the coastal plain of the southeastern U.S. (Huck 1987). The soil, quartzipsamments, consists of deep, unstructured, acidic sand deposits with little to no organic horizon and is low in available nutrients especially phosphorous (Huck 1987, Myers and Ewel 1990). Disturbance is a critical component for persistence of populations. A significant increase of individuals is often seen following perturbation of a site, mechanical or pyrogenic (Huck 1987, Menges 1992, personal observation). Annual species are almost exclusively found in disturbed areas near road cuts, power line rights-of-way, firebreaks, or trails. Sites are typically characterized by having a sparse forb and graminoid community, minimal to no overstory canopy cover, and low water availability. Plants can be found within natural openings in former sandhill or high pine communities but are more common in disturbed areas adjacent to dense, closed-canopy, later-successional variants of sandhill or high pine (xeric hammock or turkey oak barrens) communities or adjacent to pine plantations presumably occupying former sandhill or other high pine community types (Myers and Ewel 1990, FNAI 2010).

Dicerandra species are considered obligate outcrossers. Self compatibility experiments showed 1.9% of crosses resulted in mature flowering plants, while selfing studies showed 0.7% of selfed offspring reached maturity (Huck 1987).

Artificial hybridization experiments involving all *Dicerandra* species revealed a low frequency of viable interspecific offspring, approximately 12%, with the vast majority resulting from crosses between annual species (Huck 1987). In most cases the

allopatric distribution of species prohibits reproductive contact; however, several contact zones have been considered to contain natural hybrid individuals. In eastern Georgia *D. linearifolia* var. *linearifolia* has been observed to hybridize with *D. odoratissima* (Huck 1987, personal observation), and in the Florida panhandle hybrids between *D. fumella* and *D. linearifolia* var. *robustior* have been proposed (Huck 2010).

Dispersal capability of *Dicerandra* is thought to be limited. The pericarp wall is coved with hydrophobic mucilaginous cells, which provide buoyancy to seeds, suggesting that water is an important dispersal mechanism (Huck 1987). This is corroborated by many populations of *Dicerandra* being located near waterways or on sand ridges deposited by ancient waterways and may provide a mechanism to explain the overall distribution of species from a historical standpoint (Figures 2-1 & 2-2). Limited dispersal is also supported by phylogenetic data, which show a strong correlation between clades of annual species and the watershed in which they occur (Oliveira et al. 2007). Yet localized dispersal by water is likely limited to precipitation events of sufficient volume to create sheet flows extending beyond the population perimeter. While precipitation events of this scale do occur, the isolated nature of many populations and their relatively small sizes suggest long-distance seed dispersal is limited. These observations suggest that the primary mechanism for genetic exchange between populations is pollinator-mediated. However, relatively little is known about the pollinators for the annual species. Various insects have been observed systematically visiting flowers, with the majority being generalists, such as the cloudless sulphur butterfly, Gulf fritillary, and various bees (personal observation, Huck 1987), but detailed

study would be required to determine the precise composition and frequency of visits by these species and their efficiency as pollinators.

Past Research on *Dicerandra*

Taxonomic History

The first description of *Dicerandra* was a specimen of *D. linearifolia* var. *linearifolia,* from eastern Georgia, collected by Stephen Elliot in 1821. Originally *Ceranthera linearifolia*, the name was changed by Bentham (1848) to *Dicerandra*. The description of additional species continued over the next century with the genus receiving taxonomic treatment by Shinners (1962) and then Kral (1982). Huck (1987) provided the first thorough investigation of the taxonomy, phylogenetics, reproductive biology, and extent of hybridization. This work resulted in the recognition of two sections within the genus. Section *Dicerandra* contained all of the perennial species as well as *D. linearifolia* var. *linearifolia*, *D. linearifolia* var. *robustior*, and *D. densiflora*. Section *Lecontea* contained the remaining two species, *D. odoratissima* and *D. radfordiana*. These sectional designations were based primarily on a suite of floral characters and their primary type of secondary compound; either menthol-based, which produced a minty odor, or cineol-based, which produced a cinnamon like odor.

A molecular phylogenetic analysis (Oliveira et al. 2007) showed the genus was in fact composed of two sister clades; one contained the perennials, redesignated as subgenus *Kralia*, and the other the annuals, designated subgenus *Dicerandra*. Huck (2010) elevated the Florida panhandle populations that had been referred to as *D. linearifolia* var. *linearifolia* to the rank of species, naming them *D. fumella*, based on phylogenetic results in which the panhandle populations of *D. linearifolia* var. *linearifolia*

were in a separate clade from the populations found in Georgia (Oliveira et al. 2007) and morphological characteristics (Huck 2010).

Biological Research

Due to conservation concerns, the perennial species have received considerable research attention with studies investigating reproductive and pollination biology (Evans et al. 2004), production of secondary leaf compounds (McCormick 1993, Huck 1989), microhabitat preference and response to disturbance (Menges et al. 1999, Menges 1992), and genetic diversity across scrub endemics in past and present landscapes (McDonald and Hamrick 1996, Menges et al. 2001, Menges 2010). The annuals have received less attention, although *D. linearifolia* was the focus of a series of studies investigating development and plasticity of leaf variation (Winn 1996a, 1996b, 1999) and functional trait comparisons (Merchant 2006). Both perennials and annuals were included in a chromosome analysis showing 2*n*=32 (*D. cornutissima*, *D. christmanii*, *D. frutescens*, *D. linearifolia* var. *robustior*, *D. odoratissima*, *D. radfordiana*, and *D. thinicola*) and 2*n*=48 in (*D. densiflora*, *D. linearifolia* var. *linearifolia*) with the 2*n*=48 species considered to be the result of recent genome duplication events within the genus (Huck and Chambers 1997).

Motivation for this Study

The phylogenetic analysis performed by Oliveira et al. (2007) clearly separated annuals and perennials into distinct clades; however, it provided limited resolution in the annual clade especially within geographic regions (Figure 1-2). Within the annuals, three clades were recovered which correspond to major geographic regions; river basins draining to either the Atlantic Ocean or Gulf of Mexico and the highlands of west-

central Georgia. While regional geographic structure was evident, polytomies involving multiple species were common within the clades (Figure 1-2).

The analysis of Oliveira et al. (2007) also revealed several instances where accessions of the same species from different geographic areas did not form a clade (Figure 1-2). While these results are not particularly uncommon in species-level phylogenies when multiple populations of a species are sampled (Comes and Abbott 2001, Jakob and Blattner 2006, Syring et al. 2007, Ramdhani et al. 2011), they raise questions regarding the mechanisms responsible for such phylogenetic discontinuities and the validity of current taxonomic circumscriptions. Factors such as incomplete lineage sorting or gene flow from ancient or recent hybridizations can result in intraspecific non-monophyly (Edwards et al. 2008). Distinguishing between incomplete lineage sorting and hybridization is often difficult if not impossible (Rieseberg and Soltis 1991, Wendel and Doyle 1998) yet progress is being made with the development of new analytical methods to address these questions (see Joly et al. 2009, Yu 2011). However, many currently used loci for phylogenetic analysis cannot determine the degree of gene flow between species or populations because they often lack sufficient variability to discriminate between recently diverged species or clades. Assessing the extent of gene flow or degree of isolation is a critical component of many species concepts and thus is necessary for addressing species boundaries (Wheeler and Meier 2000). Population level studies with highly variable markers are the best way to determine admixture, migration, and population structure, all of which may ultimately aid in determining species limits.

This study aims to address the discontinuities observed in the phylogeny of Oliveira et al. (2007) utilizing microsatellite markers and population level-analyses. With this approach, we will use analyses of population structure, divergence, and gene flow to address species boundaries.

Study 1: Dicerandra linearifolia Complex

The most widespread and morphologically diverse annuals are members of the *Dicerandra linearifolia* complex. This complex is composed of *D. linearifolia* var. *linearifolia*, *D. linearifolia* var. *robustior,* and the recently named *D. fumella* (Huck 2010). *Dicerandra linearifolia* var. *linearifolia* occurs in west-central Georgia and eastern Georgia, *D. linearifolia* var. *robustior* is primarily found in the eastern panhandle of Florida between the Apalachicola and Suwannee Rivers, extending into southern Georgia, and *D. fumella* occurs primarily in the Florida panhandle west of the Apalachicola River and into southern Alabama (Figure 1-1) (Kral 1982, Huck 1987, 2010).

To frame the questions for this study, consider the phylogenetic analysis of Oliveira et al. (2007) with the Florida panhandle populations of *D. linearifolia* var. *linearifolia* relabeled to reflect the recent taxonomic revision to *D. fumella* (Figure 1-3). The phylogenetic tree shows the Florida panhandle taxa, *D. fumella* and *D. linearifolia* var. *robustior*, forming a clade (Figure 1-3). Interestingly, population 2, representing *D. fumella*, has identical DNA sequences for the markers used as population 8, identified as *D. linearifolia* var. *robustior*. The voucher specimens for these populations were consulted to verify proper identification of these species, and despite identical sequences, they are properly identified based on the morphological characters used to delimit each species. Additionally, the voucher specimen for population 3 labeled as *D.*

linearifolia var. *linearifolia* in the Oliveira et al. (2007) phylogeny, has subsequently been determined to be *D. linearifolia* var. *robustior* and is thus relabeled to reflect this change in identification. However, even with these changes, *D. linearifolia* var. *linearifolia* remains unresolved because of population 4, from western Georgia (Figure 1-3). We were interested in investigating at a population level the degree of molecular support for species designation of *D. fumella* and the relationship of the Georgia populations of *D. linearifolia* var. *linearifolia*.

Recognition of *D. fumella* as a species was based on its phylogenetic placement in a separate clade from the Georgia populations and on morphological differentiation. However, the morphological similarity between *D. fumella* and *D. linearifolia* var. *linearifolia* is rather striking upon first examination. The differences within the following characters are subtle, but have been cited by Huck (2010) as distinct for each species: leaf width, inflorescence architecture, number of flowers per cyme, corolla upper lobe shape, corolla tube length, and anther color (Table 1-1). A greater morphological distinction exists between *D. fumella* and *D. linearifolia* var. *robustior*, despite their parapatric distribution in the Florida panhandle and identical sequences at the surveyed loci (Table 1-1). However, sufficient morphological intergradation occurred near the geographic contact point between these two species for Huck (2010) to propose the existence of a relatively wide hybrid zone in the central panhandle of Florida.

Within *D. linearifolia* var. *linearifolia*, the west-central and eastern populations did not form a clade, and hybrids with *D. odoratissima* have been observed at the eastern contact zone where populations of the two species are found in proximity or are intermixed (Huck 1987). These rare hybrids display either an intermediate or chimeric

morphology between the two otherwise distinct parental species and show 90% pollen viability (Huck 1987).

To address more thoroughly the phylogenetic incongruence within the *D. linearifolia* complex and assess the molecular signature of hybrids, sampling was increased to represent the full ranges of *D. fumella*, *D. linearifolia* var. *linearifolia*, and *D. linearifolia* var. *robustior*, and *D. odoratissima*. Sampling included the proposed hybrid zone in the Florida panhandle and the two hybrid populations in eastern Georgia. Twenty populations totaling 265 individuals were genotyped at eight highly variable microsatellite loci to determine population structure, extent of gene flow, and hybridization.

Polyploidy has played a role in the evolution of *Dicerandra* (Huck and Chambers 1997). Chromosome analysis of meiotic pollen cells indicated differing chromosome numbers within the annuals: *D. fumella* 2*n* = 48, *D. linearifolia* var. *robustior* 2*n* = 32, with no available data for *D. linearifolia* var. *linearifolia* (Huck and Chambers 1997, Huck 2010). These high chromosome numbers for diploids (2*n* = 32) indicate the genus is likely the descendent of an ancient genome duplication event in an ancestral lineage having 2*n* = 16. While Huck and Chambers (1997) included all species recognized at the time, their counts were based on a limited geographic range and a small number of samples per species, in many cases only one. Thus, the actual extent of polyploidy within the annuals remains unknown. To address the ploidy of the annual species of *Dicerandra* in greater detail, live plants were collected and chromosome counts were performed from mitotically dividing root-tip cells of *D. linearifolia* var. *robustior* (2*n* = 32) and *D. densiflora* (2*n* = 48). Flow cytometry of silica-dried leaf tissue was then used to

screen a wider sample of individuals within the *D. linearifolia* complex. Verification of chromosome counts and denser sampling allow for a better understanding of the distribution of cytotypes and can provide increased insight into the role of genome duplication in the evolutionary history of the annuals.

The non-monophyly of species in the phylogeny of Oliveira et al. (2007) may reflect species divergence within the *D. linearifolia* complex, with some regions potentially representing independently evolving lineages with significant geographic barriers and distinct morphologies (as suggested by Huck's (2010) recognition of *D. fumella*). The disparate populations contained in these geographic regions may merit taxonomic recognition as distinct entities. To investigate these hypotheses, a fine-scale molecular analysis was necessary to: 1) assess the genetic structure within populations of *D. linearifolia* var*. linearifolia*, *D. fumella,* and *D. linearifolia* var. *robustior*, 2) determine if population genetic structure corresponds to geographic distribution and phylogenetic relationships, 3) evaluate the hypothesized hybrid zone between *D. fumella* and *D. linearifolia* var. *robustior* as well as between *D. linearifolia* var*. linearifolia* and *D. odoratissima,* 4) determine the extent of polyploidy within the complex, 5) assess species delimitations and potential conservation implications, and 6) give further insights into the evolutionary patterns of this clade of regional endemics.

Study 2: Dicerandra radfordiana and D. odoratissima

Dicerandra radfordiana has an extremely narrow distribution with only two recorded populations in McIntosh Co., Georgia. Its range could extend farther north within coastal plain communities but the definitive extent of its range remains unknown (Robin Huck and Tom Patrick-Georgia DNR lead botanist, pers. comm.). *Dicerandra odoratissima* is widespread in eastern Georgia and extends north to the southernmost

county of South Carolina, Jasper. *Dicerandra radfordiana* and *D. odoratissima* are morphologically divergent from other members of the genus, sharing most notably the distinct floral synapomorphies of tubular corollas, cucullate superior corolla lobe, slightly protruding anthers, and distinct cineol-based secondary compounds that producing the scent of cinnamon. All other annual species of *Dicerandra* display sharply geniculate, funnel-shaped corollas, an erect superior corolla lobe, strongly protruding anthers, and menthol-based secondary compounds that produce the scent of mint (Huck 1987). The morphological differences between *D. odoratissima* and *D. radfordiana* are primarily based on scale, with *D. radfordiana* having very similar floral and leaf architecture to *D. odoratissima,* but with larger dimensions (Table 1-2). Different pollinators have been suggested for *D. radfordiana* and *D. odoratissima*, hawkmoth and bee respectively (Huck 1987) but studies specifically addressing pollinators are needed. The chromosome analysis by Huck and Chambers (1997) showed that *D. odoratissima* and *D. radfordiana* have chromosome counts of 2*n* = 32. Currently, *D. radfordiana* is recognized as an endangered species by the state of Georgia and is listed federally as a category 2 species of concern with an unknown trend (USFWS 1993), with the need for additional work to establish range limits, assess population diversity, and verify species rank.

As with *D. linearifolia* var*. linearifolia* the analyses of Oliveira et al. (2007) revealed that *D. odoratissima* is non-monophyletic (Figure 1-2). The *D. radfordiana* sample used had identical sequences (ITS, *matK*, and *trnT-trnL*) to the most geographically proximate population of *D. odoratissima* and these samples shared one chloroplast and two nuclear synapomorphies. The remaining two populations of *D. odoratissima* had

unique, yet identical sequences. Geographically the three populations of *D. odoratissima* spanned the range of the species with considerable distance between the three sampled populations: 74, 81, and 110 km (Oliveira et al. 2007). However, the distance between *D. radfordiana* and the closest sampled population of *D. odoratissima* was only 750 m, and these two populations had identical sequences and are found at the southeastern extreme of *D. odoratissima's* range.

Multiple hypotheses could account for the observed non-monophyly and identical nucleotide sequences noted above. The shared sequences of the population of *D. odoratissima* and *D. radfordiana* could result from recent divergence of *D. radfordiana* from *D. odoratissima* with incomplete lineage sorting producing the observed similarity in DNA sequences. Alternatively, introgression of *D. radfordiana* genes into neighboring *D. odoratissima* through hybridization and backcrossing could also produce these phylogenetic results. Both of these mechanisms could produce DNA sequence similarities, yet maintain the morphological differences used to define these two entities as distinct species. Another hypothesis for the sequence similarities is clinal variation or isolation by distance. *Dicerandra radfordiana* occurs at the very southeastern portion of *D. odoratissima's* range and may appear as a distinct entity when compared morphologically or molecularly to individuals from the northern or western portion of the range. But when compared to geographically proximate populations, those distinguishing characteristics diminish. Given the small sampling from both species in the phylogeny of Oliveira et al. (2007), addressing these hypotheses requires broader geographic sampling and finer-scale molecular analyses.

We sampled 175 individuals from 10 populations of *D. odoratissima* spanning 150 km between the northern and southernmost samples and both known *D. radfordiana* populations. Seven highly variable microsatellite loci were developed to investigate the relationship between *D. radfordiana* and *D. odoratissima*, allowing us to assess: 1) if there is unique and significant genetic structure between *D. radfordiana* and *D. odoratissima*, 2) the degree of gene flow between geographically proximate populations of *D. radfordiana* and *D. odoratissima*, 3) the extent of isolation by distance, and 4) the implications of population structure or hybridization on species delimitation and conservation.

Table 1-1. Character differences between *D. linearifolia* var. *linearifolia*, *D. fumella*, and *D. linearifolia* var. *robustior* (modified from Huck (2010) *Dicerandra fumella* (Lamiaceae), a new species in the Florida panhandle and adjacent Alabama, with comments on the *D. linearifolia* complex. Rhodora, 112, 215-227).

Character	*D. linearifolia* var. *linearifolia*	*D. fumella*	*D. linearifolia* var. *robustior*
Leaf shape	Linear	Linear	Obovate, rhombic
Leaf width	0.5-1 mm	1-5 mm	2-10 mm
Leaf vestiture	Rugose, hispid	Rugose, hispid	Smooth
Inflorescence	Simple	Simple or compound	Compound
Inflorescence aspect	Open verticils	Very open verticils	Compact or loose verticils
Flowers per cyme	1-2	1-7, usually 3-5	3-7
Corolla upper lobe shape, size ratio in upright position	Ovate, appearing wider than tall	Cordate, appearing taller than wide	Cordate, appearing taller than wide
Corolla tube length	4-5 mm	6-7 mm	6-7 mm
Corolla color	White, pale pink, rarely deep pink	White, cream, pale pink, salmon-pink, pale lavender, very rarely magenta	Vivid reddish-purple, vivid purplish-red, magenta, deep pink, rarely pale pink
Anther color	Yellow	Yellow, orangish-yellow, streaked with reddish-purple	Reddish-brown, maroon

Table 1-2. Character differences between *D. odoratissima* and *D. radfordiana*. Measurements of leaf size, corolla length, tube length, and orifice width are means with ranges in parentheses (modified from Huck (1987) Systematics and evolution of *Dicerandra* (Labiatae). Phanerogamarum monographiae, 19, 1-343).

Character	*D. odoratissima*	*D. radfordiana*
Plant height	Up to 5 dm	Up to 6.5 dm
Habit	Often bushy	Virgate
Leaf shape	Linear	Narrowly obovate
Leaf size	25(15-45) mm x 1(1-3) mm	40(19-55) mm x 4(2-6) mm
Corolla length	14(14-21) mm	32(27-44) mm
Tube length	11(11-18) mm	25(20-32) mm
Orifice width	2 mm	4 mm
Pistil length	ca. 35 mm	ca. 68 mm

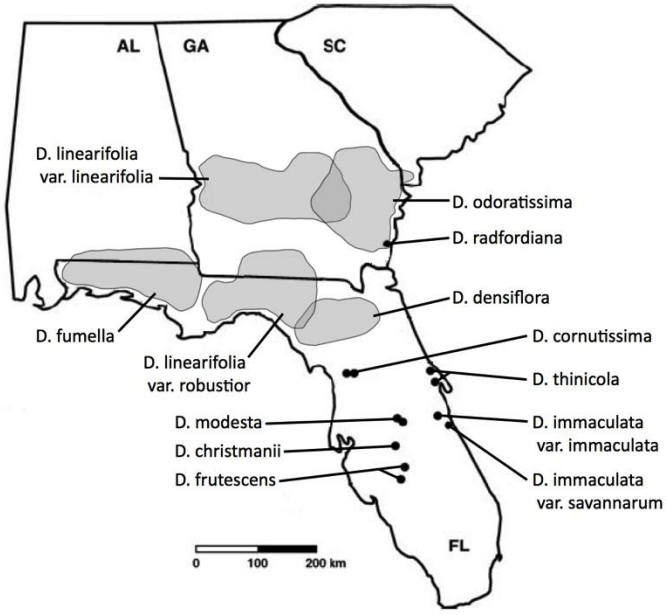

Figure 1-1. Approximate distribution of *Dicerandra* species. Species with very limited ranges are represented by points.

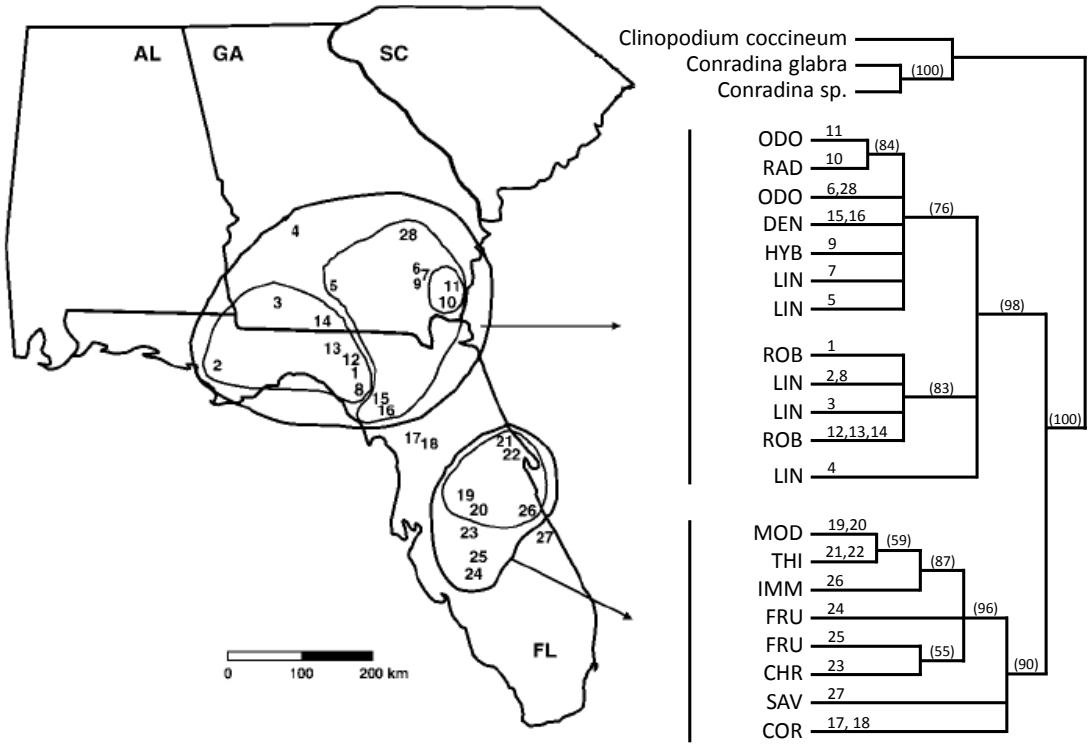

Figure 1-2. Geographic distribution of *Dicerandra* populations sampled in Oliveira et al. (2007). This cladogram is from a maximum likelihood analysis of the combined nuclear and chloroplast data sets. Numbers on branches correspond to sampling locations on the map; where several samples of the same species had identical sequences, they are represented as multiple numbers on a single branch. Bootstrap values are in parentheses. Two major clades correspond to the northern annuals and the southern perennials, respectively, with further geographic subdivision between the Atlantic and Gulf of Mexico drainages within the annual clade. Note the non-monophyly of both *D. linearifolia* var. *linearifolia* (LIN) and *D. odoratissima* (ODO) (Figure modified from Oliveira et al. (2007) Molecular phylogeny, biogeography, and systematics of *Dicerandra* (Lamiaceae), a genus endemic to the southeastern United States. American Journal of Botany, 94, 1017-1027).

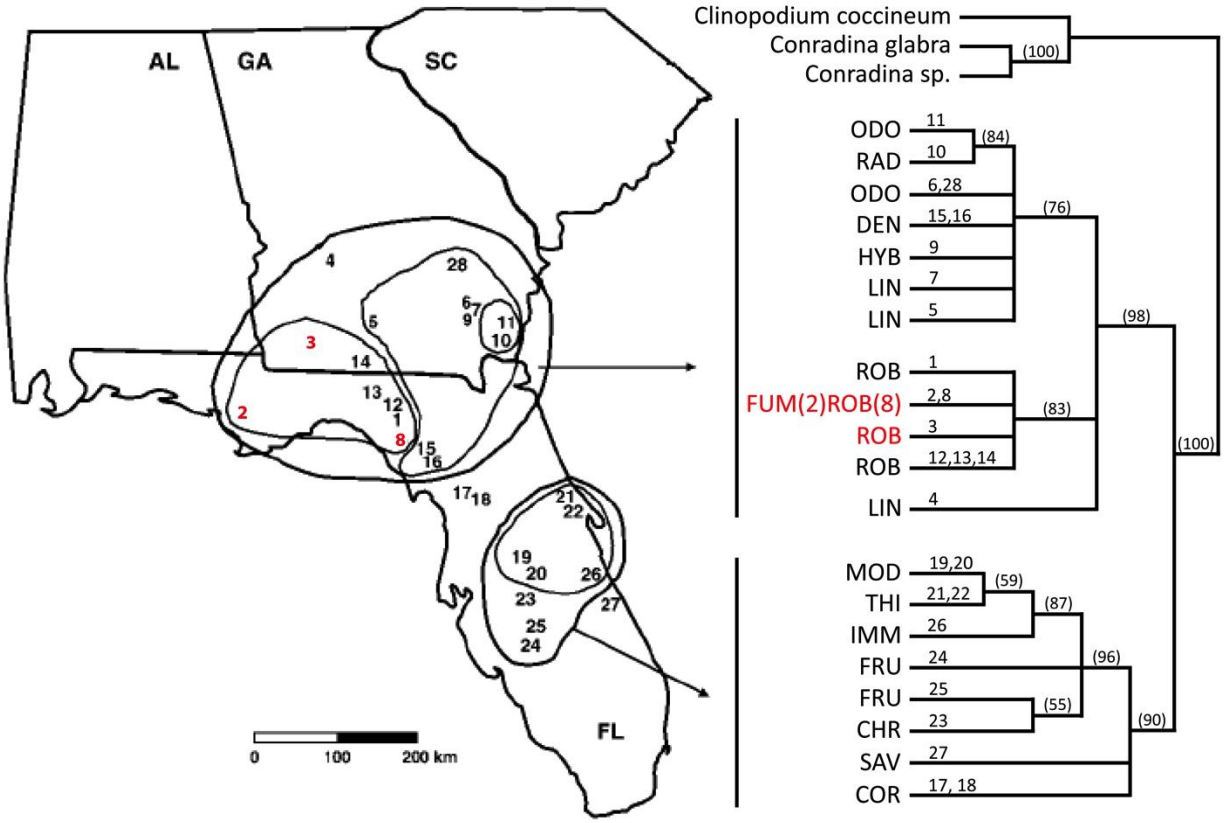

Figure 1-3. Oliveira et al. (2007) phylogeny with *D. linearifolia* var. *linearifolia* populations from the Florida panhandle relabeled to reflect the taxonomic revisions of Huck (2010) (*Dicerandra fumella* (Lamiaceae), a new species in the Florida panhandle and adjacent Alabama, with comments on the *D. linearifolia* complex. Rhodora, 112, 215-227). Taxa differing from original phylogeny are indicated in red. Population 3, formerly labeled as *D. linearifolia* var. *linearifolia* in the Oliveira et al. phylogeny, was relabeled as *D. linearifolia* var. *robustior* to reflect subsequent taxonomic revision to the original voucher specimen. Note the *D. fumella* (FUM) specimen, population 2, has an identical sequence as a specimen of *D. linearifolia* var. *robustior*, population 8. (Figure modified from Oliveira et al. (2007) Molecular phylogeny, biogeography, and systematics of *Dicerandra* (Lamiaceae), a genus endemic to the southeastern United States. American Journal of Botany, 94, 1017-1027).

CHAPTER 2
MATERIALS AND METHODS

Sampling Locations

Forty-seven populations were collected from Florida, Georgia, and southern Alabama (Figures. 2-1 & 2-2). The number of individuals collected at each population varied based on overall size and collecting conditions; 38 populations had 16 individuals sampled, and 9 populations had 8 individuals sampled (Table 2-1). Due to the annual habit, lower dispersal capability, and habitat specificity of these *Dicerandra* species, populations are almost always confined to areas of several hundred square meters or less with easily demarcated boundaries. Thus, for this study a population is defined as a sampling area where plants were determined to be consistently distributed.

Sampling locations were initially identified through herbarium collections with additional locations discovered through exploration of similar habitat in areas surrounding known samples. Collection took place from August through November 2010 and 2011, with the majority of sampling taking place in 2010. Upon arrival at a site, an outer perimeter was identified, general overview photographs were taken, and leaf material was collected from up to 16 individuals throughout the site. Leaf tissue was immediately stored in silica desiccant. Voucher specimens were collected at all sites and typically included both flowering and vegetative individuals. If flowering individuals were absent at the initial time of collection, sites were revisited to ensure sampling of identifiable individuals. Voucher specimens were deposited at the University of Florida herbarium (FLAS). Site locations were recorded utilizing GPS-derived Universal Transverse Mercator (UTM) coordinates in the North American Datum 1983 (NAD 83).

454 Sequencing

Genomic DNA was isolated from a single individual of *D. linearifolia* var. *robustior* from “site 7”, in Florida, using a modified CTAB extraction procedure (Doyle and Doyle 1987). In order to produce the desired quantity of 4 ug of high-molecular-weight DNA, multiple extractions were performed, the DNA was cleaned using a Geneclean Turbo kit, and individual extractions were combined and concentrated by sodium acetate/ethanol precipitation. Gel electrophoresis using a λ/*Hind* III ladder showed DNA fragments were between 8,000-11,500 base pairs in length. Nanodrop (Thermo Fisher Scientific, Inc. Waltham, MA, USA) analysis indicated sample concentrations of 360 ng/µL with ratios of DNA to proteins and organic compounds above the desired threshold of 2.

Based on prior studies using next-generation sequencing for microsatellite discovery, it was determined that a full lane ($1/8^{th}$ of a sequencing plate) was not necessary to generate the desired quantity of data (Castoe et al. 2010). Thus, the *Dicerandra* sample was barcoded during library preparation and combined with a sample of the monocot *Nolina brittoniana* (Nolinaceae), and pooled to run on $1/8^{th}$ of a sequencing plate. Unique MID barcodes were ligated onto the genomic DNA during library construction, prior to pooling of the two samples, allowing for accurate identification and separation of data for the two species. Sequencing was performed on a Roche 454 GS-FLX utilizing Titanium chemistry (Roche Diagnostics Corp., Indianapolis, IN, USA). Following sequencing, the data were separated by barcode using FASTX Barcode splitter (Gordon 2008), and the barcodes were removed using FASTX Trimmer (Gordon 2008).

Microsatellite Discovery and Primer Design

Microsatellite discovery and initial primer design were accomplished by utilizing a modified Perl scrip originally published by Castoe et al. (2010), and modified by M. Gitzendanner (University of Florida). This script searched the sequence data and identified repeated units of nucleotides, indicating the presence of a microsatellite region. Primer3 (Rozen and Skaletsky 2000) was then run, and primers flanking the repeat were generated. For this study di, tri, and tetra nucleotide repeats were targeted, with a minimum repeat number of 8. Primers were excluded from the 35 base pair region flanking each microsatellite. Primer design specifications, as allowed by Primer3, were as follows: optimum primer size: 20 (range: 18-30), product size range: 60-400, primer GC content 40-60%, GC clamp: 2 bp, optimal melting temp: 62 (range: 58-65), maximum self-complementarity: 4, maximum pair complementarity: 4, maximum end complementarity: 3, maximum homopolymers: 3.

The selected microsatellite repeats and their primers were further screened manually to ensure the microsatellites were perfect repeats, contained only short homopolymer runs (<8), and had high-quality base calling scores from the 454 software. Primers containing greater than 3 base pairs with a low-quality score were discarded. This screening produced 101 loci, 46 of which were chosen for PCR amplification testing across all species.

To ensure microsatellite loci were not located in known genic regions or in the chloroplast genome the 101 candidate loci were subjected to a BLAST search in GenBank. Any loci showing greater than 80% similarity with a known coding region or to a region within the chloroplast genome would be discarded. This approach ensured

that loci would be as neutrally evolving as possible, avoiding the potential for selection driving false population structure for a given locus.

Primer Testing

Primers selected for amplification testing had an M13 label, 5'CACGACGTTGTAAAACGAC, added to the 5' end of the forward primer to allow for genotyping utilizing fluorescently labeled probes. Primers were tested against a standardized panel of DNA composed of samples from *D. linearifolia* var. *linearifolia, D. linearifolia* var. *robustior*, *D. odoratissima*, and *D. radfordiana*. Initial testing took place with M13 primers not labeled with the florescent probe. All primers were tested for amplification twice against the same panel under the same PCR conditions to assess reliability of primer pairs.

PCR reactions were carried out in 10 µL volumes consisting of 1.5 µL H_2O, 2 uL 5 M betaine, 2 µL 5xGoTaq buffer, 0.8 µL 25 mM $MgCl_2$, 0.6 µL 0.8 uM forward primer, 0.9 µL 5 µM reverse primer, 0.9 µL 5 mM fluorescently labeled M13 primer (FAM, VIC, NED, or PET), 0.2 µL 2.5 mM dNTP, 0.1 µL Taq polymerase, and 1 µL 15 ng/ µL DNA sample. Reactions were carried out under the following conditions: 1) 3 min at 94°C, 2) 30 sec at 94°C, 3) 30 sec at 52°C, 4) 45 sec at 72°C, 5) 35 cycles of steps 2-4, and 6) 20 min at 72°C. Initial reactions were performed in 96-well plates, and subsequent reactions used 384-well plates, once successful loci were identified.

Amplification of microsatellites developed in one species which are then used across several species may experience some degree of allele dropout, where amplification consistently fails in a particular species but is successful in others. To maximize the benefit of successfully amplifying loci, even those not amplifying in every species, loci were divided into three categories based on the questions they were

capable of addressing. The first category was loci that amplified in all species of interest, the second group was those that amplified in at least the species of the *D. linearifolia* complex, and the third group consisted of those that amplified in at least *D. odoratissima* and *D. radfordiana*. This approach ensured there was a set of loci that could be used for addressing similarity across all species, as well as a potentially larger subset that could be used to address similarity on a finer scale, such as between two taxa or between geographic regions of a single taxon.

Genotyping

All samples were amplified via PCR with fluorescently labeled primers that allow for detection of PCR product length by sequencing instruments (Schuelke 2000). Samples were genotype on an ABI 3730XL (Life Technologies Corporation, Carlsbad, CA, USA), which can detect five fluorescent dyes, one for an internal size standard and four for the sample, allowing multiple loci of a single individual to be amplified and then pooled together for analysis. Further pooling of loci can take place by using the same dye if there is a significant length difference in the PCR products (Levitt et al. 1994). When the same dye is used for loci with differing PCR fragment sizes, the resulting chromatogram will display the allele size peaks of each locus with sufficient space separating the two to allow for confident assignment of alleles to their proper locus. Using this procedure, PCR products for 8 loci of a single individual were pooled and then genotyped. Due to the phenomenon of dye shift, where the observed allele size changes solely as a result of the fluorophore used, a single dye was chosen for each locus and used consistently throughout the study (Sutton et al. 2011).

PCR products were diluted prior to genotyping based on the strength of amplification as observed from agarose gel electrophoresis and inherent dye intensity;

VIC and FAM are stronger fluorophores than PET and NED. In general, loci amplified with VIC and FAM were diluted 1 µL PCR product in 39 µL diH_2O to a final volume of 40 µL, and loci amplified with PET and NED were diluted 4 µL PCR product in 36 µL diH_2O to a final volume of 40 µL.

Genotyping was performed by the Interdisciplinary Center for Biotechnology Research at the University of Florida. GeneMarker v.1.6 (SoftGenetics LLC, State College, PA, USA) was used to visualize chromatograms and design automated allele calling panels. Additionally, all allele-calls were verified manually.

Chromosome Counts

Huck and Chambers (1997) performed meiotic chromosome counts from dividing microsporocytes and described multiple ploidal levels within the annual species of *Dicerandra*. To verify ploidy across species, as well as to calibrate flow cytometry analyses, chromosomes from dividing root tip cells from *D. densiflora* and *D. linearifolia* var. *robustior*, two species with different published ploidal levels (Huck and Chambers 1997), were counted. Live specimens of *D. densiflora* (populations 46 and 47) and *D. linearifolia* var. *robustior* (populations “site 7”, 43, and 44) were collected in the field and reared in the Department of Biology greenhouse, with voucher specimens deposited at FLAS.

Root tips were collected from vigorously growing individuals and pretreated by exposure to nitrous oxide at 150 psi for 1.5 hours in a pressurized chamber. Pretreatment disrupts the development of microtubules, resulting in an accumulation of cells arrested at metaphase with condensed chromosomes capable of being visualized. Root tips were then fixed in 90% glacial acetic acid for 10 minutes, rinsed with 70% ice cold ethanol, and stored in 70% ethanol at -20°C. To prepare the root tips for enzyme

digestion, the samples were washed for 20 minutes in 10 mL of 1x citric buffer, repeated three times. The terminal 1-2 mm of the root tip was removed and placed in a 20 µL solution of 1% pectinase and 2% cellulase. The digestion incubated for 20 min at 37°C. Root tips were then rinsed with 70% ice cold ethanol, 20 µL of glacial acetic acid were added, and the tips were broken up using a blunt dissecting probe. Three microscope slides were placed in a high-humidity chamber, and 3.5 µL of the glacial acetic acid root tip slurry were dropped onto each slide. Slides were allowed to dry over the course of 10 minutes.

Chromosomes were stained with DAPI and visualized using a Zeiss Axio Imager M2 microscope (Carl Zeiss Microscopy LLC, Thornwood, NY, USA). To ensure accuracy, counts were only made from ruptured cells that were isolated and had few to no overlapping chromosomes. Each slide produced a sufficient quantity of isolated cells with highly condensed chromosomes so that a minimum of 3 cell counts could be made per slide.

Flow Cytometry

To assess the ploidy of a large number of populations, without requiring the rearing of live material for root tip counts, flow cytometry was used to assess the genome size of individuals from populations spanning each species distribution. Tissue was prepared following the procedures outlined in Suda et al. (2006). A critical aspect of successfully extracting intact nuclei is pairing the species of interest with an appropriate buffer. Four buffers were tried with varying degrees of success; extraction with Otto's buffer, Galbraith's buffer, and General Purpose Buffer (GPB) (Gruihuber et al. 2007) were either inconsistent or consistently failed. LB01, however, provided consistent, high-quality nuclei extractions if the buffer was prepared fresh or if it was

stored frozen. LB01 used in this experiment consisted of: 15 mM tris base, 2 mM Na2EDTA, 80 mM KCl, 20 mM NaCl, 0.1% (v/v) Triton X-100, 15 mM β-mercaptoethanol. Silica-dried *Dicerandra* tissue was chopped in 1 mL LB01 buffer using a fresh disposable razor blade with fresh tissue from a single individual of *Zea mays,* which served as an internal standard. Approximately 1 cm^2 of intervenous *Zea mays* tissue was used. The quantity of *Dicerandra* leaf tissue varied, partly as a function of the amount available and partly due to the size of the leaves, with 5-9 leaves used for *D. linearifolia* var. *linearifolia*, *D. odoratissima,* and *D. fumella* and 1-3 leaves used for *D. linearifolia* var. *robustior*, *D. densiflora,* and *D. radfordiana*. Following chopping, the buffer/nuclei slurry was filtered through a 30 µm screen into a 10 mL test tube, 40 µL of 50 µg/mL propidium iodide solution was added, and the samples were incubated on ice for 60 min. Analyses were performed on an Accuri C6 flow cytometer (BD Biosciences, San Juan, CA, USA) recording a minimum of 100 events within the gated region per sample.

Population Genetic Analyses

Linkage Disequilibrium, Hardy-Weinberg Equilibrium, Genetic Diversity

Loci were tested for linkage disequilibrium (LD) and Hardy-Weinberg equilibrium (HWE) using GENEPOP v.4.1 (Rousset 2008). LD was evaluated to test the null hypothesis that genotypes at one locus are uncorrelated with genotypes at another locus. Significant departure from HWE was assessed for each population. Markov chain sampling was used to evaluate significance (α=0.05) for both LD and HWE with Markov chain search parameters set to run 300 batches with 5,000 dememorization steps and 100,000 iterations per batch, and P-values were integrated using Fisher's exact test. Arlequin v.3.5.1.2 (Excoffier & Lischer 2010) was used to calculate observed

and expected heterozygosity. GenAlEx v.6.4 (Peakall & Smouse 2006) was used to generate the descriptive statistics of alleles per locus, allele frequency by population and locus, allelic richness, and presence of private alleles. *F*-statistics were calculated with GENEPOP v.4.1.

The genetic divergence metric of *D* (Jost 2008) was calculated in GENODIVE v.2.0 (Meirmans and Van Tienderen 2004). *D* was chosen over the traditional F_{ST} because the latter is less suited for highly polymorphic microsatellite loci. The upper value for F_{ST} of 1, denoting complete population differentiation, is mathematically related to the expected heterozygosity (H_e) of the population. As heterozygosity increases, the maximum value of F_{ST} decreases proportionally; this relationship can be described mathematically as F_{ST} (max)=1- H_e (Meirmans and Hedrick 2011). For example, if a H_e value of 0.8 is found then the maximum value of F_{ST} is 0.2, which may appear to describe a relatively low degree of population structure, yet this value actually represents the maximum possible population differentiation. Jost's (2008) *D* does not rely on H_e, but rather on the effective number of alleles. Jost recognized that increases in H_e do not scale linearly with increases in genetic divergence and thus metrics relying on H_e, such as F_{ST}, do not intuitively reflect population divergence (Jost 2008). For example, a two-fold increase in H_e does not correspond to a two fold increase in F_{ST} and when comparing multiple populations with differing H_e values, the magnitude of differentiation becomes difficult to interpret because the metric does not scale linearly, plus it is affected by a reduced upper limit (F_{ST} (max)=1-H_e). Because of these properties of F_{ST}, a metric not reliant on H_e is desirable. Jost's *D* relies on the effective number of alleles, which does scale linearly with diversity and is not affected by a

diminishing upper boundary as the number of alleles increases. Thus, *D* is a more appropriate metric than F_{ST} for describing microsatellite-based variation. As there is a long history of including F_{ST} values in population genetic studies even when highly polymorphic microsatellite loci are used, we present F_{ST} as well as *D* metrics for our analysis.

To visually assess the magnitude of genetic differentiation among populations, pairwise genetic distance matrices using F_{ST} and *D* were used to construct unrooted neighbor joining trees in PAUP* v.4.0 (Swafford 2003).

To assess the distribution of genetic variance, an analysis of molecular variance (AMOVA) was performed using Arlequin v3.5.1.2 (Excoffier and Lischer 2010). The 32 populations were assigned to groups based on current species delimitations to quantify how the total observed genetic variation was distributed: among groups, among populations within groups, among individuals within populations, and within individuals. Significance was tested using 10,100 permutations of the data.

Population Structure

To determine population structure the Bayesian clustering program STRUCTURE v.2.3.1 (Pritchard et al. 2000) was used to determine the likelihood for the placement of individuals into a series of increasing numbers of clusters (*K*). This method is used to help determine the maximum amount of organization, or distinct units, within a sample. The data were evaluated at varying scales, depending on the questions being addressed, for a total of 5 nested analyses. The most inclusive analysis tested all individuals across all species using 8 loci. This initial analysis was used to evaluate large-scale genetic structure primarily at the species level. The next analysis consisted of samples from the *D. linearifolia* complex using the same 8 loci, and was run with and

without the *D. linearifolia x D. odoratissima* hybrid populations from eastern Georgia (pop 23 and 25). The final nested analysis contained samples of *D. odoratissima* and *D. radfordiana* using data from 7 loci, with the analysis run with and without the hybrid populations (pop 23 and 25).

For each analysis STRUCTURE was run for 1,000,000 generations with a 200,000 generation burn-in for values of *K* ranging from 1-8. Each analysis was performed for 10 iterations per *K* value using the admixture model, correlated allele frequencies, with no a priori grouping of individuals by population ID. Estimation of the highest level of population structure (*K*) was determined using the Delta *K* method of Evanno et al. (2005) as implemented in Structure Harvester (Earl and VonHoldt 2011). The program CLUMP v.1.1.2 (Jakobsson and Rosenberg 2007) was used to average the probability of cluster assignment for each individual across the 10 independent iterations, and DISTRUCT v.1.1 (Rosenberg 2004) was used for plot visualization.

Pairwise comparisons of F_{ST} for all populations were calculated using Arlequin v3.5.1.2 (Excoffier and Lischer 2010). A permutation test of 10,000 replicates followed by a Bonferroni correction for multiple comparisons was used to determine significance differences between populations (α=0.05) (Rice 1989). Currently, no pairwise permutation analysis is available for Jost's *D*.

Isolation by Distance

To test the effect of geography on genetic differentiation, a Mantel test was used to evaluate the effects of isolation by distance. GenAlEx V.6.4 (Peakall and Smouse 2006) was used to test the correlation between pairwise Jost's D_{est} statistic and log-transformed linear geographic distances between site locations. Log geographic distance was used over linear distance because log distance best describes spatial

relationships in a two-dimensional space, such as a landscape. Additionally, to accurately evaluate the correlation between a metric that varies between 0 and 1, such as Jost's D_{EST}, comparisons must be made against log-transformed distances. These data were randomly permuted 9,999 times to construct a null distribution to evaluate the hypothesis that increasing geographic distance correlates with increasing genetic differentiation. This analysis was performed on the complete data set as well as on a data set containing species within the *D. linearifolia* complex and a data set of *D. odoratissima* and *D. robustior*, with the latter two excluding the hybrid populations (pop 23 and 25).

Table 2-1. Sampling locations, number of individuals per populations, and whether the population was utilized for this study.

State	Population Number	Species	Num. of Individuals Sampled	Used in Molecular Analysis
GA	1	LIN	2	N
GA	2	LIN	11	Y
GA	3	LIN	16	Y
GA	4	RAD	16	Y
GA	5	ODO	16	N
GA	6	ODO	16	N
GA	7	ODO	16	N
GA	8	ODO	10	N
GA	9	ODO	16	Y
GA	10	LIN	16	Y
GA	11	ODO	16	N
GA	12	ODO	16	N
GA	13	ODO	16	Y
GA	14	ODO	16	Y
GA	15	ODO	16	Y
GA	16	ODO	16	Y
GA	17	ODO	16	Y
FL	18	FUM	16	Y
FL	19	FUM	16	N
GA	20	ODO	16	Y
GA	21	ODO	16	Y
GA	22	LIN	16	Y
GA	23	HYB,LIN,ODO	22	Y
GA	24	LIN	16	Y
GA	25	HYB, ODO	16	Y
GA	26	LIN	16	N
GA	27	LIN	16	Y
GA	28	LIN	16	Y
GA	29	LIN	16	Y
GA	30	ODO	8	Y
GA	31	ODO	8	Y
GA	32	LIN	8	Y
GA	33	LIN	8	Y
FL	34	FUM	16	Y
AL	35	FUM	14	Y
FL	36	FUM	16	Y
FL	37	FUM	8	Y
FL	38	FUM	8	Y
FL	39	FUM	8	N
FL	40	ROB	5	Y
FL	41	ROB	5	Y
GA	42	RAD	15	Y

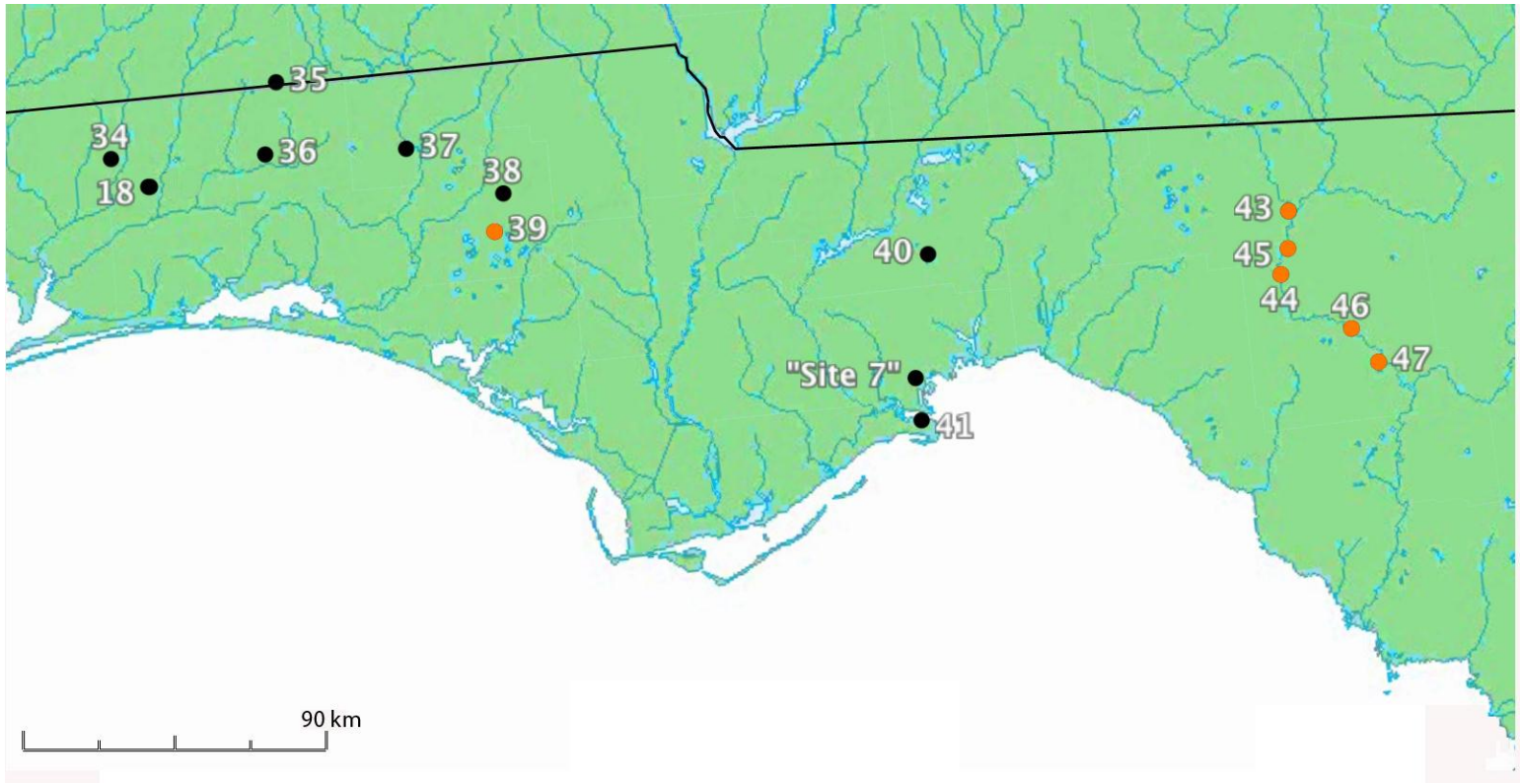

Figure 2-1. Locations sampled in Florida panhandle. Populations designated by orange circles were sampled but not included in the population-level analysis. Site 7 is the location from which the sample for the 454 sequencing run was collected; this population was not used for further analyses.

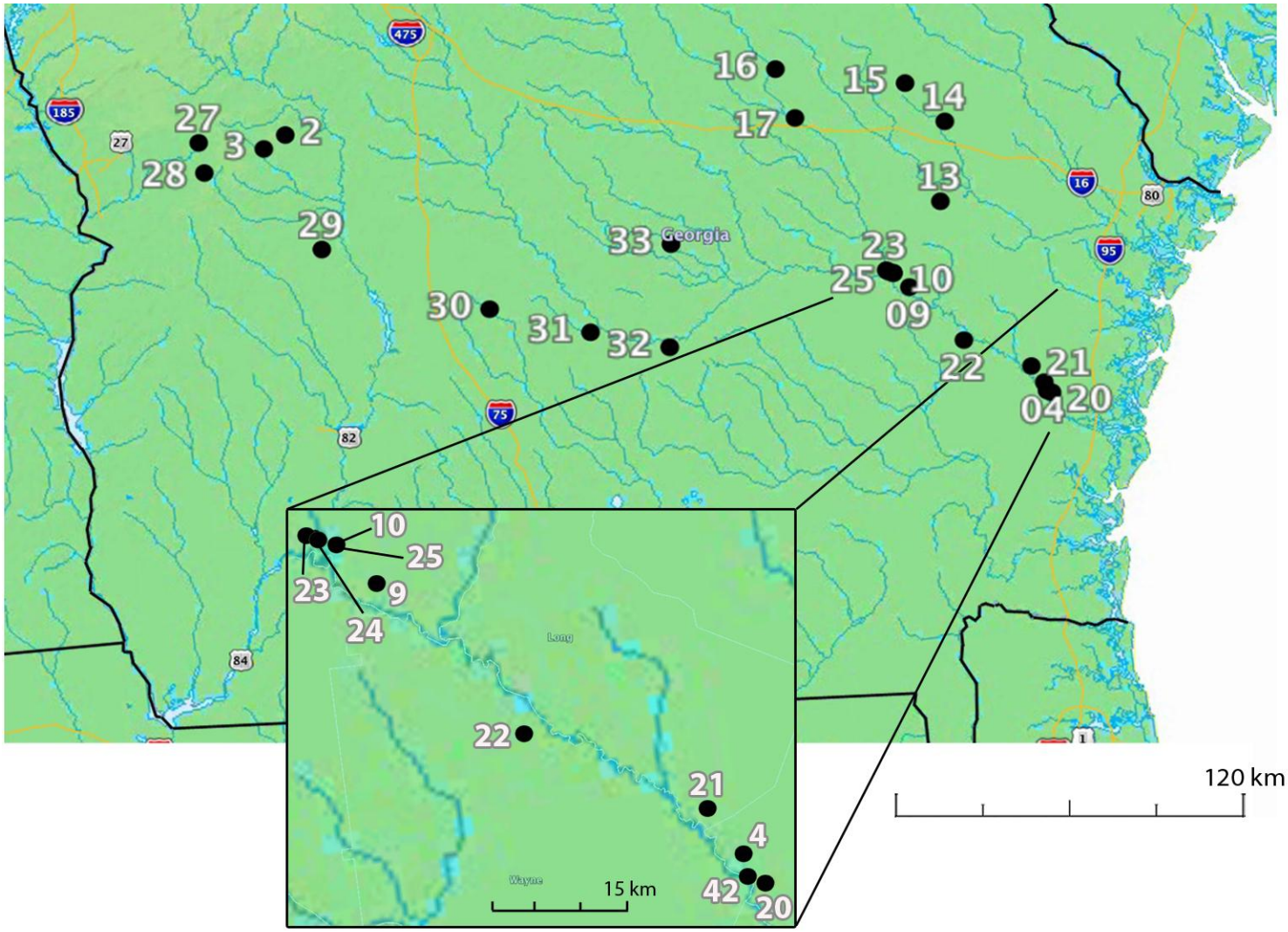

Figure 2-2. Locations sampled in Georgia. All populations shown were used in the analyses. The insert displays the populations along the Altamaha River, including the two populations of *D. radfordiana* (pop 4 and 42) and the two populations containing putative hybrid individuals (pop 23 and 25). Population 10 and population 25 are approximately 30 meters from one another and thus are represented by a single point.

CHAPTER 3
RESULTS

Next Generation Sequencing

The 454 titration run sequenced a total of 51,540,789 base pairs in 137,609 reads with an average length of 374 bp and an average quality score of 33.3. Of the total reads; 63,196 were from the *Dicerandra* sample, 72,761 were from *Nolina*, and 1,652 reads could not be matched during the sorting of reads by barcode. Of the total reads assignable to species, 46.48% were *Dicerandra* and 53.51% were *Nolina*, close to the ideal 1:1 ratio for samples pooled in the same sequencing lane.

Microsatellite Discovery, Primer Design, and Testing

Of the 63,196 *Dicerandra* reads generated during the sequencing run; 3,349 contained microsatellite loci, 1,065 dinucleotide loci of 6 repeats or greater, 1,247 trinucleotide loci of 4 repeats or greater, and 1,037 tetranucleotide loci of 3 repeats or greater. Despite the high number of microsatellites detected, primers could only be designed for 1,668 loci, and of these, 1,567 contained microsatellites with fewer than the desired 8 repeats. This resulted in 101 reads with 8 or more repeats with primers meeting out specifications (Tables 3-1 & 3-2).

The BLAST search did not reveal any of the 101 loci to be within known coding regions, thus lowering the probability of them being under selection, nor were any within the chloroplast genome. Each locus and its associated primer sites were visually inspected, with 46 loci being selected for amplification testing. These 46 loci were divided into two categories: 1) 25 loci with high-quality base calls in the primer sites, unbroken microsatellite repeats, and a low frequency of homopolymer runs; and 2) 21

loci with poorer-quality base calls or longer homopolymer runs, but all contained unbroken repeats.

Within the first category, 6 of the 25 loci did not amplify in any of the 4 species in the test panel. Eight of the loci showed amplification across the entire test panel, and 11 amplified in 2 or 3 of the species. Of the 21 loci in the second category, 14 failed to amplify in any of the species in the test panel, but the remaining 8 loci amplified in all of the species. The 16 loci that amplified in all species of the test panel were then genotyped.

Of the species within the test panel, *D. linearifolia* var. *robustior* had the highest proportion of amplifying loci, consistent with the fact that this was the species sequenced. However, when comparing the number of loci which amplified in some but not all of the species, there was no pattern between phylogenetic relatedness and probability of amplification. Taxa more closely related to *D. linearifolia* var. *robustior* did not have a greater proportion of loci amplify than more distantly related taxa.

Genotyping

The 16 loci that amplified across the test panel were genotyped in a subset of individuals from populations across the range of each species. This enabled the early detection of monomorphic loci as well as loci producing chromatograms that could not be scored easily and consistently. Of the 16 loci genotyped, 3 appeared monomorphic across all species and were excluded from subsequent tests. Five loci produced chromatograms that could not be accurately scored. The morphology of these chromatograms either consisted of several scattered peaks across a wide range of base pairs, often with considerable background noise, or consisted of a smooth mound of peaks, each one base pair apart, with no distinguishable features to identify the true

fragment length (Figure 3-1). These 5 loci were also excluded from further analysis. The remaining 8 loci (locus 2, 5, 8, 10B, 16B, 18, 20, 23) produced clear chromatograms that were scored easily and consistently (Figure 3-1).

Despite showing amplification across all species in the test panel, not all of these 8 loci consistently amplified across all the species when the number of populations was increased. Once all individuals were genotyped, 5 loci amplified in greater than 95% of individuals. The remaining 3 amplified in either the *D. linearifolia* complex or in *D. odoratissima* and *D. radfordiana* (Table 3-3).

Chromosome Counts

Huck (1997) reported *D. densiflora* and *D. linearifolia* var. *robustior* as 2*n*= 48 and 2*n*=32, respectively. To verify these observations we sampled 4 individuals of *D. densiflora* from populations 46 and 47 and 6 individuals of *D. linearifolia* var. *robustior* from populations “Site 7”, 43, and 44. However, we found no difference in chromosome number between these species; all 10 individuals had 2*n*=32 (Figure 3-2). Each individual’s count was corroborated by at least three metaphase or prometaphase cells, and often more, with minimal physical overlap.

Flow Cytometry

Analyses of genomic DNA concentrations in relation to a standard of known size were performed to assess variation in ploidal level. The log difference between the standard and a sample of *D. densiflora*, with a confirmed chromosome count of 2*n*=32, was calculated and used to compare against samples to determine if there was an increase in DNA content. The difference in chromosome numbers between 2*n*=32 and 48 represents a 33.3% increase in DNA content. Screening for specimens with 2*n*=48 chromosomes was based on comparing a sample’s log difference in absorbance to the

standard, with the difference of the known 32 chromosome sample and the standard. Any samples with an approximately 33% increase in DNA content would have been considered a polyploid.

Flow cytometry was performed on 2 individuals each from 15 populations: Populations 46 and 47 of *D. densiflora*, populations 40 and 41 of *D. linearifolia* var. *robustior*, populations 34, 18, and 36 of *D. fumella*, populations 27, 29, 22, and 33 of *D. linearifolia* var. *linearifolia*, populations 15, 13, and 21 of *D. odoratissima*, and population 4 of *D. radfordiana*. Across all samples no significant deviation in the observed absorbance compared to the known 32-chromosome sample was found. This lack of observed polyploids from the flow cytometry analysis is consistent with the observations from the chromosome counts, indicating all annual species of *Dicerandra* have a chromosome number of 2*n*=32. The lack of chromosome number variation was corroborated by the microsatellite data, where all loci evaluated had a maximum of two alleles. If polyploids existed it is likely there would be loci that contained more than two alleles, since the genome would consist of more than two copies of a locus resulting from the duplication event; however this was not observed in any of the loci examined.

Locus Characteristics and Fragment Variation

The number of alleles observed at each locus varied considerably, ranging from a maximum of 39 alleles at locus 20 to a minimum of 5 alleles at locus 23 (Table 3-3). The nature of microsatellite expansion, increasing in discrete intervals, predicts that a locus with a high number of alleles will also show high variation in fragment length, since differences in fragment length are assumed to reflect differences at the microsatellite repeat. Locus 20 displayed the largest variance in fragment size, 48bp, and also had the greatest number of alleles. Similarly, locus 23, with the fewest

observed alleles, exhibited the smallest variance in fragment size, 13bp. To verify if the large variation in fragment length seen at certain loci was the result of increases in tandem repeats at the microsatellite, instead of indels in the flanking region, a limited number of homozygous individuals exhibiting large differences in allele length were sequenced. Loci 2, 5, and 10B, all of which showed a high number of alleles (Table 3-3), were sequenced in 12 individuals. Within this relatively small sample, multiple sources contributed to the variation in fragment length other than expansion of the microsatellite repeat. Indels in the flanking region and point mutations in the repeat or the flanking sequences were found in all loci sequenced. Locus 2 contained 20 alleles across all species, with some being significantly smaller fragments than expected: for example, 173 bp and 174 bp in length compared to the expected 234-266 bp. Sequencing revealed the small allele size was the result of a 72 bp deletion, which included the microsatellite repeat. Locus 5 contained 31 alleles spanning a fragment size range of 45 bp. The vast majority of variation at this locus was due to expansion of tandem repeats; however, a 3 bp deletion in the flanking region was present in several sequenced individuals. Locus 10B displayed 21 alleles with a fragment size range of 26 bp. Of the individuals sequenced, variation in the repeat region was low, with only point mutations being detected. However, indels in the repeat's flanking sequences were common and ranged from 1-14 bp.

Overall, sequencing of these microsatellite loci showed the variation in fragment length was attributed to both expansion of the microsatellite as well as indels in the flanking sequences. While not strictly attributed to expansion of the microsatellite repeat, mutations in the flanking regions are heritable and thus still remain valuable in

assessing population structure and genetic similarity. For this reason, no attempt to correct allele size for variation in the flanking region was made. Additionally, because the source of fragment length variation was not exclusively from the microsatellite repeat, a stepwise mutation model is not valid for analyzing this data set and thus R_{ST} was not used as a divergence metric. Instead, F_{ST} and Jost's *D*, which both rely on an infinite alleles model, were used.

Linkage Disequilibrium, Hardy-Weinberg Equilibrium, and Heterozygosity

None of the loci analyzed showed evidence of significant linkage disequilibrium (α=0.05). Of the 32 populations sampled 9 (28%) were in Hardy-Weinberg equilibrium (HWE) (Table 3-4). The low percentage of populations in HWE is not of great surprise given the HWE assumptions and the life-history traits of these organisms. Populations not in HWE may have resulted from founders events or have undergone bottlenecks. Current factors such as selection or small population size can also result in deviation from HWE; however, the size of the population did not predict whether a population was or was not in HWE. Populations 28, 32, and 33 which contained 100's-1,000's of individuals at the time of sampling were in HWE. In contrast populations 30, 37, 38, and 40 contained fewer than 50 individuals and were also in HWE. The same is true for those populations not in HWE; size at time of sampling does not appear to predict if a population is in HWE or not. However, given the annual nature of these plants, the isolation of most populations, and relatively small size of many populations, genetic drift could be a considerable factor contributing to the observed levels of diversity within each population.

Mean observed heterozygosity (H_o) among populations was 0.584 with a range of 0.275-0.8. In 28 of the 32 populations, H_o was lower than expected heterozygosity (H_e)

(Table 3-4). Within these 28 populations, 23 showed a statistically significant deficiency of heterozygotes, including both populations of *D. radfordiana* (α=0.05). The four populations showing higher H_o than H_e were all *D. fumella*, with two showing a statistically significant excess of heterozygotes.

The populations observed to be deficient in heterozygotes, and thus not in HWE, may be deficient due to biological factors mentioned previously (founders events, bottlenecks, small population sizes). However, an alternative explanation for the observed deficiency of heterozygotes is population substructure producing a Wahlund effect. If there was strong differences in allele frequencies between two portions of a population which were analyzed as a single population, this could result in a significant deviation from HWE (Wahlund 1928). This would be perceived as a deficiency in heterozygotes even though each subpopulation may individually be in HWE. In order for a significant deviation in HWE to be solely from a Wahlund effect there must be a high frequency of individuals homozygous for different alleles within the population. Close evaluation of the 8-locus data set for all populations with a significant deviation from HWE revealed that at each locus the vast majority of homozygous individuals were homozygous for the same allele. This indicates the deviation from HWE and the observed deficiency of heterozygotes is not due to unrecognized population substructure but more likely the result of biological factors acting on the population as a whole. The overwhelming presence of a single allele for all homozygotes is unusual, especially since there was little to no representation of homozygotes for the alternative alleles seen in the heterozygotes. The low frequency of homozygotes with the

alternative alleles and fewer than expected heterozygotes is likely the cause for the significant deviation from HWE and the deficiency in heterozygotes.

Genetic Diversity Among Populations

Genetic differentiation was measured using two metrics, F_{ST} and D_{EST}, both with a range of 0-1. D_{EST} is an estimator of Jost's *D* that takes into account small population size and the uncertainty of not knowing the true allele frequencies. These values were calculated using the 5-locus data set, which had the greatest coverage across all species (>95% of individuals could be genotyped at these 5 loci), and allows for a direct comparison across species. In general, F_{ST} and D_{EST} values showed a similar pattern (Table 3-5) despite having different values. This is to be expected since F_{ST} calculations are likely lower due to the highly polymorphic nature of the loci used and the inability of F_{ST} to scale with increasing levels of heterozygosity.

D_{EST} values for *D. odoratissima* and *D. linearifolia* var. *linearifolia* showed the highest degree of interpopulational variation, D_{EST} = 0.560 and 0.437 respectively. Both species had an equal number of populations, n=10, and spanned the widest geographic area of all sampled species (Table 3-5). *Dicerandra radfordiana* and *D. linearifolia* var. *robustior* had lower D_{EST} values, 0.304 and 0.205, respectively. However, both species were represented by only two populations, by necessity for *D. radfordiana* given that only two populations are known to exist. But *D. linearifolia* var. *robustior* is more widespread, and its corresponding D_{EST} value could change if additional populations were included. There is high uncertainty associated with the D_{EST} value for *D. linearifolia* var. *robustior* as demonstrated by the standard error, 0.223, which is higher than the D_{EST} value itself, 0.205; thus this value should be interpreted with significant caution. *Dicerandra fumella* had the lowest D_{EST}, 0.263, indicating less interpopulational

differentiation than the other species, despite the samples representing a relatively wide geographic area.

Mean F_{ST} values varied among species, ranging from a minimum of 0.149 in *D. fumella* to a maximum of 0.254 in *D. odoratissima* (Table 3-5). Overall, F_{ST} values were highest in species with the greatest geographic range as seen in *Dicerandra linearifolia* var. *linearifolia* and *D. odoratissima*. Reciprocally, species with smaller geographic ranges, *D. fumella, D. linearifolia* var. *robustior*, and *D. radfordiana* showed less genetic differentiation among populations (Table 3-5).

The AMOVA showed that of the total genetic variance 12.49% was among species, 11.39% was among populations within species, 9.36% was among individuals within populations, and 66.76% was within individuals. Permutations tests indicated all values were significantly different than zero (α=0.05).

Software to perform pairwise comparisons of Jost's *D* using permutation based tests for statistical significance is not available. However, Bonferroni-corrected pairwise comparisons of F_{ST} values, using all individuals for the 5-locus data set, revealed mostly statistically significant differences between populations. Of the 496 pairwise comparisons only 15 were not statistically different (Table 3-6). Of these 15 pairwise comparisons, one was between two populations of *D. odoratissima*, seven were between populations of *D. linearifolia s.l.*, four were between populations of *D. fumella*, two were between *D. linearifolia* var. *linearifolia* pop 33 and both hybrid populations (pop 23 and 25), and the final non-significant comparison was between two populations of different species, *D. linearifolia* var. *linearifolia* and *D. odoratissima*. Most of the populations displaying non-significant pairwise comparisons are found in the same

geographic region, but are not necessarily proximate. The shortest distance between two of these populations is 9 km, but many are separated by over 30 km. In most cases there are populations geographically closer, yet these have differentiated sufficiently to be considered statistically different. Given the distances involved, these non-significant associations are probably not due to current gene flow, especially given that the insects observed visiting flowering plants are generalists, and thus not specifically seeking *Dicerandra* individuals. Additionally, the habitat is often discontinuous, further reducing the probability of pollinator-mediated gene flow.

Visualization of the magnitude of genetic differentiation was achieved using neighbor-joining trees constructed from pairwise distance matrices of F_{ST} and Jost's *D*. Overall, the trees are similar in topology and branch length and both place individuals into two main groups, one containing the *D. linearifolia* complex and the other containing *D. odoratissima* and *D. radfordiana* (Figure 3-3). Yet within these two groups populations of *D. linearifolia* var. *linearifolia* do not cluster together, nor do all populations of *D. odoratissima*. The *D. linearifolia* var. *linearifolia* cluster also contains the populations of *D. linearifolia* var. *robustior* nested within it (Figure 3-3). Within the *D. linearifolia s.l.* cluster there are geographic trends where populations in the same region are typically more genetically similar to each other than to populations in a different region; however, most geographic regions do not exclusively contain populations from that area. For example, there is general cohesion among the western populations of *D. linearifolia* var. *linearifolia* (Figure 3-3, blue); however, population 22 from eastern Georgia is nested among this predominantly western cluster. Similarly, the central Georgia populations (Figure 3-3, orange) also do not form an exclusive

cluster. Thus, there is not support, in the form of exclusive geographic clusters, for a clear genetic distinction between geographic regions of *D. linearifolia* var. *linearifolia.* While there is support for the exclusive clustering of *D. linearifolia* var. *robustior* (Figure 3-3, light blue), these populations are nested closely among the central Georgia *D. linearifolia* var. *linearifolia* and do not show an affinity to the proximate *D. fumella* populations from the Florida panhandle.

Both the F_{ST} and D_{EST} trees show all populations of *D. fumella* forming an exclusive cluster. Interestingly, the length of the branch leading to the *D. fumella* clade is as great or greater than branches leading to other named species. This supports Huck's (2010) recognition of these Florida panhandle populations as their own distinct species. There is also considerable genetic distinction between *D. fumella* and *D. linearifolia* var. *robustior*, despite their geographic proximity. The genetic distinction of *D. fumella* in these trees is also well supported in the STRUCTURE analysis (Figure 3-5), giving further credence to the recognition of *D. fumella* as a species.

Similar to the situation with the *D. linearifolia* complex, *D. odoratissima* populations also do not from an exclusive cluster. While the two populations of *D. radfordiana* consistently are sister, they are found on the same branch as two geographically proximate populations of *D. odoratissima* (pops 20 and 21). Additionally, these four populations are the same ones recognized as a distinct cluster in the STRUCTURE analyses (Figure 3-5).

There is little geographic cohesion among the populations of *D. odoratissima*: the northeastern (Figure 3-3, green), the central (Figure 3-3, orange), and one southeastern population (Figure 3-3, yellow) are all nested together and do not form geographic

based clusters. The possible exceptions are the two *D. radfordiana* populations and their proximate *D. odoratissima* populations (pop 20 and 21). These populations represent the southeastern extreme of these species' ranges and are found at the end of a relatively long branch, suggesting their location at the edge of the range is also associated with increased genetic differentiation. However, there is a 50 km gap between *D. odoratissima* population 21 and the next closest *D. odoratissima* population, 9. This gap was not sampled but is know to contain populations of *D. odoratissima*, thus the length of the branch may be misleadingly long due to the absence of these populations.

Isolation by distance, evaluated with a Mantel test, showed a significant correlation between geographic distance and genetic distance for *D. odoratissima* and a combined analysis of *D. linearifolia* var. *linearifolia* and *D. linearifolia* var. *robustior* (Figure 3-4). No correlation was seen with *D. fumella*, even though the sampled populations represent the extent of the species' range.

Population Structure

STRUCTURE analysis of all individuals using both the 5- and 8-locus data sets showed significant genetic structure which generally followed currently delimited species boundaries (Figure 3-5). The Delta *K* method for determining the maximum number of genetically significant clusters showed the highest support for *K*=4 (Evanno et al. 2005). However, determining the true number of clusters can be contentious as different methods for determining the maximum number of clusters often produce differing results (Ball et al. 2010). It is important to consider the biological aspects of the organism when evaluating the clustering pattern at a given *K*. As a result, it is often prudent to examine the clustering patterns for values of *K* surrounding what was

determined by the given method, especially if there could be biologically pertinent differences captured by an alternative clustering pattern. In this analysis, *K*=4 was consistently chosen by the Delta *K* method. However, the Delta *K* method relies heavily on the observed variance of the log likelihood value of multiple independent STRUCTURE runs. The *K* value with the smallest variance is often mathematically favored when compared to *K* values with even slightly larger variances. In this analysis, *K*=4 showed very low variance. When compared to *K*=5, which also had a very low variance but was larger than that of *K*=4, *K*=4 was selected as having the highest probability of being the "correct" *K*. Since *K*=5 also exhibited a very low variance the clustering pattern at this value was also evaluated.

Within the combined STRUCTURE analysis the first cluster, red, includes all populations of *D. fumella*, 18 and 34-38 from the Florida panhandle, which form a distinct genetic cluster with low to no signs of genotypic admixture with other clusters (Figure 3-5). These populations were collected along a 115 km stretch of the Florida panhandle from the Escambia River in the west to the center of the Marianna lowland in the east. This includes samples from the type locality of *D. fumella*, population 34, as well as two populations, 37 and 38, within the zone of hybridization proposed by Huck (2010) between *D. fumella* and *D. linearifolia* var. *robustior*. However, our analyses found no evidence of admixture in any of the individuals of *D. fumella* with individuals of *D. linearifolia* var. *robustior* or *D. linearifolia* var. *linearifolia*.

Individuals of *D. linearifolia* var. *linearifolia* and *D. linearifolia* var. *robustior* formed the second cluster, blue (Figure 3-5, *K*=4). This cluster spans the greatest geographic area, ranging from the Florida panhandle to west-central Georgia and east to the

Altamaha River. This cluster displays a high degree of cohesion even though the most extreme populations are separated by as much as 300 km, with many separated by 100 km or more.

The final two clusters, purple and yellow, contain populations of *D. odoratissima* and *D. radfordiana*. These clusters, while well supported, do not follow current species delimitations. The yellow cluster includes the two *D. radfordiana* populations as well as two nearby populations of *D. odoratissima*. The remaining eight *D. odoratissima* populations constitute the purple cluster.

The placement of *D. radfordiana* and *D. odoratissima* populations in the same cluster correlates strongly with geography. The distance between the two *D. radfordiana* populations is 3 km, with the two closest *D. odoratissima* populations (pop 20 and 21) being 1.8 km and 7.3 km from the closest *D. radfordiana* population. All four of these populations (4, 42, 20, and 21) occur across 11.7 km. The next closest *D. odoratissima* population sampled (pop 9) is 50 km away and is a member of a different cluster (purple), which does not show any signs of admixture with any of the individuals from the yellow *D. radfordiana* cluster. The two *D. odoratissima* populations within the same cluster as *D. radfordiana* are also supported in the neighbor-joining trees based on Jost's *D* and F_{ST} distance matrices (Figure 3-3).

The clusters observed in the analysis of all individuals were identical to those observed in the two less-inclusive analyses focusing on the *D. linearifolia* complex and *D. odoratissima* and *D. radfordiana*. Additionally, there was no change in cluster assignment of populations between the 5- and 8-locus data sets.

In general, the cluster assignments of most populations do not change from *K*=4 to *K*=5 except for certain populations of the *D. linearifolia* var. *linearifolia* and *D. linearifolia* var. *robustior* cluster (Figure 3-5). At *K*=5, the two populations of *D. linearifolia* var. *robustior* are pulled out in a cluster, green, along with two populations of *D. linearifolia* var. *linearifolia* from eastern Georgia, as well as the *D. linearifolia* var. *linearifolia* individuals of the putative hybrid populations, pop 23 and 25, also from eastern Georgia. The western populations of *D. linearifolia* var. *linearifolia* remain in a well-defined cluster, blue. Populations 32 and 33 are *D. linearifolia* var. *linearifolia* from central Georgia, and both populations show a high degree of admixture with the eastern (green) and western (blue) clusters. Based on this pattern, it could be hypothesized that there is clinal variation in *D. linearifolia* var. *linearifolia*, with the east and west being genetically distinct regions and central Georgia being an area of intergradation. This pattern would be supported ecologically as well, because dispersal capability is limited and pollinator-mediated gene transfer is likely low, allowing for genetic drift, or selection, to reinforce particular genotypes at either extreme of the range. However, population 22, which is the eastern most population of *D. linearifolia* var. *linearifolia,* is made up of individuals primarily of the western genotype. The placement of population 22 in the same cluster as the western populations is also seen in both the F_{ST} and Jost's *D* trees (Figure 3-3).

Hybrid Detection

To evaluate the molecular support for a hybridization zone between *D. fumella* and *D. linearifolia* var. *robustior*, STRUCTURE analyses containing all populations of *D. fumella, D. linearifolia* var. *linearifolia,* and *D. linearifolia* var. *robustior* were performed (Figure 3-6). In all circumstances the highest support was for *K*=2, with one cluster

containing populations of *D. fumella* and the second cluster containing all the *D. linearifolia* populations. However, regardless of the *K* value there was no support for any individuals of *D. fumella* or *D. linearifolia* var. *robustior* being of hybrid origin.

Additional STRUCTURE analyses containing individuals from mixed populations of *D. odoratissima* and *D. linearifolia* var. *linearifolia* in eastern Georgia were performed to determine if there was a detectable molecular signature for putative hybrid individuals, identified in the field based on intermediate or chimeric morphology. The two putative hybrid populations, pop 23 and 25, contained individuals that were morphologically identified as *D. linearifolia* var. *linearifolia, D. odoratissima*, and as putative hybrids. The putative hybrids had corolla architecture similar to *D. linearifolia* var. *linearifolia* except the stamens were positioned along the upper lobe or were upwardly curving with minimal stamen exertion, a characteristic of *D. odoratissima*. Additionally, the putative hybrid individuals had a distinct odor more reminiscent of the cineol compounds of *D. odoratissima* than the menthol compounds of *D. linearifolia* var. *linearifolia*, despite showing floral morphology more similar to *D. linearifolia* var. *linearifolia*. The STRUCTURE analysis of the populations containing the putative hybrids showed many individuals with mixed genotypes with portions from both of the suspected parental species, as well as individuals with genotypes composed primarily of *D. linearifolia* var. *linearifolia* (blue and green depending on the *K* value employed) or *D. odoratissima* (purple) (Figure 3-6). The Delta *K* method identified *K*=2 as having the greatest support. The clustering pattern for *K*=3 was also investigated, but did not reveal additional meaningful patterns. When the individuals in populations 23 and 25 are segregated based on identification in the field, we see there is not a clear molecular distinction

between the two parental species. Individuals identified as *D. linearifolia* var. *linearifolia* and *D. odoratissima* have large portions of their genotype that are assigned to the other species (Fig 3-6). In population 23 at *K*=2, the genotypes for two of the eight samples of *D. linearifolia* var. *linearifolia* are almost entirely assigned to the *D. odoratissima* cluster. Similarly, 4 of the 8 samples of *D. odoratissima* have genotypes almost exclusively assigned to the *D. linearifolia* var. *linearifolia* cluster. Within the putative hybrids only two individuals show a genotype that is significantly mixed, each with approximately 65% assigned to *D. linearifolia* var. *linearifolia* and 35% assigned to *D. odoratissima*. This pattern suggests there has been backcrossing of hybrids to the parental lines, creating individuals that are phenotypically more similar to one parent but genetically more similar to another. The observation of individuals that are not a 50:50 mixture of each parental species suggests the observed hybrids are likely not F_1 hybrids and may represent F_2 or later generation progeny and/or backcrosses.

Population 25 consists of individuals of *D. odoratissima* and putative hybrids, but population 10 which is approximately 30 m from population 25 is exclusively *D. linearifolia* var. *linearifolia*. The STRUCTURE analysis strongly supports population 10 being exclusively *D. linearifolia* var. *linearifolia*, with no evidence of admixture within this population (Figure 3-6). This suggests there may be a directional component to hybridization, where hybrids are produced with *D. linearifolia* var. *linearifolia* exclusively being the pollen donor and *D. odoratissima* exclusively being the pollen recipient. Yet within population 25, of the eight individuals identified as *D. odoratissima*, three are genotypically identical to *D. linearifolia* var. *linearifolia*, two display a near 50:50 genotypic mixture, and the remaining three display approximately 80:20 genotypic

mixture between the two parental species. A similar situation is observed in the individuals identified as hybrids with five of the eight being exclusively *D. linearifolia* var. *linearifolia* genotypes, one displaying almost an exclusively *D. odoratissima* genotype, and two show a 35:65 genotypic mix.

In both of these putative hybrid populations there is a clear molecular signature showing certain individuals are of hybrid origin. However, the individuals identified as hybrids based on possessing what appeared to be intermediate or mosaic morphology did not necessarily carry a molecular signature for a hybrid origin.

Table 3-1. 101 reads with microsatellites containing 8 or more repeats for which primers could be constructed.

Monomer Size, bp	Num. of Repeats	Occurrences
2	8	27
	9	24
	10	13
	11	1
	12	4
Total		69
3	8	17
	9	6
	10	3
	12	1
	14	1
	16	1
	17	2
	19	1
Total		32

Table 3-2. 1,567 reads with microsatellites containing fewer than 8 repeats for which primers could be constructed.

Monomer Size, bp	Num. of Repeats	Occurrences
2	6	143
	7	64
Total		207
	4	528
3	5	159
	6	54
	7	38
Total		779
4	3	520
	4	40
	5	19
	7	2
Total		581

Table 3-3. Description of the primer sequences, loci characteristics, and allelic complement for the 8 loci utilized across all the annual species. Single asterisk denotes loci that amplified consistently in all species, double asterisk only amplified in species of the *D. linearifolia* complex, triple asterisk loci amplified across species but was inconsistent between populations.

Locus Num.	Locus Name		Primer Sequence 5'-3'	Repeat Motif	Dye Used	Num. of Alleles	Fragment Size (bp)
2*	H2BD8	F:	AAAGGAAGACAAGGGTCAAGG	(TC)8	VIC	20	173, 174, 234-266
		R:	AGAAGCCAGCAACAAGAACG				
5*	H41X1	F:	TCGCAGAACAAGTAGTCAAAGC	(TC)10	PET	31	211-256
		R:	GGAAGCTCAATCCATAGAAAGG				
8*	H61XY	F:	GAAGCACCATATCAAGATTCACC	(AC)8	FAM	23	198-245
		R:	GGACATTCAGCTCTCCATCG				
10B***	IAXR6	F:	TGGTGCTACTGCTACTGCTGG	(AC)9	NED	21	211-237
		R:	CGACATTGGTCTGAATCATGG				
16B***	IGK5Z	F:	TTGTGGAAAGTTATTGACACCC	(TC)9	PET	11	184-209
		R:	CCTCACGGACTCGACTGG				
18*	IMEMS	F:	CTCACCTATGCGGAGATGG	(TC)9	NED	19	237-265
		R:	AACTCGACTCAAAGCAATCG				
20*	IAPG4	F:	AGGCGTCCCATTCTGAAGC	(TC)12	FAM	39	257-305
		R:	AGGTTGGTGAGGGTGTTGG				
23**	IJQ5D	F:	TTTCGTAGGAGGTAGATGTAGAAG	(AT)9	VIC	5	302-315
		R:	GAGCAAACTTGGGAAGGAGG				

Table 3-4. Hardy-Weinberg equilibrium p-values, inbreeding coefficient (F_{IS}), heterozygosity deficiency and excess tests, and observed and expected heterozygosity (H_o, H_e). Asterisk denotes significant values.

POP Num.	Species	HWE p-value	F_{IS}	H_e def. p-value	H_e excess p-value	H_o	H_e
2	LIN	0	0.361	0	1.0	0.475	0.764
3	LIN	0	0.198	0	1.0	0.520	0.643
4	RAD	0	0.270	0	0.999	0.417	0.619
9	ODO	0	0.243	0	1.0	0.489	0.644
10	LIN	0.208	0.059	0.069	0.931	0.582	0.618
13	ODO	0	0.229	0	1.0	0.530	0.703
14	ODO	*0.001	0.203	0	0.999	0.518	0.616
15	ODO	*0.043	0.120	*0.032	0.969	0.553	0.685
16	ODO	0	0.308	0	1.0	0.512	0.735
17	ODO	0	0.255	0	1.0	0.574	0.779
18	FUM	*0.014	-0.070	0.222	0.779	0.743	0.698
20	ODO	0	0.294	0	1.0	0.423	0.628
21	ODO	0	0.201	0	0.999	0.602	0.765
22	LIN	*0.0208	0.111	*0.015	0.985	0.613	0.701
23	HYB	0	0.245	0	1.0	0.623	0.821
24	LIN	0	0.190	0	0.999	0.645	0.793
25	HYB	0	0.242	0	1.0	0.635	0.830
27	LIN	0	0.313	0	1.0	0.455	0.699
28	LIN	0.062	0.076	*0.041	0.956	0.643	0.711
29	LIN	0	0.269	0	1.0	0.566	0.771
30	ODO	*0.101	0.074	0.152	0.845	0.535	0.588
31	ODO	0.019	0.156	*0.031	0.967	0.589	0.704
32	LIN	0.955	0.031	0.295	0.707	0.716	0.739
33	LIN	0.325	0.103	0.059	0.935	0.692	0.783
34	FUM	*0.008	-0.086	0.591	0.413	0.739	0.684
35	FUM	*0.026	-0.216	0.995	*0.003	0.757	0.627
36	FUM	*0.037	-0.124	0.878	0.129	0.781	0.694
37	FUM	0.371	-0.207	0.993	*0.007	0.800	0.671
38	FUM	0.590	0.114	*0.046	0.956	0.675	0.756
40	ROB	0.143	0.586	*0.001	0.999	0.275	0.597
41	ROB	0.376	0.200	*0.037	0.966	0.600	0.733
42	RAD	*0.035	0.263	*0.002	0.998	0.393	0.526

Table 3-5. Genetic diversity metrics D_{EST} and F_{ST}.

	Num. of pops.	Mean Num. of alleles	Eff. Num. of alleles	D_{EST}	D_{EST} Std. error	Mean F_{ST}	F_{ST} Std. error
D. fumella	6	12.667	3.412	0.263	0.137	0.149	0.036
D. linearifolia var. linearifolia	10	16.857	3.453	0.437	0.046	0.221	0.067
D. linearifolia var. robustior	2	4.167	2.096	0.205	0.223	0.151	0.07
D. odoratissima	10	17.000	2.824	0.560	0.059	0.254	0.033
D. radfordiana	2	4.429	2.134	0.304	0.13	0.158	0.05

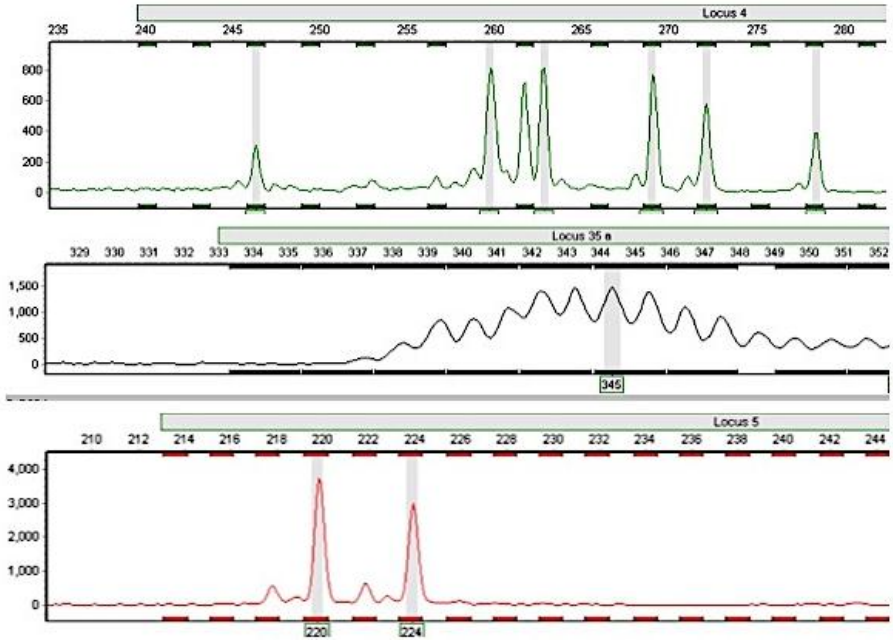

Figure 3-1. Examples of microsatellite chromatograms showing peaks from loci that were unscorable (top and middle) as well as those easily and consistently scorable (bottom).

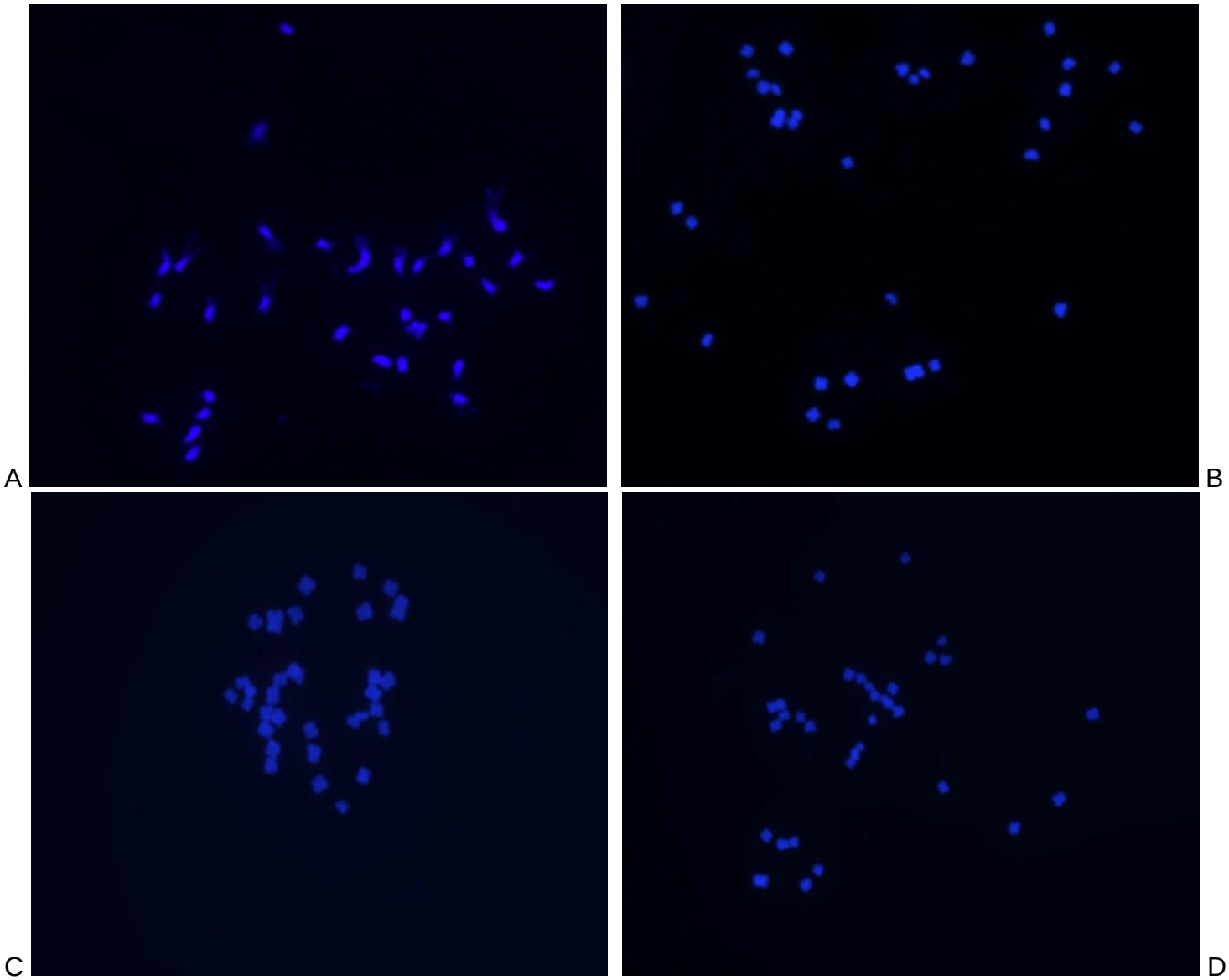

Figure 3-2. Chromosome images. A) 32 prometaphase chromosomes from *D. linearifolia* var. *robustior*. B) 32 metaphase chromosomes from *D. linearifolia* var. *robustior*. C&D) 32 metaphase chromosomes from *D. densiflora*.

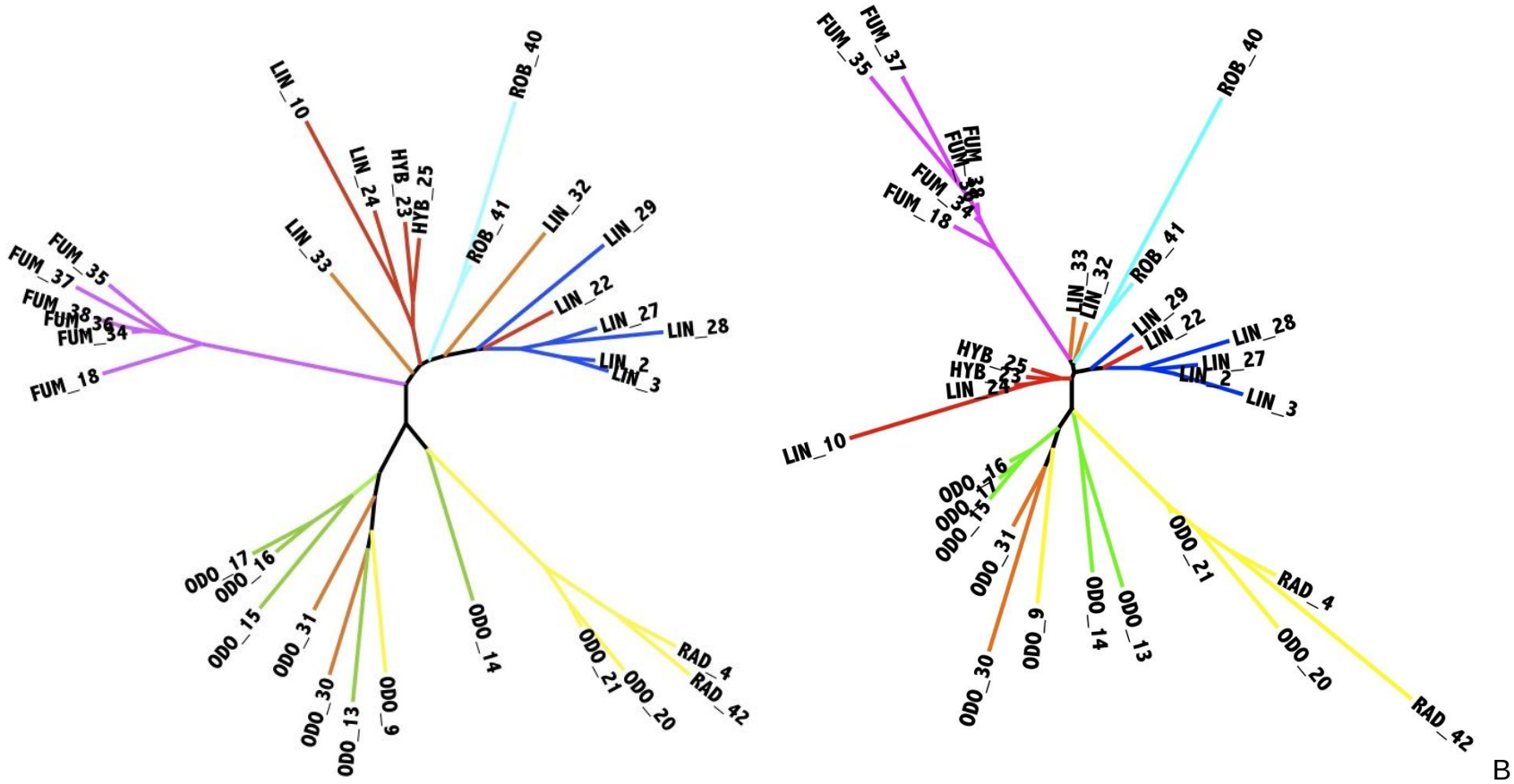

Figure 3-3. Genetic distance based neighbor-joining trees of all annual populations of *Dicerandra* using 5 loci. A) Jost's *D* distance matrix. B) F_{ST} distance matrix. Both trees two main clades containing 1) the *D. linearifolia* complex 2) *D. odoratissima* and *D. radfordiana*. Colors represent geographic regions: Pink: Florida panhandle, Light Blue: central Florida, Yellow: southeast Georgia, Red: eastern Georgia, Orange: central Georgia, Blue: western Georgia, Green: northeastern Georgia.

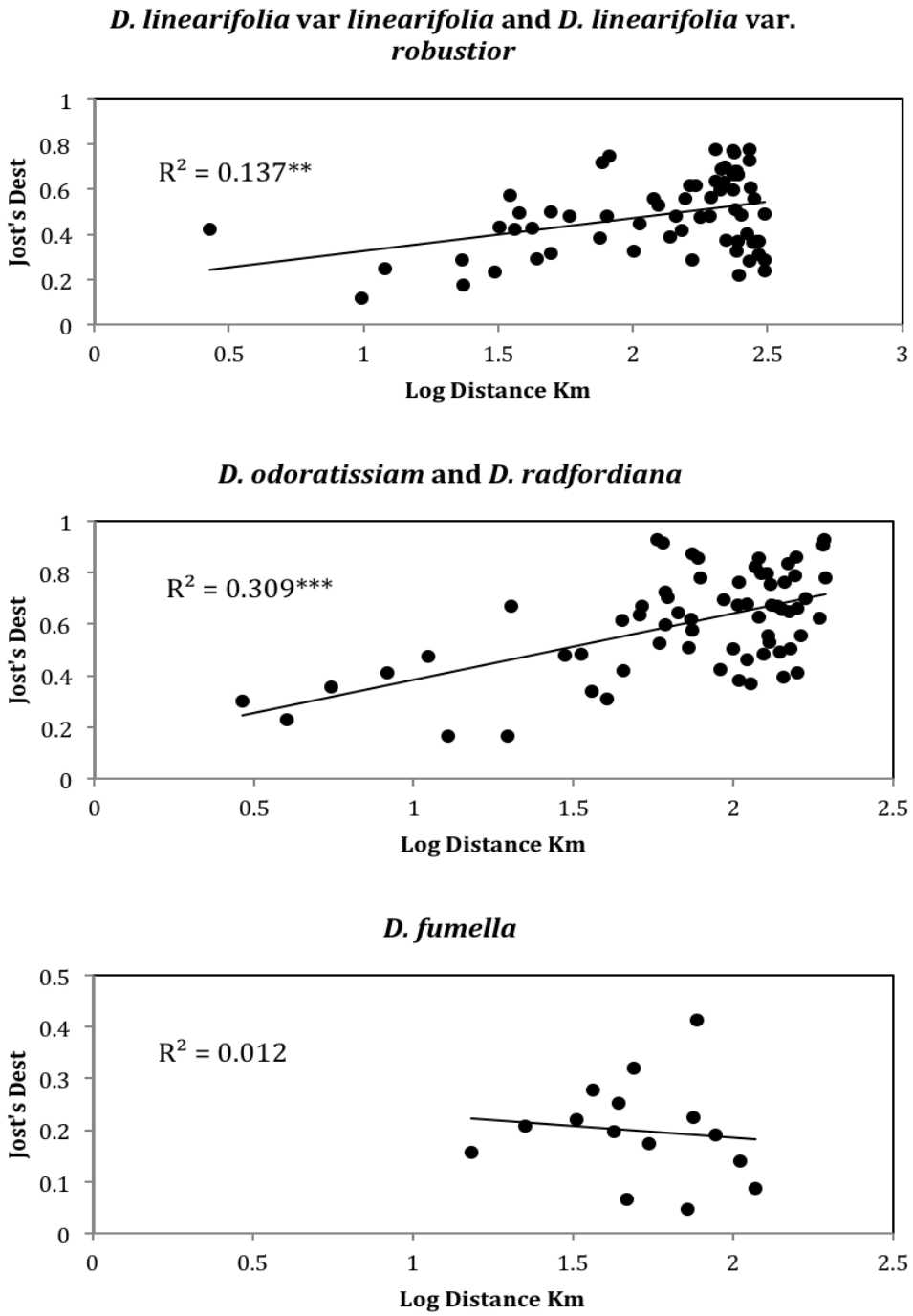

Figure 3-4. Mantel test results evaluating isolation by distance. Asterisk indicates significance of the correlation: *p= < 0.05, **p= < 0.005, ***p= < 0.0005. A significant correlation between increasing geographic distance and genetic distance was found in *D. linearifolia* var. *linearifolia* and *D. linearifolia* var. *robustior*, as well as *D. odoratissima* and *D. radfordiana*.

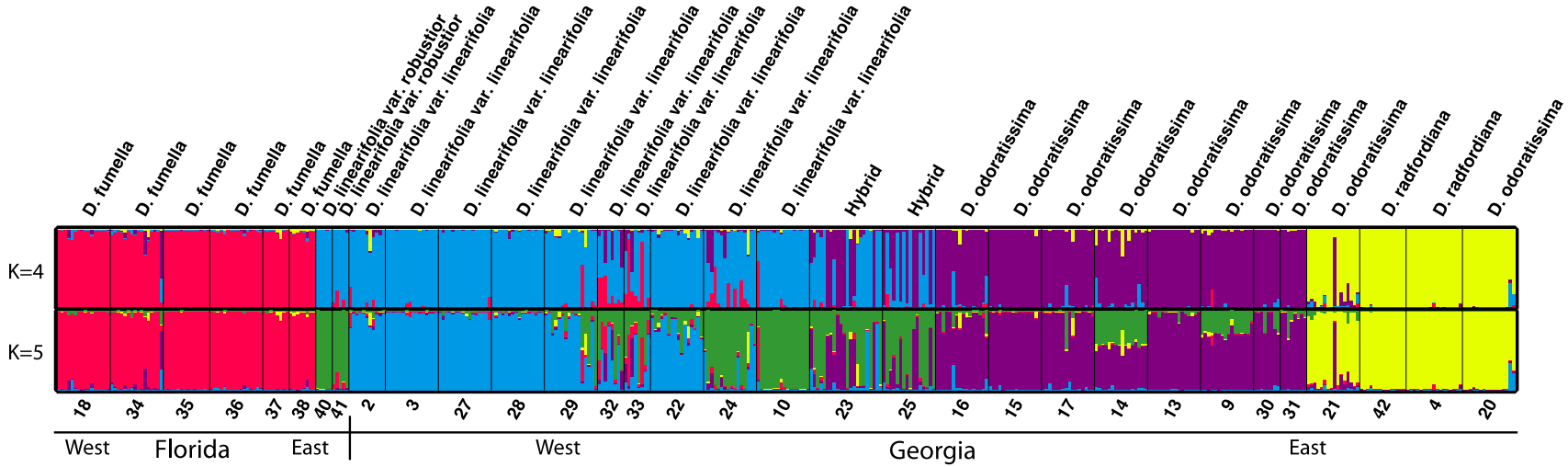

Figure 3-5. Structure analysis of all 448 individuals at 8 loci, Delta *K* analysis indicates maximum population structure occurs at *K*=4. Population numbers are listed below, corresponding species names are listed above. Populations are arranged geographically.

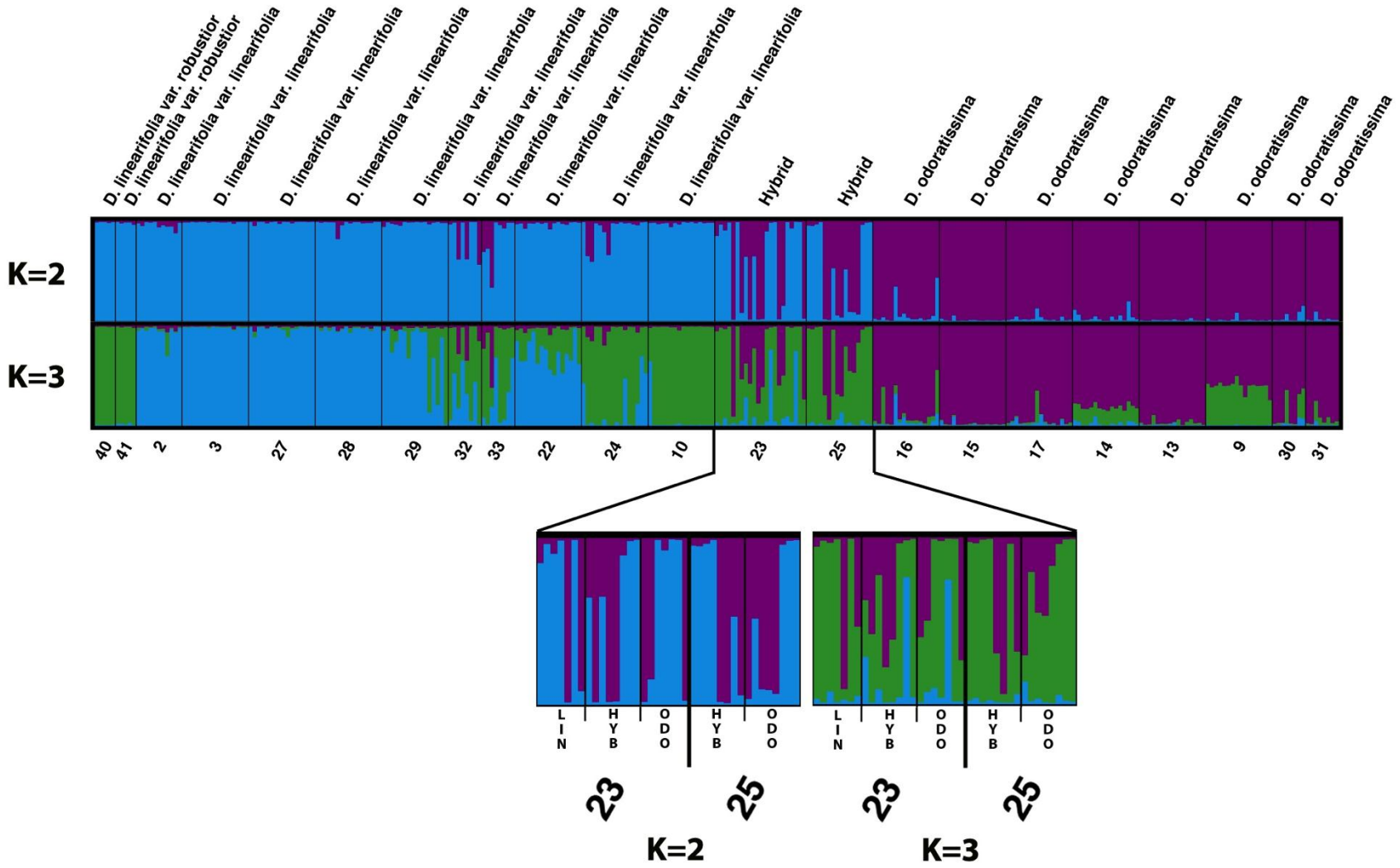

Figure 3-6. Structure analysis of *D. linearifolia* var. *linearifolia* and *D. odoratissima* populations with a focus on the hybrid populations 23 and 25 which contain individuals of suspected hybrid origin as well as putative parental species, *D. linearifolia* var. *linearifolia* and *D. odoratissima*. Blue and green represents the clusters containing individuals of *D. linearifolia* var. *linearifolia* purple represents the cluster containing individuals of *D. odoratissima*.

CHAPTER 4
DISCUSSION

Next-Generation Sequencing for Microsatellite Discovery

The use of next-generation sequencing for the discovery and development of population-level markers has seen a remarkable increase in use since its start only a few years ago (e.g. Adbelkrim et al. 2009, Allentoft et al. 2009, Castoe et al. 2010). Our recovery of 3,349 microsatellite loci from one-half of 1/8 of a sequencing plate further shows the utility of this method for marker discovery. However, for a variety of reasons, the detection of a large quantity of loci does not directly result in a high number of usable loci. The lack of sufficient flanking sequence for primer design and flanking sequences not meeting primer design specifications reduced the number of potential loci to 1,668. When loci with under eight repeats were then removed, this significantly reduced the number of loci to 101, yet this is still more than sufficient for population genetic analyses. Reducing the threshold for the minimum number of repeats would have provided a larger pool of candidate loci if additional loci were desired. This was especially true for tri- and tetranucleotide repeats. After removing loci with fewer than eight repeats, two-thirds of the remaining loci were dinucleotide repeats. Relaxing the number of repeats to seven would have produced an additional 104 loci (64 di-, 38 tri-, and two tetranucleotide loci), all of which had primers meeting the desired specifications (Tables 3-1 & 3-2).

The quality of the base calls in the sequence from which the primers were designed was strongly associated with the success of PCR amplification of the locus. The 46 loci for which primers were tested were divided into two categories. The first category containing primers designed from sequences with few to no poor-quality base

calls had 25% that failed to amplify in every species. However, the second category containing loci with primers designed from sequences with some lower-quality base calls had 66% that did not amplify in any species. While several factors could contribute to the differential success in amplification between the two categories, the primary justification for assigning loci into their respective category was the quality of the sequence from which the primers were designed. Based on this observation, we suggest only testing primers designed from sequences with the highest quality score to maximize successful amplification of candidate loci.

Overall, amplification of loci across species boundaries was successful. Of the 26 loci that amplified, only three were exclusively found in *D. linearifolia* var. *robustior,* the species used for the sequencing effort, and 11 amplified in one or two additional species. Consistent cross-species amplification was required for this study, and thus only loci showing amplification in all members of the test panel were considered for further testing. Because loci that amplified in all species were specifically selected, this has likely introduced some bias into our estimates of genetic diversity. These estimates are likely conservative since more rapidly evolving sites have a greater probability of accumulating mutations in the primer binding sites and thus cease to amplify within those species or populations. Those loci, that did not show amplification in all species were removed from the analysis. Thus, it is unlikely that our estimates of genetic diversity are elevated as a result of marker selection.

Genetic Diversity

The AMOVA results showed there was substantial genetic differentiation found among populations within a species, 11.39%, and this quantity is similar to the amount of variation found between species, 12.49%. This further illustrates the importance of

populations for maintaining overall genetic diversity within a species. Population level variation is further illustrated with comparisons of F_{ST} *and* D_{EST} where genetic differentiation between populations was found to be high within most species. This is especially true in the widely distributed species *D. odoratissima* and *D. linearifolia* var. *linearifolia*, where D_{EST} values were 0.56 and 0.437, respectively, and F_{ST} values were 0.254 and 0.221 (Table 3-5). This level of differentiation was also reflected in the branch lengths of the genetic distance trees (Figure 3-5). Long branches often separate populations within the same species, indicating the magnitude of genetic differentiation among clusters of the same species. Furthermore, the Mantel test revealed a positive correlation between geographic distance and genetic distance in *D. odoratissima* and *D. linearifolia* var. *linearifolia*, indicating proximate populations are more genetically similar to one another than they are to populations farther away. This correlation is consistent with high D_{EST} values and the long branch lengths observed in the distance trees. The patterns seen across these three measures of genetic divergence consistently show there is considerable genetic diversity contained at the population level. In the case of *D. odoratissima* and *D. linearifolia* var. *linearifolia,* 56% and 43.7%, respectively, of the total variation is contained within populations, as indicated in the D_{EST} values, the remainder is found among individuals (Table 3-5).

In *D. fumella* and *D. radfordiana* there was less variation at the population level: 26.3% and 30.4%, respectively, based on D_{EST} values. The magnitude of this reduced variation between populations in *D. fumella* is represented in the distance tree where populations are found on considerably shorter branches than those of *D. odoratissima* and *D. linearifolia* var. *linearifolia* (Figure 3-5). Additionally the lack of a correlation

between geographic distance and genetic distance in the Mantel test further illustrates *D. fumella* has less genetic divergences among the populations (Figure 3-4). This is further corroborated by pairwise F_{ST} comparisons, where three of the six populations were not significantly different from one another. This is noteworthy because the vast majority of pairwise comparisons, 481 out of 496, between all the populations were statistically different from one another. Interestingly, even though population level variation is lower within *D. fumella* than other species, *D. fumella* exhibits high divergence from all other *Dicerandra* specie. This is exemplified by the length of the branch in the distance trees leading to the *D. fumella* group (Figure 3-5).

While the degree of genetic differentiation between populations varies by species, the process driving these patterns is likely genetic drift. These species' annual habit, low dispersal capability, isolated populations, and visitation by generalist pollinators all contribute to reduce gene flow between populations. These life-history characteristics would readily allow individual populations to differentiate from one another. Additionally, this tendency for populations to be genetically divergent is accompanied by a reduction in heterozygosity. In 23 out of 32 populations, there was a significant deficiency in heterozygotes (Table 3-4). This pattern of population differentiation, along with a reduction in heterozygosity, is consistent with what would be expected from isolation and drift, potentially accelerated by fluctuations in population size, non-random mating, or consistently low population numbers. These factors eventually could lead to the observed reduction of heterozygosity. However, the obligate outcrossing nature of these plants has probably aided in maintaining heterozygosity despite drift and limited gene flow. Recruitment from the seed bank may also aid in maintaining heterozygosity

within these populations. Within the perennial species *D. frutescens*, reestablishment of populations solely from the seed bank occurs following fires which eliminate all individuals from the population (Menges 1992). This indicates that at least some *Dicerandra* species possess seeds capable of persisting for several seasons within the seed bank. This type of recruitment from the seed bank would produce populations with individuals from multi generations living at a given time. The presences of a multi-generational population would aid in maintaining heterozygosity in an annual outcrossing plant because individuals with previous years allele combinations would be present and capable of reproducing with more recent individuals.

Population Structure and Species Delimitation

Delimiting species boundaries requires adherence to a theoretical framework that provides both a sound description of what constitutes a species and evolutionarily meaningful evidence used to circumscribe populations as a single species. The unified species concept recognizes that most competing species concepts share the common feature of describing separately evolving metapopulation lineages, but they differ in the criteria employed to recognize those lineages or delimit metapopulations (De Queiroz 2007). While the difference among distinguishing criteria may seem to simply shift the name of the debate from species concepts to criteria, recognizing that the collective goal is to name populations of organisms that represent units moving through time provides a critical ontological foundation for recognizing and delimiting species. Under this species concept, many lines of evidence can be used to recognize metapopulations, such as reciprocal monophyly, diagnosability, reproductive isolation, apomorphic characters, or niche specialization (De Queiroz 2007). These are biologically and evolutionarily important characteristics that can delimit particular groups

of organisms as species, but they often are in conflict with one another for one reason: the speciation process occurs over a long timescale. These characteristics do not necessarily evolve in a consistent order, over a consistent time period, nor do they all manifest themselves in every taxon. They all capture important elements of the evolutionary process, which can be used as evidence for delimiting species, the difference being that each criterion recognizes a different aspect of the speciation process. Evolution is often a gradual process, and the stage of diversification at which an organism is observed will dictate the magnitude of divergence that has occurred. Under this approach, there is no single defining criterion, but the totality of the evidence is appreciated, recognizing the validity and biological significance of differing sources of data used to evaluate the question: "is this a metapopulation lineage?" In the case of *Dicerandra* four criteria best enable the evaluation of the degree of divergence and thus species boundaries: monophyly of populations in phylogenetic and distance based-trees, clustering patterns based on allele frequencies, gene flow, and morphological differentiation.

The STRUCTURE analysis and genetic distance trees that depict the magnitude of differentiation tell a congruent story regarding the relationships among species and among populations within species. Overall, most current taxonomic designations are well supported at their respective rank.

There is strong support for the recognition of *D. fumella* at the species level. The STRUCTURE analyses at both *K*=4 and *K*=5 are consistent in placing all populations of *D. fumella* within a single cluster. There was no evidence of admixture even with *D. fumella's* geographic neighbor, *D. linearifolia* var. *robustior* (Figure 3-5). Clustering of

all populations of *Dicerandra fumella* in the genetic distance trees further supports recognition at the species level. Additionally *D. fumella* has experienced substantial genetic differentiation from other *Dicerandra* species; this is depicted in the length of the branch leading to the *D. fumella* clade. The branch length is as great as or greater than the branches separating other *Dicerandra* species, i.e., *D. odoratissima* and *D. linearifolia s.l.* (Figure 3-3), which have long been recognized by systematists (Huck 1987). The genetic evidence presented here in conjunction with the morphological characteristics presented by Huck (2010) provide substantial support for the maintenance of *D. fumella* at the species rank.

Populations of *D. linearifolia s.l.* are supported at the species rank in both the STRUCTURE analysis and the genetic distance trees (Figures 3-3 & 3-5). In the distance trees, all populations of *D. linearifolia s.l.* form an exclusive cluster. Strong support is also seen from the clustering of these populations together in STRUCTURE at *K*=4, the value indicated by the Delta *K* method as the maximum number of statistically significant genetic clusters (Figure 3-5). There is also support at *K*=5 where the additional cluster results purely from the subdivision of the *D. linearifolia* cluster of *K*=4. This new cluster contains populations of *D. linearifolia* var. *robustior* and several populations of *D. linearifolia* var. *linearifolia* from central Georgia.

Taxonomic designations within *D. linearifolia s.l.* are complex. The varietal designation of *D. linearifolia* var. *robustior* is supported by morphology and in the distance trees where populations of this variety are consistently sister to each other (Figure 3-3). However, the recognition of var. *robustior* results in a paraphyletic *D. linearifolia* var. *linearifolia*. This is not necessarily of great concern, as the function of

varietal designations is to describe meaningful variation that is seen within a species, and monophyly in such cases should neither be assumed nor required. For *D. linearifolia* var. *robustior* the taxonomic designation as a variety, and not a species, best characterizes this group of populations for the following reasons: morphological characteristics are capable of differentiating var. *robustior* from other taxa of *Dicerandra,* including var. *linearifolia*, populations of var. *robustior* are more genetically similar to each other than they are to var. *linearifolia* but var. *robustior* is nested well within var. *linearifolia*, and populations of var. *robustior* are geographically disjunct from those of var. *linearifolia.*

Similarly, *D. linearifolia* var. *linearifolia* is best kept as a variety. Elevation of regional groups of var. *linearifolia* to species rank is not supported because there is not evidence for consistently cohesive geographic clusters within var. *linearifolia* (Figures 3-3 & 3-5). Nor is there a meaningful way to formally classify the exclusive groups that are observed in the distance trees without introducing an additional two taxonomic units, to designate the eastern, western, and central populations, which are not morphologically or genetically distinct.

Populations of *D. odoratissima* also did not for exclusive clusters in the distance trees and in the STRUCTURE analyses (Figures 3-3 & 3-5). Populations of *D. radfordiana* were nested among populations of *D. odoratissima* in the distance trees and *D. radfordiana* and proximate populations of *D. odoratissima* formed a distinct cluster in STRUCTURE analyses. This cluster (Figure 3-5, yellow) comprises populations at the southeastern extreme of the range of *D. odoratissima* and shows little admixture with more northern or western populations of *D. odoratissima*, except in

population 21 where there is an individual with a large contribution of its genotype coming from the other *D. odoratissima* cluster (Figure 3-5, purple). Population 21 marks one edge of the 50 km sampling gap separating these two clusters. This gap may contain further admixed populations suggesting clinal variation or the beginning of speciation by isolation at the extreme southeastern portion of the range, yet additional data are needed to address this hypothesis. Regardless of the sampling gap, the recognition of *D. radfordiana* at the species rank causes *D. odoratissima* to contain populations which are more genetically similar to a different species, *D. radfordiana*, which is in conflict with our species concept. This is further complicated because some *D. odoratissima* populations are more genetically similar to *D. radfordiana*, yet they are morphologically more similar to all other populations of *D. odoratissima*. There are two potential solutions to this problem, which are each considered below. 1) Recognize the genetically distinct populations, represented by the yellow cluster of the STRUCTURE analysis (Figure 3-5), as a single species, *D. radfordiana*. If this change were made, certain populations of *D. radfordiana s.l.* (placed in that species because of their genetic similarity) would be morphologically indistinguishable from populations of *D. odoratissima*. To a degree this defeats the purpose of a classification, because it makes some populations impossible to identify accurately without genetic analysis. Therefore, the recognition of *D. radfordiana* in an expanded circumscription, based on this genetic information, would only introduce confusion and erroneous identification of populations because the species would lack distinguishing morphological characteristics. Finally, the state of Georgia recognizes *D. radfordiana* as an endangered species, and continuing to recognize it as a species by widening the

circumscription to include genetically similar, but morphologically different, populations would only negatively impact its management, as one could no longer definitively recognize an individual of that species in the field. 2) Recognize the morphologically distinct populations of *radfordiana* as a variety of *D. odoratissima*. This approach recognizes the monophyly of *D. odoratissima*, maintains the morphological diagnosability of the *radfordiana* entity, allowing for consistent and positive identification of individuals, and continues support for the nomenclatural recognition of *radfordiana* populations as different from *D. odoratissima*. Maintaining the nomenclatural recognition at an infraspecific rank not only acknowledges its distinct morphology it also alludes to its differing evolutionary trajectory. This is significant because the genetic data suggest that these populations are at the beginning stages of speciation by isolation, and may evolve into separate species, if sufficient time elapses and the processes driving the divergence persist. However, currently the genetic differentiation is insufficient for recognition at the species level.

Within *D. linearifolia* var. *linearifolia* a surprising outcome from the population structure analyses was the placement of the far eastern population 22 within populations found exclusively in western Georgia (Figures 3-3 & 3-5). There are several biogeographic hypotheses that could have given rise to this pattern. The first is long-distance dispersal of individuals from the western populations to the far eastern edge of the range. However, this is implausible due to the dispersal limitations discussed earlier, the sheer geographic distance (260 km separate population 22 and the western populations), as well as geography (the two regions are located in different watersheds one draining to the Atlantic the other to the Gulf of Mexico). Even though

seed transport by water has been proposed as a significant influence in current species distributions, the probability of a long-distance dispersal event capable of establishing a viable population of annual plants appears small. A second scenario is the reverse of the first, i.e., one in which the western populations were derived from an eastern population, and while the western populations thrived and spread, their eastern relatives remained rare and local. But, this scenario is also improbable for the same reasons as the first. The third scenario is shaped by large-scale local extinction. Prior to the practice of widespread fire suppression and land conversion for anthropogenic activities, it is likely there was a broader distribution of suitable habitat with a greater frequency of *Dicerandra* populations. Clinal variation may still have existed with a predominantly western genotype and a predominantly eastern genotype with an area of admixture in the intermediate geographic region. Under this scenario, populations in the intermediate region could consist of populations dominated by one genotype or the other, or contain truly admixed individuals. Since dispersal is low, populations of a single genotype could easily be maintained genetically due to the plants annual habit; removal of alleles from migrants can be swift, after four generations of backcrossing an individual contains only 6.25% of the introduced alleles. If local extinction increased, eliminating the mosaic of eastern and western genotypes in the intermediate region, we would be left with a genotypic distribution that would look like clinal variation with an occasional outlier, which is precisely what is seen today (Fig 3-5, K=5). These populations would seem genetically out of place given the current distribution of genotypes, but they may represent relict populations from an era where intergradation of the two genotypes was more common. However, all three of these hypotheses

require additional testing in a phylogeographic framework to assess them more effectively.

Conservation Implications

Dicerandra, as an endemic genus to the southeast, is a valuable biologic and intrinsic component of this region. Given the narrow distribution of some of it's species and the isolated nature of may of it's populations it is important to consider the conservation implication of any research focusing on this genus. This study has addressed the population genetic aspect of these organisms' biology and has shown there is considerable diversity found at the population level for the species investigated. As with all conservation decisions it is important to consider multiple biologic aspects of an organism when deciding how we can best aid in the maintenance of these species.

All *Dicerandra* species rely on perturbation to maintain their habitat. Fieldwork conducted for this study found many populations along road-cuts, power line rights-of-way, and in recently disturbed natural areas with almost all populations adjacent to later successional closed canopy forest; testament to these species requirement for disturbance. Even within the widely distributed species of *D. linearifolia* var. *linearifolia*, *D. fumella*, and *D. odoratissima* the greatest threat to the long-term persistence of populations is habitat loss. This loss most often appears to be the result of habitat succession resulting from fire suppression or wholesale land conversion for anthropogenic purposes (typically silviculture or agriculture). Research by Menges (1992) has shown the rapid response of populations following disturbance and allowing natural processes, such as periodic fire, to remain within ecosystems is a critical component for the long-term persistence of these species.

Within the annuals examined for this study, *D. linearifolia* var. *linearifolia*, *D. fumella*, and *D. odoratissima* have the largest distribution with many populations containing 100's-1,000's of individuals. These populations are likely stable and will persist if suitable habitat remains available. While each species has a wide distribution certain areas have a greater presence of populations. For *D. linearifolia* var. *linearifolia* the greatest density of populations occur in western Georgia, near the upper Flint River, centered in Talbot and Taylor counties. *Dicerandra linearifolia* var. *robustior* is common in the eastern portion of the Florida panhandle between the Ochlockonee and Suwannee Rivers. *Dicerandra fumella* is widespread in the western portion of the Florida panhandle with populations commonly found in Okaloosa, Walton, Holmes, and Washington counties east of the Escambia Rive and surrounding the Choctawhatchee River. *Dicerandra* odoratissima populations are common in eastern Georgia along the Altamaha River north to Emanuel county. These regions represent the core populations for these species where efforts to maintain habitat would provide maximum benefit in terms of conserving the greatest quantity of populations and thus genetic diversity.

Alternatively *D. radfordiana* has a very narrow distribution only occurring in McIntosh county, Georgia and is currently listed by the state of Georgia as an endangered species. This study found the populations of *D. radfordiana* are not genetically distinct from proximate populations of *D. odoratissima* despite showing morphological differentiation. We presented two approaches for addressing the taxonomy of *radfordiana*. Each approach has distinct conservation implications; either *radfordiana* and the proximate populations of *D. odoratissima* are recognized as *D. radfordiana* (based on their genetic similarity (Figure 3-5)) but would lack morphological

characters to distinguish populations from one another making identification in the field impossible or *D. radfordiana* is recognized as a variety of *D. odoratissima*, thus removing its species rank and any conservation protections this rank carries, but would continue to recognize it's unique morphology and evolutionary trajectory.

It can often be argued when addressing designation and conservation of threatened or endangered species that the potential damage associated with a type II error (accepting the null of no difference where a difference actually exists) is significantly greater than a type I error (rejecting the null when there really is no difference) (Dayton 1998; McGarvey 2007). However, there also exists a delicate balance between available resources for conservation and the degree of need of each species. As funding for conservation is typically limited, it is prudent to apply those resources to species where the benefit will be greatest. Part of determining which species and projects will be most beneficial is identifying previously unrecognized or under-studied species that truly merit species status or concerted efforts to aid their survival. Alternatively, understanding which species do not require conservation effort is equally as important to ensure funding is only directed toward organisms that truly need it. For example, understanding the biological requirements of these species, determining conservation actions, and establishing priority among species represent significant investments of resources.

Dicerandra radfordiana has been the focus of recent conservation efforts by the Georgian Division of Natural Resources to open habitat surrounding certain populations and has benefited from the corporation of the Nature Conservancy and the Georgia Rare Plant Alliance in assessing population trends and identifying the extent of it's

range. These efforts have provided invaluable benefits to aid in the persistence of *D. radfordiana*. Regardless of the taxonomic designation of *radfordiana* these kinds of investments are critical to maintaining the unique variation that resulted in the initial recognition of *radfordiana* populations as being unique from other populations of *Dicerandra* and any taxonomic revision should not be interpreted as diminishing the evolutionary significance of these populations.

Future Research

This study has made significant advances by increasing the quantity of genetic resources available for this genus and furthering the understanding of genetic distribution and relatedness of species within the annual clade of *Dicerandra*. Yet, additional questions still remain unanswered. The population-level sampling for this study sets the stage for further more detailed work toward understanding the mechanisms responsible for the distributions of these species. With additional sequence data, the current geographic sampling could easily be used in a phylogeographic study to address the pattern of expansion within *D. linearifolia* var. *linearifolia* in Georgia. Similar questions could be addressed in connection with the relationship and pattern of expansion between *D. fumella* and *D. linearifolia* var. *robustior* in Florida and *D. linearifolia* var. *linearifolia* Georgia. Additional sampling in a few key geographic areas such as central Georgia would help evaluate the strength of the eastern *D. linearifolia* var. *linearifolia* populations as a cohesive unit and provide additional insight into the potential role local extinction, land use changes, or habitat reduction have played in the current geographic distribution of this taxon. These types of analyses could be extended to include the perennial species of peninsular Florida to provide a more complete understanding of how these species have radiated and

evolved, from their hypothesized glacial refugium in peninsular Florida, following the end of the last glacial maximum. With the addition of a relatively small number of samples, the genus could provide a unique opportunity to investigate the dynamics associated with the climate shift following glacial retreat and the subsequent radiation of a genus of plants endemic to the southeastern United States.

While there is substantial genetic evidence for the recognition of *D. fumella* as a species, presented in this study as well as in the phylogenetic analysis of Oliveira et al. (2007), the morphological basis for species identification and circumscription would significantly benefit from a more detailed analysis. Variation among the populations of *D. fumella* is substantial, as noted by Huck (2010), so much so that she proposed a hybrid zone with *D. linearifolia* var. *robustior*, yet hybridization was not supported by genetic data in this study. Several factors could drive this morphological variation: drift, phenotypic plasticity, environmental variation in resource availability, or local adaptation. Further studies with ordination analyses based on detailed morphological comparisons, reciprocal transplant or common garden experiments, and microsite comparisons could help advance our understanding of the genetic and ecological roles associated with the observed phenotypic variation. These types of studies could also be used to assess the rather minimal pattern of morphological differentiation between the western and central Georgia genotypes of *D. linearifolia* var. *linearifolia*.

Conclusions

This study aimed to investigate phylogenetic incongruences observed by Oliveira et al. (2007) between multiple accessions of the same species of *Dicerandra* using a population level approach. Each of the following questions were proposed at the onset of the study.

Within *Dicerandra linearifolia s.l.*: 1) Did population structure mirror the east-west structure observed in the phylogeny? We found there was general geographic cohesions of Georgia populations of *D. linearifolia* var. *linearifolia*. However, populations did not form exclusive clusters, as illustrated by STRUCTURE analyses and distance based trees. Central Georgia populations showed admixture of both eastern and western genotypes and population 22, found at the eastern edge of the range was composed exclusively of individuals with western genotypes. 2) Is there molecular evidence for hybrids between *D. linearifiolia* var. *linearifolia* and *D. odoratissima*? Putative hybrid individuals identified in the field did carry a mixed genotypic signature, however not all putative hybrids showed this. Some hybrids displayed genotypes that were almost exclusively *D. linearifolia* var. *linearifolia* or *D. odoratissima* despite showing intermediate or chimeric morphology in the field. Additionally hybridization may be directional, as seen by the presences of individuals with mixed genotypes within a *D. odoratissima* population but the absence of genotypically mixed individuals within the *D. linearifolia* var. *linearifolia* population separated by only 30 m. 3) What is the extent of polyploidy within the genus? All annual species of *Dicerandra* have a ploidal level of 2*n*=32. While this number indicates this genus is the result of historic genome duplication in an ancestral lineage there was no evidence of recent polyploidy within the extant members of this genus. 4) Is there molecular support for the recently described species *D. fumella*? There was strong support for the recognition of *D. fumella* at the species rank. Populations of *D. fumella* formed exclusive clusters in STRUCTURE analyses and within the distance based trees. Additionally, *D. fumella* had considerable genetic differentiation from all other *Dicerandra* species investigated. 5) Is there

support of a hybridization zone between *D. fumella* and *D. linearifolia* var. *robustior*? While there is considerable morphological variation at the eastern edge of *D. fumella's* range giving the appearance of a potential hybridization zone with neighboring *D. linearifolia* var. *linearifolia* there was no genetic support for this hypothesis.

Within *D. odoratissima* and *D. radfordiana*: 1) Is there significant genetic structure separating populations of *D. odoratissima* and *D. radfordiana*? The two *D. radfordiana* populations were found to be genetically distinct from most but not all populations of *D. odoratissima*. Two populations of *D. odoratissima* that occur within 11 km of *D. radfordiana* consistently clustered with both *D. radfordiana* populations in the STRUCTURE analyses and the distance trees. While the two *D. radfordiana* populations and the two proximate *D. odoratissima* populations are genetically distinct from the remaining *D. odoratissima* populations, morphologically these two *D. odoratissima* populations are indistinguishable from the remaining populations of *D. odoratissima*. It is likely these are in the process of diverging from the remaining *D. odoratissima* populations; however, insufficient time has elapsed for them to be both genetically and morphological distinct. 2) Does geography play a significant role in explaining the observed non-monophyly of *D. odoratissima*? Mantel tests showed there was a significant correlation between geographic distance and genetic distance within *D. odoratissima* and *D. radfordiana*. The *D. radfordiana* populations and their proximate *D. odoratissima* populations are found at the far southeastern edge of *D. odoratissima's* range and may represent the beginning of speciation by isolation.

Overall, there is significant genetic diversity found at the populations level for all *Dicerandra* annuals, 11.39% of the total genetic diversity is found among populations

within species. This value is similar to the proportion of genetic variation found between species, 12.49%. The isolated nature of many populations, limited dispersal capability, and frequent visitation from generalist pollinators produce conditions where drift can cause significant differentiation between populations, even within the same geographic area. However the obligate outcrossing nature of these plants and the potential for multi-generational populations resulting from recruitment from the seed bank are likely stabilizing factors maintaining heterozygosity within populations.

With the exception of *D. radfordiana* all the annual species are widely distributed within their respective ranges and often populations containing 100's-1,000's of individuals exist. Disturbance is required for the maintenance of their habitat and allowing for continued natural disturbance, such as periodic fire, it is likely many of these populations will persist well into the future.

APPENDIX
PAIRWISE POPULATION COMPARISONS

Table A-1. Pairwise F_{ST} values. Asterisks indicate non-significant Bonferroni corrected p-values (p>0.05). Column and row headers are color-coded by species: Blue: *D. linearifolia* var. *linearifolia*, Yellow: *D. radfordiana*, Purple: *D. odoratissima*, Red: *D. fumella*, Green: *D. linearifolia* var. *robustior*, White: hybrid populations.

	LIN_2	LIN_3	RAD_4	ODO_9	LIN_10	ODO_1	ODO_1	ODO_1	ODO_1	ODO_1	FUM_18	ODO_2
LIN_3	0.0475											
RAD_4	0.2475	0.3164										
ODO_9	0.2569	0.3131	0.3345									
LIN_10	0.2307	0.2765	0.3195	0.3445								
ODO_1	0.2283	0.2965	0.3113	0.2074	0.3301							
ODO_1	0.1799	0.1931	0.1997	0.2737	0.3076	0.1768						
ODO_1	0.1654	0.2322	0.3039	0.1843	0.2754	0.1788	0.1918					
ODO_1	0.1359	0.2098	0.2192	0.1460	0.2497	0.1672	0.1854	0.0561				
ODO_1	0.1500	0.2144	0.2370	0.1755	0.2428	0.1553	0.2057	0.0542	0.0481*			
FUM_18	0.1991	0.2548	0.2932	0.2573	0.2831	0.2477	0.2699	0.2221	0.1881	0.2049		
ODO_2	0.3066	0.3587	0.1815	0.3140	0.3804	0.3311	0.3114	0.2908	0.2290	0.2444	0.3086	
ODO_2	0.2062	0.2628	0.1268	0.2180	0.2591	0.1967	0.2214	0.1775	0.1695	0.1670	0.2277	0.0797
LIN_22	0.0896	0.0834	0.2498	0.2227	0.1930	0.2136	0.1410	0.1658	0.1263	0.1318	0.2019	0.2774
HYB_23	0.0997	0.1475	0.1997	0.1629	0.1417	0.1176	0.1671	0.1084	0.1114	0.1350	0.1653	0.2215
LIN_24	0.1310	0.1774	0.2056	0.1521	0.1310	0.1648	0.1806	0.1535	0.1515	0.1550	0.1681	0.2389
HYB_25	0.0772	0.1137	0.2106	0.1764	0.1331	0.1652	0.1553	0.0968	0.0916	0.1264	0.1685	0.2691
LIN_27	*0.0664	0.0685	0.3149	0.2719	0.2430	0.2530	0.1848	0.2564	0.1548	0.1738	0.2161	0.3235
LIN_28	0.1316	0.1150	0.2785	0.2885	0.2687	0.2662	0.1992	0.1476	0.2081	0.1988	0.2595	0.3259
LIN_29	0.0673	0.1209	0.1922	0.2173	0.1597	0.1891	0.1361	0.1417	0.1314	0.1428	0.1721	0.2436
ODO_3	0.2610	0.3374	0.3839	0.2164	0.3759	0.2236	0.2934	0.2266	0.1983	0.1904	0.3076	0.3815
ODO_3	0.1775	0.2422	0.2692	0.1830	0.2625	0.2002	0.2201	0.1371	0.1136	0.1155	0.2348	0.2710
LIN_32	0.0836	0.1518	0.2150	0.1854	0.2088	0.1817	0.1612	0.1097	0.0964	0.0948	0.1473	0.2594
LIN_33	*0.0846	0.1439	0.2650	0.1565	0.2139	0.1621	0.1553	0.1715	*0.0688	0.1155	0.1311	0.2614
FUM_34	0.1863	0.2331	0.2791	0.2708	0.2610	0.2548	0.2375	0.2152	0.2007	0.2113	0.0507	0.3176
FUM_35	0.2899	0.3110	0.3897	0.3943	0.3876	0.3766	0.3134	0.3421	0.3271	0.3238	0.1604	0.4283
FUM_36	0.1818	0.2194	0.2823	0.2871	0.2734	0.2603	0.2555	0.2159	0.2202	0.2111	0.0952	0.3079
FUM_37	0.2981	0.3392	0.3714	0.3756	0.3442	0.3872	0.3617	0.3416	0.3098	0.3046	0.1804	0.3894
FUM_38	0.2054	0.2617	0.3024	0.2629	0.2912	0.2531	0.2625	0.2221	0.2021	0.2066	*0.0514	0.3059
ROB_40	0.2676	0.3229	0.4280	0.3482	0.3782	0.4187	0.3606	0.3288	0.2426	0.2803	0.3091	0.4509
ROB_41	0.1168	0.1497	0.3046	0.2632	0.2033	0.3036	0.2317	0.2260	0.1621	0.1865	0.1918	0.3455
RAD_42	0.3805	0.4333	0.1834	0.4410	0.4696	0.4311	0.3512	0.3798	0.3499	0.3299	0.4042	0.2258

Table A-1. Continued

	ODO_21	LIN_22	HYB_23	LIN_24	HYB_25	LIN_27	LIN_28	LIN_29	ODO_30	ODO_31	LIN_32	LIN_33
LIN_2												
LIN_3												
RAD_4												
ODO_9												
LIN_10												
ODO_13												
ODO_14												
ODO_15												
ODO_16												
ODO_17												
FUM_18												
ODO_20												
ODO_21												
LIN_22	0.1779											
HYB_23	0.1158	0.0914										
LIN_24	0.1385	0.0921	0.0531									
HYB_25	0.1707	0.0668	0.0435	0.0821								
LIN_27	0.2196	0.0883	0.0919	0.1163	0.1089							
LIN_28	0.2244	0.1228	0.1233	0.1688	0.1410	0.0836						
LIN_29	0.1593	0.0802	0.1007	0.0905	0.0899	0.1010	0.1272					
ODO_30	0.2413	0.2558	0.1769	0.2149	0.2204	0.2917	0.2930	0.2309				
ODO_31	0.1511	0.1654	0.1014	0.1359	0.1171	0.2109	0.2323	0.1484	0.1700			
LIN_32	0.1610	0.0602	0.0860	0.0792	0.0717	0.0974	0.1455	0.0807	0.2399	0.1322		
LIN_33	0.1568	0.0995	*0.0571	0.0709	*0.0616	0.0933	0.1528	0.0888	0.2302	0.1603	0.0668	
FUM_34	0.2171	0.1885	0.1456	0.1507	0.1417	0.2129	0.2267	0.1641	0.3167	0.2327	0.1445	0.1628
FUM_35	0.3296	0.3047	0.2726	0.2891	0.2742	0.3187	0.3011	0.2565	0.4518	0.3599	0.2802	0.2764
FUM_36	0.2253	0.1869	0.1723	0.1754	0.1682	0.2063	0.2247	0.1658	0.3119	0.2286	0.1543	0.1660
FUM_37	0.2979	0.2856	0.2304	0.2356	0.2469	0.3067	0.3273	0.2595	0.4230	0.3245	0.2460	0.2666
FUM_38	0.2184	0.1993	0.1688	0.1591	0.1799	0.2048	0.2346	0.1637	0.3145	0.2312	0.1400	0.1380
ROB_40	0.3283	0.2497	0.2355	0.2442	0.2228	0.2901	0.3138	0.2441	0.4446	0.3408	0.2329	0.2293
ROB_41	0.2155	0.0946	0.0958	*0.0677	0.0948	0.1121	0.1806	0.1153	0.3155	0.1908	*0.0843	*0.0899
RAD_42	0.2453	0.3417	0.3267	0.3454	0.3612	0.4078	0.3823	0.3080	0.5031	0.4085	0.3533	0.3866

Table A-1. Continued

	FUM_34	FUM_35	FUM_36	FUM_37	FUM_38	ROB_40	ROB_41	RAD_42
LIN_2								
LIN_3								
RAD_4								
ODO_9								
LIN_10								
ODO_13								
ODO_14								
ODO_15								
ODO_16								
ODO_17								
FUM_18								
ODO_20								
ODO_21								
LIN_22								
HYB_23								
LIN_24								
HYB_25								
LIN_27								
LIN_28								
LIN_29								
ODO_30								
ODO_31								
LIN_32								
LIN_33								
FUM_34								
FUM_35	0.1007							
FUM_36	0.0221	0.1077						
FUM_37	0.0976	0.1759	0.0913					
FUM_38	0.0294	0.1215	*0.0186	0.1043				
ROB_40	0.3294	0.4474	0.3179	0.4602	0.3431			
ROB_41	0.1756	0.3125	0.1763	0.2872	0.1979	*0.1922		
RAD_42	0.3975	0.4914	0.4020	0.5085	0.4261	0.5477	0.4482	

Table A-2. Pairwise comparisons of Jost's *D*. Column and row headers are color-coded by species: Blue: *D. linearifolia* var. *linearifolia*, Yellow: *D. radfordiana*, Purple: *D. odoratissima*, Red: *D. fumella*, Green: *D. linearifolia* var. *robustior*, White: hybrid populations.

	LIN_2	LIN_3	RAD_4	ODO_9	LIN_10	ODO_13	ODO_14	ODO_15	ODO_16	ODO_17	FUM_18	ODO_20
LIN_3	0.1167											
RAD_4	0.8440	0.8966										
ODO_9	0.8236	0.8546	0.9264									
LIN_10	0.6879	0.6980	0.8439	0.9261								
ODO_13	0.8012	0.8585	0.8739	0.4827	0.9232							
ODO_14	0.5613	0.4848	0.5036	0.7254	0.8399	0.4800						
ODO_15	0.6591	0.6490	0.8214	0.6192	0.7597	0.6150	0.6679					
ODO_16	0.5276	0.6918	0.7637	0.4240	0.8136	0.5771	0.5991	0.4189				
ODO_17	0.5970	0.6922	0.7977	0.5078	0.7622	0.5248	0.6706	0.3088	0.1670			
FUM_18	0.7600	0.8031	0.9801	0.7889	0.8711	0.8577	0.8906	0.8915	0.7318	0.8008		
ODO_20	0.9188	0.9311	0.3578	0.7054	0.9508	0.7811	0.7622	0.7948	0.6478	0.6757	0.8850	
ODO_21	0.8330	0.8586	0.4099	0.6352	0.7946	0.6421	0.6956	0.6762	0.6698	0.6266	0.8681	0.1673
LIN_22	0.3268	0.2177	0.7440	0.7070	0.5744	0.7749	0.4278	0.5837	0.5290	0.5242	0.7972	0.8030
HYB_23	0.4513	0.5241	0.8250	0.5658	0.4837	0.5284	0.6140	0.5804	0.5264	0.6327	0.7404	0.7222
LIN_24	0.5984	0.6320	0.8019	0.5191	0.4240	0.6906	0.6488	0.7551	0.7252	0.7205	0.7237	0.7596
HYB_25	0.3278	0.3668	0.8474	0.5973	0.4318	0.6937	0.5400	0.5131	0.4142	0.5868	0.7488	0.9121
LIN_27	0.2321	0.1770	0.9148	0.8102	0.6797	0.7951	0.5545	0.7823	0.6104	0.6140	0.7963	0.8595
LIN_28	0.4303	0.2853	0.8600	0.8589	0.7616	0.8617	0.5563	0.5086	0.7712	0.7088	0.9267	0.9116
LIN_29	0.2889	0.4255	0.7728	0.8365	0.5624	0.8602	0.4846	0.7800	0.6525	0.7114	0.7966	0.8400
ODO_30	0.7327	0.8234	0.9072	0.3932	0.8734	0.4122	0.6984	0.5549	0.5291	0.4832	0.8706	0.7799
ODO_31	0.6555	0.7147	0.7902	0.4616	0.7256	0.5543	0.6573	0.4915	0.3705	0.3810	0.8624	0.6607
LIN_32	0.4160	0.5584	0.8534	0.6371	0.7498	0.7453	0.6032	0.6637	0.4793	0.4699	0.6907	0.8729
LIN_33	0.3887	0.4824	0.9117	0.4589	0.7187	0.5268	0.5330	0.7215	0.3239	0.5032	0.5577	0.7186
FUM_34	0.7094	0.7076	0.8532	0.8307	0.7638	0.8447	0.7517	0.7450	0.8025	0.8328	0.1558	0.8912
FUM_35	0.7277	0.6554	0.8937	0.9102	0.8541	0.8574	0.6863	0.8106	0.9140	0.8839	0.3194	0.9284
FUM_36	0.6697	0.6446	0.9169	0.8932	0.8161	0.8872	0.8055	0.7677	0.8704	0.8061	0.2785	0.8745
FUM_37	0.8537	0.8134	0.8695	0.9088	0.7541	0.9554	0.9349	0.9395	0.9448	0.9034	0.4126	0.8393
FUM_38	0.8342	0.8388	0.9754	0.8029	0.8891	0.8260	0.8688	0.8745	0.8366	0.8367	0.1403	0.8419
ROB_40	0.6658	0.6806	0.9694	0.6817	0.7778	0.8963	0.8497	0.7206	0.6105	0.7404	0.8025	0.9653
ROB_41	0.3710	0.3675	0.9027	0.7383	0.4915	0.9364	0.6825	0.6533	0.5896	0.6867	0.6494	0.9315
RAD_42	0.8847	0.9274	0.2307	0.9170	0.9841	0.8573	0.6746	0.8552	0.8356	0.7561	0.9734	0.3001

Table A-2. Continued

	ODO_21	LIN_22	HYB_23	LIN_24	HYB_25	LIN_27	LIN_28	LIN_29	ODO_30	ODO_31	LIN_32	LIN_33
LIN_2												
LIN_3												
RAD_4												
ODO_9												
LIN_10												
ODO_13												
ODO_14												
ODO_15												
ODO_16												
ODO_17												
FUM_18												
ODO_20												
ODO_21												
LIN_22	0.6981											
HYB_23	0.4938	0.4263										
LIN_24	0.5888	0.4230	0.2496									
HYB_25	0.7840	0.2934	0.2081	0.4164								
LIN_27	0.7808	0.2838	0.3965	0.5103	0.4638							
LIN_28	0.7891	0.4023	0.4767	0.6643	0.5465	0.2473						
LIN_29	0.7567	0.3755	0.5742	0.4813	0.5011	0.4803	0.5016					
ODO_30	0.6223	0.7242	0.5534	0.6751	0.7032	0.7417	0.7765	0.7966				
ODO_31	0.5060	0.6055	0.4169	0.5774	0.4809	0.6967	0.7909	0.6867	0.3397			
LIN_32	0.7962	0.3274	0.5318	0.4800	0.4291	0.4744	0.6179	0.5300	0.7730	0.6111		
LIN_33	0.6435	0.4456	0.3021	0.3853	0.3597	0.2864	0.6174	0.5576	0.6206	0.6396	0.4947	
FUM_34	0.8127	0.7196	0.6640	0.6641	0.6350	0.7449	0.7770	0.7815	0.8635	0.8236	0.7012	0.6997
FUM_35	0.8669	0.7987	0.8233	0.8516	0.7959	0.7685	0.7039	0.7602	0.9539	0.8822	0.8065	0.7376
FUM_36	0.8310	0.7006	0.7460	0.7344	0.7150	0.7175	0.7460	0.7311	0.8756	0.8140	0.7030	0.7169
FUM_37	0.8362	0.8258	0.7396	0.7365	0.7823	0.8003	0.8798	0.8872	0.8945	0.8462	0.7935	0.8144
FUM_38	0.8431	0.8152	0.7973	0.7173	0.8534	0.7388	0.8324	0.8011	0.8683	0.8648	0.7037	0.6165
ROB_40	0.8854	0.6058	0.7053	0.7272	0.6265	0.6654	0.7735	0.7753	0.8006	0.7796	0.6355	0.5962
ROB_41	0.7792	0.2867	0.3638	0.2373	0.3588	0.3094	0.5561	0.4877	0.7772	0.6123	0.3667	0.3655
RAD_42	0.4759	0.7696	0.8965	0.9128	0.9789	0.9401	0.8423	0.8075	0.9299	0.8615	0.8854	0.9545

Table A-2. Continued

	FUM_34	FUM_35	FUM_36	FUM_37	FUM_38	ROB_40	ROB_41	RAD_42
LIN_2								
LIN_3								
RAD_4								
ODO_9								
LIN_10								
ODO_13								
ODO_14								
ODO_15								
ODO_16								
ODO_17								
FUM_18								
ODO_20								
ODO_21								
LIN_22								
HYB_23								
LIN_24								
HYB_25								
LIN_27								
LIN_28								
LIN_29								
ODO_30								
ODO_31								
LIN_32								
LIN_33								
FUM_34								
FUM_35	0.1727							
FUM_36	0.0654	0.2071						
FUM_37	0.1897	0.2513	0.1973					
FUM_38	0.0873	0.2253	0.0467	0.2202				
ROB_40	0.8237	0.8335	0.8254	0.9163	0.9086			
ROB_41	0.5751	0.6865	0.5738	0.6907	0.7135	0.3159		
RAD_42	0.9050	0.8804	0.9514	0.9567	0.9877	0.9593	0.9525	

LIST OF REFERENCES

Abdelkrim J, Robertson BC, Stanton JL, Gemmel NJ (2009) Fast, cost-effective development of species-specific microsatellite markers by genomic sequencing. *BioTechniques*, **46**, 185–192.

Allentoft ME, Schuster SC, Holdaway RN *et al.* (2009) Identification of microsatellites from an extinct moa species using high-throughput (454) sequence data. *BioTechniques*, **46**, 195–200.

Bentham G (1848) *Dicerandra*. *DC Proceedings*, **12**, 242-243.

Castoe T, Poole A, GU W, *et al.* (2010) Rapid identification of thousands of copperhead snake (*Agkistrodon contortrix*) microsatellite loci from modest amounts of 454 shotgun genome sequence. *Molecular Ecology Resources*, **10**, 341–347.

Comes H, Abbott R (2001) Molecular phylogeography, reticulation, and lineage sorting in Mediterranean *Senecio* sect. *Senecio* (Asteraceae). *Evolution*, **55**, 1943–1962.

Dayton P (1998) Reversal of the Burden of Proof in Fisheries Management. *Science*, **279**, 821–822.

De Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology*, **56**, 879-886.

Doyle JJ, Doyle JL (1987) A rapid isolation procedure for small quantities of fresh tissue. *Phytochemical Bulletin,* **19**, 11-15.

Earl D, VonHoldt B (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359-361.

Edwards C, Solits D, Soltis P (2008) Using patterns of genetic structure based on microsatellite loci to test hypotheses of current hybridization, ancient hybridization and incomplete lineage sorting in *Conradina* (Lamiaceae). *Molecular Ecology*, **17**, 5157-5174.

Estill J, Cruzan M (2001) Phytogeography of rare plant species endemic to the southeastern United States. *Castanea*, **66**, 3–23.

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

Evans M, Menges E, Gordon D (2004) Mating systems and limits to seed production in two *Dicerandra* mints endemic to Florida scrub. *Biodiversity and Conservation*, **13** 1819-1832.

Excoffier L, Lischer H (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564-567.

Florida Natural Areas Inventory (FNAI) (2010) Guide to the natural communities of Florida. 2010 edition, pp. 32-44. Florida Natural Areas Inventory, Tallahassee, Florida.

Gordon A (2008) Fastx-toolkit. http://hannonlab.cshl.edu/fastx_toolkit/index.html. Last accessed July 2012.

Greilhuber J, Temsch E, Loureiro J (2007) Nuclear DNA content measurements. In: Flow cytometry with plant cells: analysis of genes, chromosomes and genomes (ed. Dolezel and Suda), pp. 67-99. Wiley-VCH, Weinheim, Germany.

Huck R (1987) Systematics and evolution of *Dicerandra* (Labiatae). *Phanerogamarum monographiae*, **19**, 1-343.

Huck R, Chambers H (1997) Polyploidy: a factor in the evolution of *Dicerandra* Benth.(Labiatae). *Edinburgh Journal of Botany*, **54**, 217-229.

Huck R, Judd W, Whitten W, *et al.* (1989) A new *Dicerandra* (Labiatae) from the Lake Wales Ridge of Florida, with a cladistic analysis and discussion of endemism. *Systematic Botany*, **14**, 197-212.

Huck, R (2010) *Dicerandra fumella* (Lamiaceae), a new species in the Florida panhandle and adjacent Alabama, with comments on the *D. linearifolia* complex. *Rhodora*, **112**, 215-227.

Jakob S, Blattner F (2006) A chloroplast genealogy of *Hordeum* (Poaceae): Long-term persisting haplotypes, incomplete lineage sorting, regional extinction, and the consequences for phylogenetic inference. *Molecular Biology and Evolution*, **23**, 1602–1612.

Jakobsson M, Rosenberg N (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, McLenachan P, Lockhart P (2009) A statistical approach for distinguishing hybridization and incomplete lineage. *American Naturalist*, **174**, 54-70.

Jost L (2008) GST and its relatives do not measure differentiation. *Molecular Ecology*, **17**, 4015-4026.

Kral R (1982) Some notes on *Dicerandra* (Lamiaceae). *SIDA*, **9**, 238-262.

Levitt R, Kiser M, Dragwa C, *et al.* (1994) Fluorescence-based resource for semiautomated genomic analyses using microsatellite markers. *Genomics*, **24**, 361–365.

McCormick K, Deyrup M, Menges E *et al.* (1993) Relevance of chemistry to conservation of isolated populations: the case of volatile leaf components of Dicerandra mints. *Proceedings of the National Academy of Science*, **90**, 7701-7705.

McDonald D, Hamrick J (1996) Genetic variation in some plants of Florida scrub. *American Journal of Botany*, **83**, 21–27.

McGarvey D (2007) Merging precaution with sound science under the Endangered Species Act. *BioScience*, **57**, 65–70.

Meirmans P, Van Tiederen P (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792-794.

Meirmans P, Hedrick P (2011) Assessing population structure: FST and related measures. *Molecular Ecology Resources*, **11**, 5–18.

Menges E (1992) Habitat preferences and response to disturbance for *Dicerandra frutescens*, a Lake Wales Ridge (Florida) endemic plant. *Bulletin of the Torrey Botanical Club*, **119**, 308-313.

Menges E, Dolan R, Yahr R, Gordon D (2001) Comparative genetics of seven plants endemic to Florida's Lake Wales Ridge. *Castanea*, **66**, 98-114.

Menges E, McIntyre P, Finer M, *et al.* (1999) Microhabitat of the narrow Florida scrub endemic *Dicerandra christmanii*, with comparisons to its congener *D. frutescens*. *Journal of the Torrey Botanical Society*, **126**, 24–31.

Menges ES, Dolan RW, Pickert R, *et al.* (2010) Genetic Variation in Past and Current Landscapes: Conservation Implications Based on Six Endemic Florida Scrub Plants. *International Journal of Ecology*, **2010**, 1–13.

Merchant A (2006) Variation in functional morphology and physiological responses of *Dicerandra* (Lamiaceae) congeners native to sandhill habitats and Florida scrub. Doctoral dissertation, University of Florida, Gainesville, Florida.

Myers R, Ewel J (1990) In: Ecosystems of Florida (ed. Myers and Ewel), pp. 35-69, 150-193. University Presses of Florida, Gainesville, Florida.

Oliveira L, Huck R, Gitzendanner M, *et al.* (2007) Molecular phylogeny, biogeography, and systematics of *Dicerandra* (Lamiaceae), a genus endemic to the southeastern United States. *American Journal of Botany*, **94**, 1017-1027.

Peakal R, Smouse P (2006) GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*, **6**, 288-295.

Ramdhani S, Barker N, Cowling R (2011) Revisiting monophyly in *Haworthia* Duval (Asphodelaceae): Incongruence, hybridization and contemporary speciation. *Taxon*, **60**, 1001-1014.

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

Rice W (1989) The sequential Bonferroni test. *Evolution*, **43**, 223-225.

Rieseberg L, Soltis D (1991) Phylogenetic consequences of cytoplasmic gene flow in plants. *Evolutionary Trends in Plants*, **5**, 65–84.

Rosenberg N (2004) Distruct: a program for the graphical display of population structure *Molecular Ecology Notes*, **4**, 137-138.

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

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. In: Bioinformatics Methods and Protocols: Methods in Molecular Biology. (ed. Krawetz S, Misener S), pp. 365-386. Humana Press, Totowa, New Jersey.

Schuelke M (2000) An economic method for the fluorescent labeling of PCR fragments. *Nature biotechnology*, **18**, 223-224.

Shinners L (1962) Synopsis of Dicerandra (Labiatae). *SIDA*, **1**, 89-91.

Suda J, Trávnícek P (2006) Reliable DNA ploidy determination in dehydrated tissues of vascular plants by DAPI flow cytometry--new prospects for plant research. Cytometry. *The journal of the International Society for Analytical Cytology*, **69**, 273–280.

Sutton J, Robertson B, Jamieson I (2011) Dye shift: a neglected source of genotyping error in molecular ecology. *Molecular Ecology Resources*, **11**, 514–520.

Swofford D (2003) PAUP* Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, Massachusetts.

Syring J, Farrell K, Businsky R, *et al.* (2007) Widespread genealogical nonmonophyly in species of *Pinus* subgenus *Strobus*. *Systematic Biology*, **56**, 163–181.

USFWS (1993) Endangered and threatened wildlife and plants; review of plant taxa for listing as endangered or threatened species. *Federal Register*, **58**, 51144-51190.

USFWS (2007) Endangered and threatened wildlife and plants; 5-year review of 22 southeastern species. *Federal Register*, **72**, 20866-20868.

Wahlund S (1928) Composition of populations and correlation appearances view in relation to the studies of inheritance. *Hereditas*, **11**, 65-106.

Wendel J, Doyle J (1998) Phylogenetic incongruence: window into genome history and molecular evolution. In: Molecular Systematics of Plants II (ed. Soltis, Soltis, and Doyle), pp. 265-296. Kluwer Academic Publishers, Norwell, Massachusetts.

Wheeler Q, Meier R (2000) In: Species Concepts and Phylogenetic Theory (ed. Wheeler and Meier), pp. 1-208. Columbia University Press, West Sussex, New York.

Winn A (1996a) Adaptation to fine-grained environmental variation: an analysis of within-individual leaf variation in an annual plant. *Evolution*, **50**, 1111–1118.

Winn A (1996b) The contributions of programmed developmental change and phenotypic plasticity to within-individual variation in leaf traits in *Dicerandra linearifolia*. *Journal of Evolutionary Biology*, **9**, 737-752.

Winn A (1999) Is seasonal variation in leaf traits adaptive for the annual plant *Dicerandra linearifolia*? *Journal of Evolutionary Biology*, **12**, 306-313.

Yu Y, Than C, Degnan JH, Nakhleh L (2011) Coalescent histories on phylogenetic networks and detection of hybridization despite incomplete lineage sorting. *Systematic Biology*, **60**, 138–149.

BIOGRAPHICAL SKETCH

Adam Payton completed high school in Sandpoint, Idaho and attended Western State College of Colorado in Gunnison, for his undergraduate education. It was his experiences living and recreating in the western US that sparked his interest in biology. This personal interest depended into a professional interested and as he pursued his bachelor's degree he began working for the Bureau of Land Management as a field botanist during the summers and analyzed data during the school semesters. His work with the Bureau of Land Management was pivotal in shaping his interests and desire to further his education in graduate school. In 2008 he graduated with a bachelor's degree in ecology and evolutionary biology. In 2009 he began his graduate career at the University of Florida under the supervision of Douglas Soltis, where he studied population genetics and the evolution of the southeastern endemic genus *Dicerandra*. Adam graduated with a Master of Science in 2012 and will continue his work with population genetics and evolution in the McDaniel lab at the University of Florida.